THE GPS RADIO OCCULTATION CONCEPT: THEORETICAL PERFORMANCE AND INITIAL RESULTS

Thesis by

E. Robert Kursinski

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1997

(March 17)

Acknowledgments

Willy Bertiger for the modifications of the GYPSY system for orbit determination and calibration of the occultation results.
Courtney Duncan for early interactions on combining microsatellites with GPS occultations
Chuck Greenhall for discussions and help in completing the fourier transform representation of the retrieval process.
Ken Hardy for his positive interactions particularly in the early going.
Roger Linfield for discussions on diffraction and nonideal behavior.
Stephen Leroy for idl idl idl, unix, the cubic interpolator and general help and discussions concerning atmospheric science.
Dan McCleese for helping to get my initial proposal to NASA funded and prodding me to do a far better error analysis than my initial effort.
Tom Meehan for many discussions regarding receiver development.
Brooks Thomas for discussions regarding receiver development and signal processing.
Tom Yunck for coming up with the GPS occultation concept.
Gunnar Lindal for suggesting I become involved with the GPS occultation concept.
Ab Davis for hiring me to do so.
Mike Exner who was very effective in getting the initial GPS occultation system off the ground after several years of wallowing at JPL and NASA.
Ramesh Kakar at NASA headquarters for funding most of the work presented here.
Bruce Murray for continually trying to kick me out the door (and finally succeeding).
Andy Ingersoll useful discussions regarding dynamics, waves and common sense and reducing my load to within the JPL annual tuition coverage.
Dewey (note spelling) for getting the division to allow me to come to Caltech while working half time at JPL, many useful discussions on a variety of topics, consenting to be the PI on the original President's fund proposal to analyze the GPS-MET data and a positive attitude toward completing the Ph.D. and doing it quickly.
Yuk Yung for teaching me radiative transfer, atmospheric chemistry and having many enjoyable discussions of a wide variety of science topics.
Larry Romans for developing various aspects of the occultation system at JPL including the fast and accurate raytracing, the data editor and parts of the calibration system.
George Hajj as the co-developer/conspirator of the JPL occultation system, many late nights at JPL and at home developing and debugging the system, automating the system.
Gail for putting up with me and the Caltech experience over the past 4.5 years.

Abstract

Implementation of the Global Positioning System (GPS) network of satellites and small, high performance instrumentation to receive GPS signals have created an opportunity for low cost, active remote sounding of Earth's atmosphere by radio occultation. The first goal of the present research is to estimate the spatial coverage, resolution and accuracy expected for atmospheric profiles derived from GPS occultations. Typically, vertical resolution ranges from 0.5 km in the lower troposphere to 1.4 km in the middle atmosphere. Useful profiles of refractivity should be derivable from ~60 km altitude to the surface with the exception of regions less than 250 m in vertical extent associated with high vertical humidity gradients. Above the 250 K altitude level in the troposphere, where the effects of water are negligible, sub-Kelvin temperature accuracy is predicted up to ~40 km depending on the phase of the solar cycle. Predicted accuracy of geopotential heights of constant pressure levels is ~10 m or better between 10 and 20 km altitudes. Deep in the warm troposphere the contribution of water to refractivity becomes sufficiently large for the accurate retrieval of water vapor given independent temperatures from weather analyses. We discuss several applications of the unique qualities of the technique including numerical weather prediction and long term monitoring of Earth's climate.

The second goal is to demonstrate some features using data from the prototype GPS-MET occultation investigation. We demonstrate ~1 km vertical resolution and temperature consistency with global weather analyses generally at the 0.5 to 1 K level. We discuss some initial observations of equatorial waves in the lower stratosphere and possible implications for exchange between the troposphere and stratosphere. During the June-July 1995 period, occultations typically extend to within 1 to 3 km of the surface and are used to derive a brief climatology of water vapor in latitude versus height, the first truly global view of water vapor at ~1 km vertical resolution. A low latitude bias structure

in the weather analyses is revealed centered near 2 km altitude where analysis humidities are larger than those derived from the occultations in the subtropics but smaller in the tropics apparently associated with a systematic error in the boundary layer height in the analyses.

Table of Contents

List of Figures

Acronyms

AHL	Andrews, Holton and Leovy
AMP	Atmospheric multipath
CC	Clausius-Clapeyron relation for saturation vapor pressure
DMSP	Defense Meteorological Satellite Program
ECE	Complete ECMWF structure for the June 21-23 and June 27-July 4, 1995, period
ECMWF	European Centre for Medium-Range Weather Forecasts
GEW	Water vapor climatology derived from GPS occultation refractivity and ECMWF analysis temperatures
GPS	Global Positioning System
IEC	Interpolated ECMWF structure interpolated to occultation locations
IGW	Inertio-gravity wave
ITCZ	Inter-Tropical Convergence Zone
JGOS	JPL GPS Occultation System
JJA	June-July-August
JPL	Jet Propulsion Laboratory
LEO	Low Earth Orbit
NH	Northern Hemisphere
P&O	Peixoto and Oort climatology
q	Specific humidity
RGW	Rossby-Gravity Wave
SB96	Soden and Bretherton (1996)
SH	Southern Hemisphere
SNR	Signal to Noise Ratio
SSM/I	Special Sensor Microwave Imager flown aboard DMSP polar orbiters
TOVS	TIROS Operational Vertical Sounder flown aboard NOAA polar orbiters
U	Relative humidity

1. Introduction

Given the non-linear character of Earth's climate system, an understanding of both its natural variability and its response to anthropogenically driven changes in radiative forcing requires long-term observations of temperature structure and water vapor concentrations in the troposphere and stratosphere. These parameters must be measured with sufficient accuracy, resolution, and spatial and temporal coverage to improve our knowledge of the fundamental processes involved in global change [Houghton *et al.*, 1990]. With the development of the Global Positioning System (GPS) satellite network, an opportunity now exists to acquire valuable new temperature, pressure and humidity data using radio occultation. Timely delivery of occultation data would represent an important, new contribution to the set of observations available for operational meteorology and, in particular, could play a significant role in global numerical weather prediction (NWP). Since the occultations are widely distributed, the data from even one receiver in low Earth orbit (LEO) represents a significant contribution to the global observing system particularly in remote areas where little high vertical resolution information presently exists. If many such receivers were in LEO, they would represent a very major source of information on the atmosphere's temperature and humidity fields [Eyre, 1994]. The unique combination of global coverage, high precision, high vertical resolution, long-term stability and all-weather viewing enabled by the long GPS wavelengths will complement and enhance data sets obtained by existing and planned *in situ* and remote sounding techniques.

In radio occultation, the radio path between an orbiting transmitter and an orbiting receiver traverses the Earth's atmosphere and is deflected primarily by the vertical gradient of atmospheric refractivity. At radio wavelengths, bending angle is derived from the Doppler shift in the refracted beam, given sufficiently accurate frequency and orbit geometry measurements. Refractivity can then be retrieved from the bending angle profile. From the mid-troposphere to the mesosphere, profiles of atmospheric bulk density, and

therefore pressure and temperature, can be derived from refractivity profiles. In the lower troposphere, water vapor also contributes significantly to refractivity, so that temperature and water cannot be retrieved independently using the present GPS frequencies. However, radio occultation bending angle and refractivity do provide important constraints on temperature and humidity profiles for numerical weather prediction models. At low latitudes in the lower troposphere, refractivity variations are dominated by water vapor and high vertical resolution humidity profiles can be retrieved, given independent temperature data.

Radio occultation measurements of the terrestrial atmosphere using signals transmitted by GPS satellites and received by one or more satellites operating in LEO can yield approximately 500 profiles per LEO orbiter daily. High vertical resolution is provided by limb sounding geometry and a point-like field-of-view, although in practice resolution is limited by diffraction effects to approximately 1 km at the limb or lowest height of the radio path. Precision is limited by the accuracy with which the Doppler frequency shift can be determined. Profile measurements are inherently stable in the long-term because frequency is derived either from stable atomic reference clocks or less stable clocks whose instabilities can be solved for and removed. This resolution and precision can be achieved with modest spacecraft attitude control capabilities, because Doppler shift is independent of spacecraft orientation and both the transmitter and receiver have broad fields-of-view. Finally, because GPS occultation measurements are made at radio wavelengths, they are insensitive to particulates which contaminate many tropospheric remote sounding techniques.

The application of the radio occultation technique to the remote sounding of the Earth's atmosphere was first suggested in the 1960's [Fishbach, 1965; Lusignan *et al.*, 1969]. However, to achieve the global coverage and accuracy needed to enhance the existing atmospheric data set, multiple orbiting transmitters and receivers were required at a

cost exceeding the level of interest. As a result, the practical and theoretical development of the radio occultation concept was directed towards the remote sounding of planetary atmospheres through a combination of efforts primarily at the Jet Propulsion Laboratory (JPL) and Stanford University. Radio occultation is now well established in NASA's planetary program due to its successful application to the atmospheres of Venus, Mars and the outer planets through missions including the Mariner, Pioneer and Voyager series [e.g., Fjeldbo *et al.*, 1971; Kliore and Patel, 1982; Lindal *et al.*, 1983, 1987; Hinson and Tyler, 1983; Tyler *et al.*, 1986, 1989; Lindal, 1991].

Utilization of the GPS network of satellites to characterize Earth's atmosphere by radio occultation was originally conceived for NASA's Earth Observing System in the form of the GPS Geoscience Instrument (GGI) [Yunck *et al.*, 1988]. I first heard about the concept from Jay Breidenthal in 1987 when he was working on a competing proposal to EOS. Subsequently, GPS radio occultation capabilities and applications have been discussed by a number of authors [Hardy *et al.*, 1992; Yuan *et al.*, 1993; Gorbunov and Sokolovskiy, 1993; Kursinski *et al.*, 1995] and have drawn attention in the general literature [Kerr, 1992].

A primary goal of this research has been to develop and establish first order estimates of the resolution and accuracy inherent to the profiles of refractivity, density, pressure and temperature derived from GPS occultations. This includes the development of a systematic and comprehensive examination of sources of error and their impact on retrieved quantities representing the first such analysis in the 30 year history of the spacecraft radio occultation technique. Such estimates define a baseline of expectation for developers to compare against and explain to those developing methods of assimilating occultation data into models what the sources and characteristics of error are so as to better optimize the assimilation process. Further, these estimates justify future GPS climate and

weather sensors and missions. Given the estimated resolution and accuracy, areas where these features can have a significant impact will be identified and described briefly.

While these theoretical results are extremely promising, real data must be examined to demonstrate and establish the capability. The first of a series of opportunities to demonstrate the GGI concept came with the GPS-MET investigation led by the Universities Corporation for Atmospheric Research (UCAR) with funding from NSF, NOAA and NASA [Ware *et al.*, 1996]. Future missions include second generation receivers, piggybacked Oersted and SUNSAT (launch late 97) and CHAMP and SAC-C which launch early 1999. A small 6 to 12 satellite constellation mission has also been proposed to NASA for launch in 1999. The GPS-MET mission placed a modified TurboRogue receiver developed by JPL and Allen Osborne Associates into a 70^{o} inclination, LEO, in April 1995 as an addition to the payload of the Microlab 1 spacecraft developed for NASA by the Orbital Sciences Corporation. I oversaw implementation of the GPS occultation processing system developed at JPL to support GPS-MET. The system which is described in more detail in section 6 was hung off the back end of the JPL GYPSY precision positioning system developed for TOPEX and other geodynamical applications. GYPSY was modified to remove the effects of geometry and clock errors and isolate just the additional phase shift due to the atmosphere. An occultation processing system was then developed by George Hajj and myself to take the atmospheric phase shift and geometry and produce profiles of refractivity, density, pressure and temperature and in the mid-to-lower troposphere, water vapor given independent knowledge of temperature.

Assessment of accuracy and resolution of the initial GPS-MET results has been underway since launch through comparison with radiosondes, satellite and weather model data. Comparison of occultation results with ECMWF weather model analyses appear to be the best method of evaluation because of 1.) quality-screened radiosonde and passive sounder data assimilated into the model, 2.) the translation of the very different radiosonde

and passive nadir sounding resolutions to the model resolution which is more consistent with that of the occultation observations and 3.) one of the best estimates of the global state of the atmosphere. Initial results indicate rms temperature differences w.r.t. ECMWF are roughly 1.5 K with mean differences of order 0.5 K in the northern hemisphere where model results are most accurate. Comparison of geopotential height derived from occultation data with ECMWF agree on average to ~20 m with a scatter of ~20 m rms [Leroy, 1996]. Initial water vapor comparisons indicate consistency at the 0.1 g/kg level and reveal systematic refractivity biases at the 1 to 2% level which may be associated with the model's representation of the Hadley circulation. Poorer agreement in the southern hemisphere presumably reflects the impact of sparse radiosonde data on ECMWF model performance. These comparisons establish preliminary upper limits on accuracy and resolution because of limitations in the ECMWF model, the assimilated data and the initial occultation results which should improve as occultation instrumentation and inversion techniques evolve.

The thesis is structured as follows. It begins with an overview of the occultation concept and technique (Sections 2.2 and 2.3). Sources of atmospheric attenuation which may limit the altitude range of the observations are then described together with their impact on measurement sensitivity (Section 2.4). The spatial resolution of the observations and derived quantities is then developed both in terms of the vertical weighting contribution to an individual measurement and the diffraction limit (Section 2.5). Section 3 presents a detailed examination of sources of error in derived refractivity, pressure and temperature. Section 4 discusses expected water vapor accuracy. Having established the vertical range and resolution and expected accuracy, areas in atmospheric science where the occultation observations will have a significant impact are identified and discussed briefly in Section 5. Section 6 presents initial temperature retrieval results from GPS-MET data including examination of some equatorial waves in the temperature profiles and some thoughts on their possible roles in exchange between the troposphere and stratosphere. Section 7

discusses initial water vapor results derived from GPS-MET data using temperature data from ECMWF analyses combined into a latitude versus height cross section of the atmospheric state for June 21-July 4, 1995. Results in Sections 6 and 7 include statistical comparisons against ECMWF global analyses in an attempt to sort out sources (and sinks) of error in the two data sets. Section 8 summarizes and concludes our sojourn.

2. Overview of the Occultation Technique

2.1 GPS Radio-Occultation Concept

The GPS consists of 24 satellites, distributed in 6 orbital planes about the globe. Each satellite orbit is circular, with an inclination of ~55°, a period of 12 hours and an altitude of 20,200 km. Although the GPS was designed and developed as a navigation aid by the U.S. Air Force, the civilian community has taken advantage of its capabilities for many additional high precision applications including now the study of the atmosphere. A schematic representation of atmospheric profiling by GPS radio occultation, using a receiver in LEO, is given in figure 2.1. From the standpoint of the receiver, an occultation occurs whenever a GPS satellite rises or sets and the ray path from its transmitter traverses the Earth's atmospheric limb. With 24 GPS satellites, a single GPS receiver in a near polar orbit at 800 km will observe over 500 occultations per day, distributed fairly uniformly about the globe (figure 2.2). This number can be doubled if the receiver is also capable of receiving signals from the Russian GLONASS constellation, a network similar in function and performance to the GPS. All of these occultations provide useful atmospheric data. The horizontal distribution of coverage depends chiefly on the inclination and period of the LEO satellite orbit. Better coverage is achieved by placing additional receivers in orbit. A set of 25 GPS/GLONASS receivers in diverse low Earth orbits would provide approximately 25,000 daily occultations yielding an average of one profile every 150 km square.

GPS-LEO occultation observations are made in a limb scanning mode, where vertical scanning is provided by relative motion between the GPS and LEO satellites (figure 2.3). In the geometrical optics approximation, a ray passing through the atmosphere is refracted according to Snell's law due to the vertical gradient of density and hence refractive index. The overall effect of the atmosphere can be characterized by a total

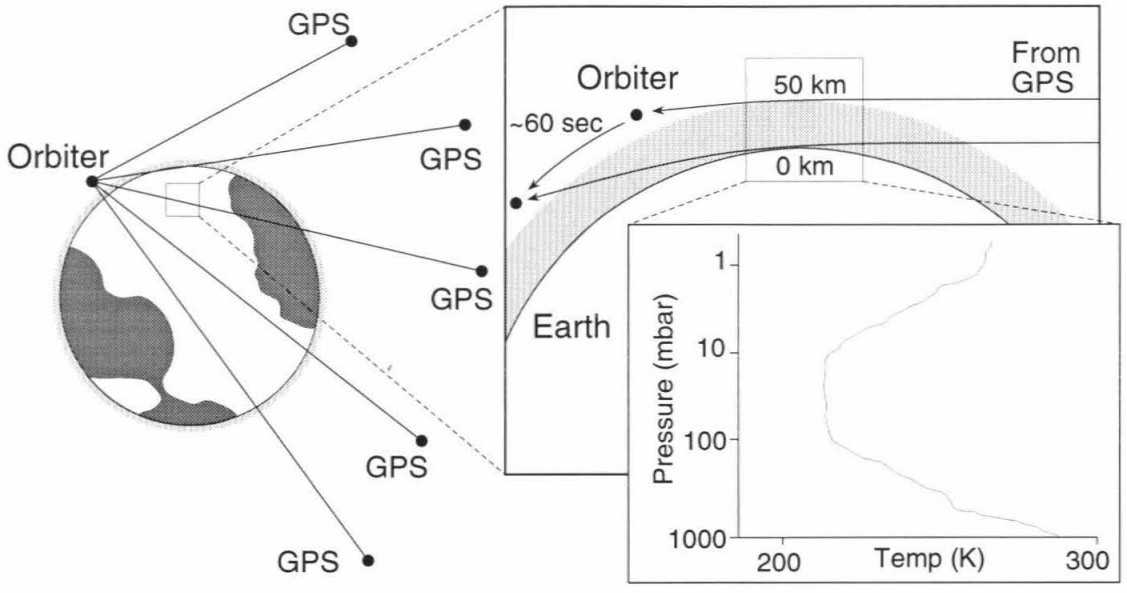

Figure 2.1 Schematic representation of atmospheric profiling using GPS satellite signal occultation.

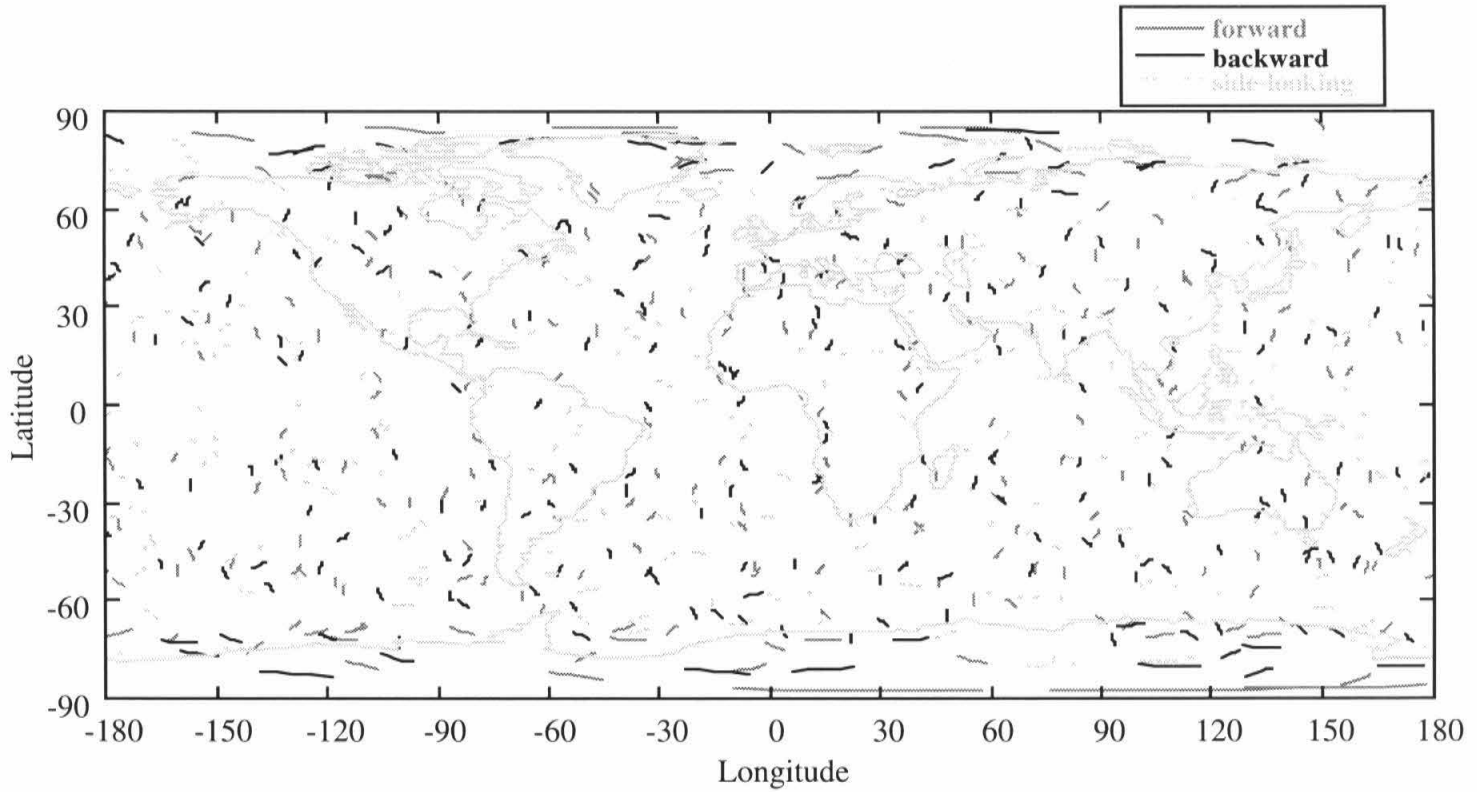

Figure 2.2 Distribution of occultations observed in one day by a receiver in low Earth orbit viewing 24 GPS satellites.

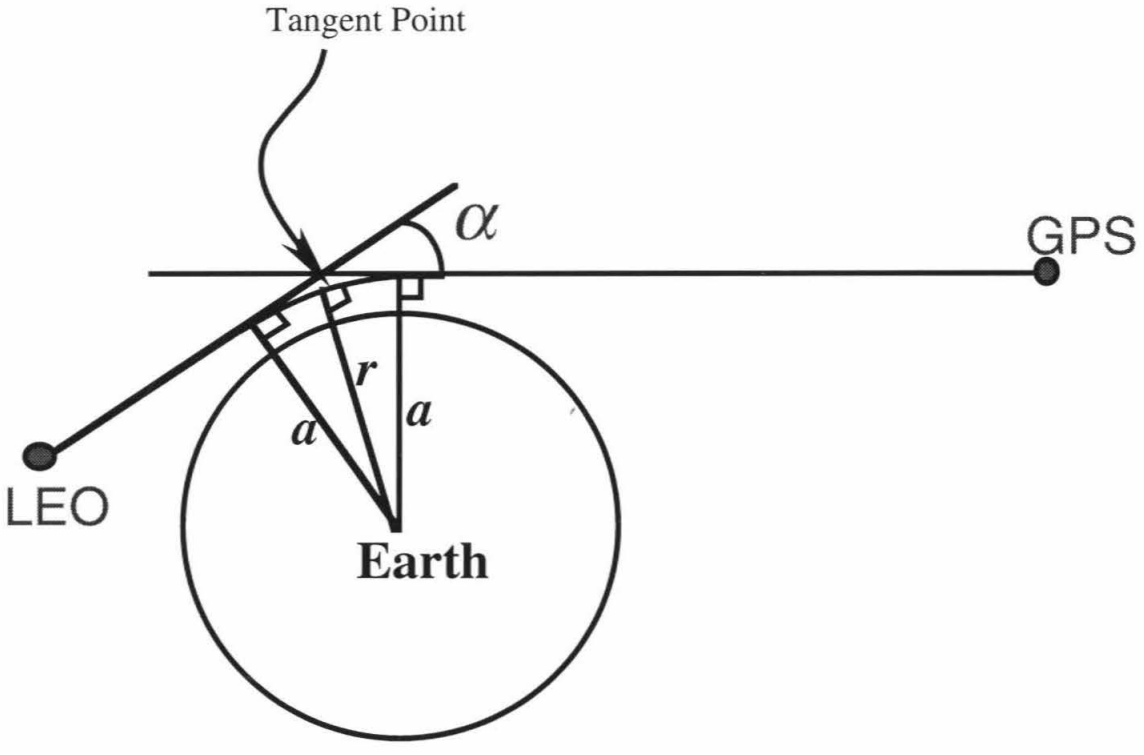

Figure 2.3 Instantaneous occultation geometry for GPS and low Earth orbiter (LEO) satellites.

bending angle, α, an asymptotic ray miss-distance or impact parameter, a, and a tangent radius, r_t defined in figure 2.3 depicting the instantaneous GPS-LEO occultation geometry. During an occultation, as the ray path descends or ascends through the atmosphere, the variation of α with r or a depends primarily on the vertical profile of atmospheric refractive index. This profile can be retrieved from measurements of α as a function of a during the occultation, subject to the assumption of local spherical symmetry (Section 2.2).

The time dependence of both α and a during an occultation can be derived from accurate measurements of the Doppler-shifted frequency of the transmitter signal at the receiver. Frequency is obtained from the derivative with respect to time of the signal phase measured by the receiver. Doppler shift is determined by the projection of spacecraft velocities onto the ray paths at the transmitter and receiver, so that atmospheric bending contributes to the measured Doppler shift. Data from several GPS transmitters can be used to establish the precise positions and velocities of the GPS and LEO satellites and to calculate the Doppler shift expected in the absence of bending. The atmospheric contribution to Doppler shift, derived by subtracting the expected shift from the measured shift, can then be combined with satellite position and velocity knowledge to give an estimate of α and a (Section 2.2).

The accurate retrieval of atmospheric refractivity profiles depends on accurate estimates of α and a which, in turn, depend on accurate measurements of transmitter signal phase and satellite position and velocity as a function of time. Each GPS satellite continuously transmits right-hand circularly polarized signals at two L-band frequencies, 1575.42 MHz (L1) and 1227.6 MHz (L2), corresponding to wavelengths of 19.0 cm and 24.4 cm respectively. These are modulated at a frequency of 10.23 MHz by a pseudo-random precision ranging code (P-code) [Spilker, 1978]. The L1 carrier signal is also modulated at a frequency of 1.023 MHz by a coarse-acquisition (C/A) ranging code. In

order to make full use of the atmospheric information contained in these signals, a GPS receiver operating in LEO must be able to measure the phase and amplitude of the L1 and L2 carrier signals. Comparison of L1 and L2 phase measurements forms the basis for separating atmospheric and ionospheric contributions to α (Section 3.7).

A small GPS receiver capable of making these measurements with the accuracy required for atmospheric occultation work has been developed at JPL. This receiver weighs approximately 2 kg, consumes less than 10 watts of power, tracks up to 8 GPS satellites simultaneously, and obtains L1 and L2 carrier phase measurements with a precision of order 0.1 mm for a one second integration period [Meehan *et al.*, 1992; Section 3.3].

2.2 Atmospheric Bending and Refractive Index Profile Retrieval - Theory

Atmospheric refraction.

In the geometric optics approximation to the propagation of electromagnetic radiation, the path of a ray passing through a region of varying refractive index is determined globally by Fermat's principle of least time and locally by Snell's law, so that:

$$\int n \, dl = \text{minimum} \qquad \text{(Fermat's Principle)} \qquad (2.2.1)$$

$$n \sin\phi = \text{constant} \qquad \text{(Snell's Law)} \qquad (2.2.2)$$

where n is the refractive index, ϕ is the angle between the gradient of refractive index and the ray path, dl is a path element, and the integral of eq. (2.2.1) is over the entire path. An expression for $d\alpha$, the incremental bending of the ray, is obtained by differentiating eq. (2.2.2).

$$d\alpha = d\phi = -\tan\phi \, dn/n \qquad (2.2.3)$$

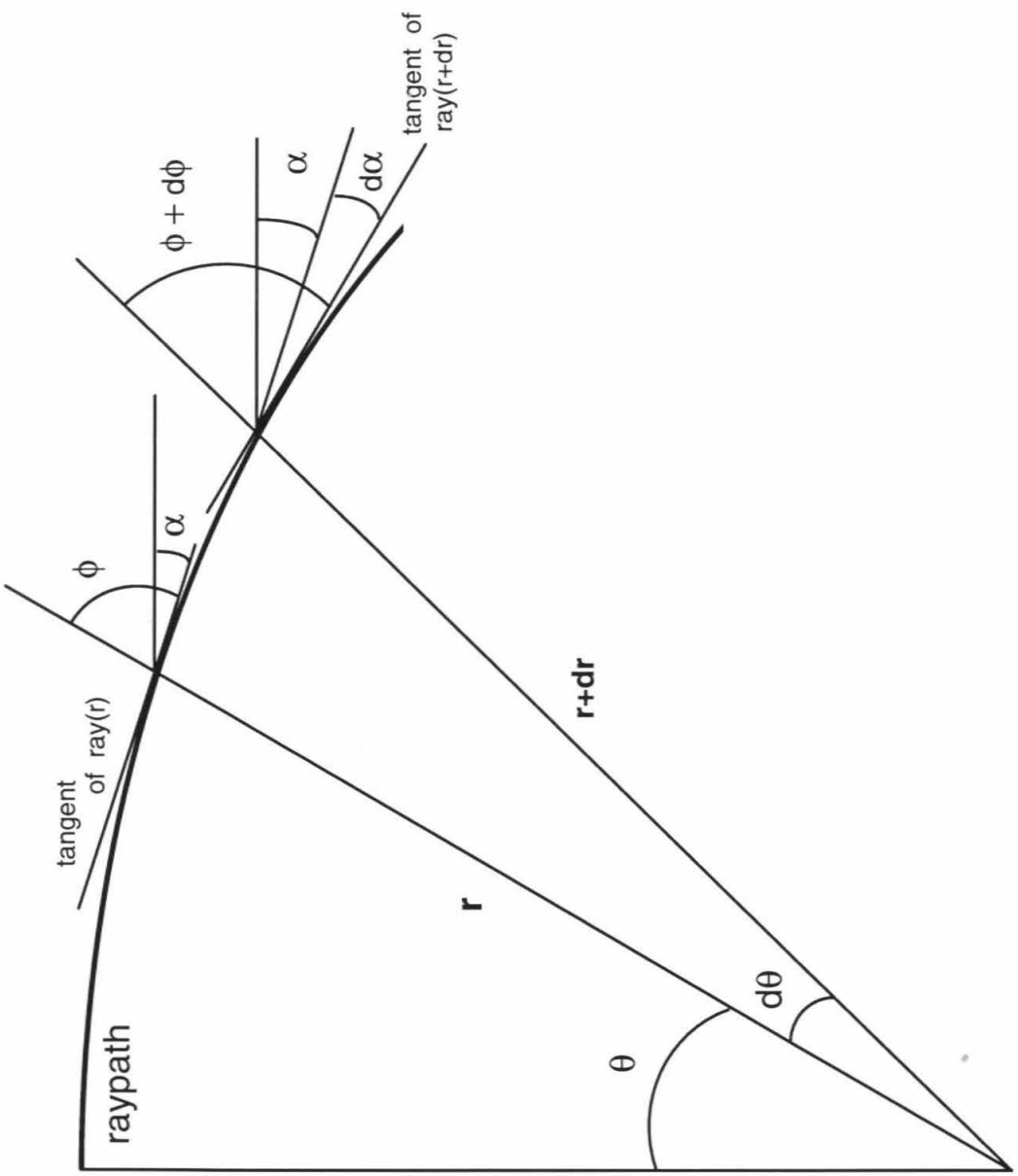

Figure 2.4 Incremental ray path bending, circular polar coordinate system.

Spherical symmetry.

In the most general case, it is not possible to retrieve the three-dimensional variation of n from measurements of α as a function of *a* during an occultation. However, the variation of n along a limb path in the Earth's atmosphere is dominated by the vertical density gradient so that, to the first order, the gradient of n is directed radially and the local refractive index field is spherically symmetrical. If spherical symmetry applies, the ray path lies in a plane, and $\alpha(a)$ can be inverted analytically to give n(r), where r is distance from the center of curvature.

Figure 2.4 defines a circular polar coordinate system lying in the ray path plane with its origin at the center of curvature of the contours of constant refractive index. In this case the change in the angle, ϕ, along the ray path has two contributions, the bending described by Snell's law, eq. (2.2.3), and the rotation of the coordinate system such that:

$$d\phi = d\phi_{snell} + \frac{\partial \phi}{\partial \theta} d\theta = d\alpha - d\theta \qquad (2.2.4)$$

where θ and ϕ are defined in figure 2.4. ϕ is the angle between the ray path and the radial vector which is coincident with the gradient of n. Combining eqs. (2.2.3) and (2.2.4), and remembering from figure 2.4 that $\tan \phi = r\, d\theta/dr$, gives:

$$dn/n + dr/r + d\phi/\tan\phi = 0$$

This can be integrated to yield:

$$n\, r \sin\phi = \text{constant} = a = n\, r_t \qquad (2.2.5)$$

Eq. (2.2.5) is known as Bouguer's rule (Born and Wolf, 1980). For a ray passing through a spherically symmetrical refractive index field, $n\, r \sin\phi$ is conserved. As $r \rightarrow \infty$, $n \rightarrow$ unity, so that the constant term in eq. (2.2.5) must equal *a*, the impact parameter,

defined in figure 2.3. Furthermore, at the ray tangent radius, r_t, also defined in figure 2.3, $\sin\phi$ = unity so that $a = nr_t$. The total refractive bending angle can be derived by substituting $\sin\phi = a/nr$ from eq. (2.2.5) into eq. (2.2.3) to give:

$$\alpha = \alpha(a) = 2\int_{r_t}^{\infty} d\alpha = 2a \int_{r_t}^{\infty} \frac{1}{\sqrt{r^2n^2 - a^2}} \frac{d\ln(n)}{dr} dr \qquad (2.2.6)$$

where the integral is over the portion of the atmosphere above r_t.

Eq. (2.2.6) addresses the forward calculation of $\alpha(a)$ given n(r). It can be inverted, using an Abelian transformation, to express n(r) in terms of α and a [Fjeldbo *et al.*, 1971].

$$n(r) = \mathrm{Exp}\left[\frac{1}{\pi}\int_{a_1}^{\infty} \frac{\alpha}{\sqrt{a^2 - a_1^2}} da\right] \qquad (2.2.7)$$

where $a_1 = nr$ is the impact parameter for the ray whose tangent radius is r. Given $\alpha(a)$, eq. (2.2.7) can be evaluated numerically.

Departures from spherical symmetry.

The derivation of n(r) from α(a) using the Abel transform inversion of eq. (2.2.7) is only valid for a spherically symmetrical atmosphere. However, the ellipsoidal shape of the Earth and horizontal gradients in atmospheric structure produce small departures from spherical symmetry in the refractive index field. Furthermore, the ray paths for a given occultation do not generally scan the atmosphere vertically nor are they coplanar. Measurements of α(a) will therefore be affected by tangential refractivity gradients and occultation geometry, and eq. (2.2.7) will introduce systematic errors into retrieved profiles of n(r) (Section 3.11).

Errors associated with Abel transform inversion depend on the accuracy with which the spherical symmetry approximation applies locally. First order errors due to the ellipsoid can be eliminated by selecting a radius of curvature appropriate to the latitude and orientation of the occultation measurement. For a given ray path, the majority of atmospheric bending occurs within a horizontal interval of 200-300 km centered on the tangent point (Section 2.5). The local spherical symmetry approximation will be accurate for a given occultation provided the horizontal scale of atmospheric structure is greater than this interval and the spread of tangent points lies well within it.

The results of the simulations described in Section 3 show that the Abel transform will generally be quite accurate as an initial retrieval scheme for GPS occultations of the Earth's atmosphere. Improved accuracy requires an inversion scheme that incorporates known local horizontal structure into the profile retrieval process. Such a scheme has been developed for the special case of elliptical symmetry to invert radio occultation observations of the atmospheric structure of the outer planets [Lindal, 1991]. More general approaches are evolving to take advantage of multiple occultations and a priori information [Gorbunov and Sokolovskiy, 1993; Eyre, 1994; Hajj *et al.*, 1994a].

The derivation of $\alpha(a)$ from Doppler shift measurements.

As noted in Section 2.1, the effect of atmospheric bending on a GPS signal received at an LEO receiver can be measured as an additional Doppler shift relative to that expected for a straight line signal path. Subject to the assumption of local spherical symmetry, Doppler shift can be combined with satellite position and velocity vector measurements to determine $\alpha(a)$.

Using the geometry and notation of figure 2.5, the Doppler shift, f_d, in the transmitter frequency, f_T, measured at the receiver is given by:

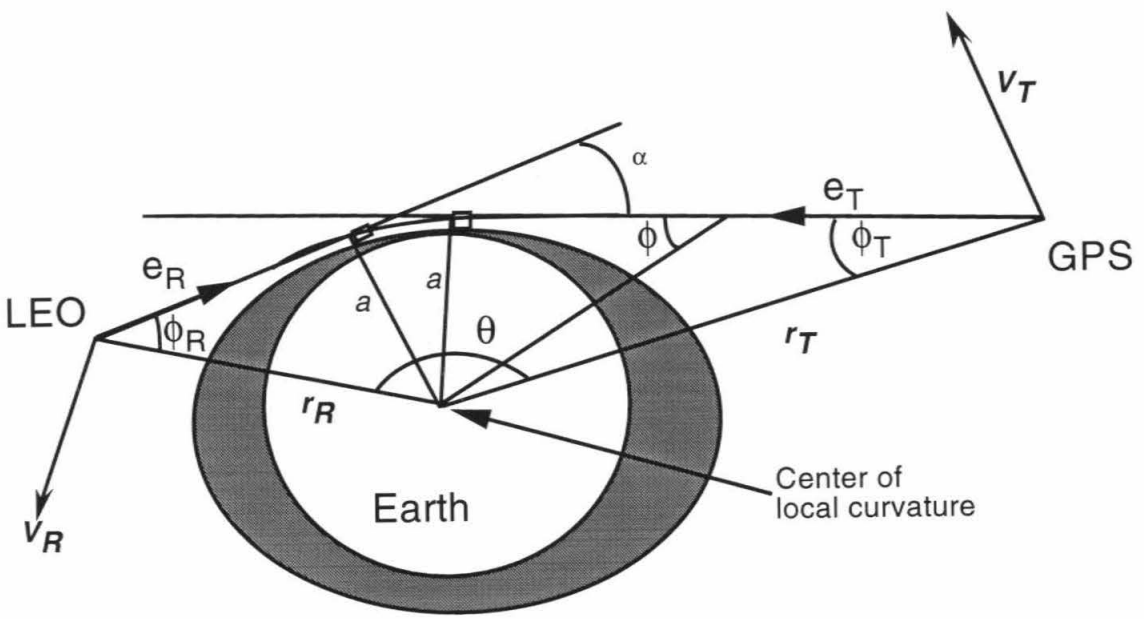

Figure 2.5 Occultation geometry and variables for the derivation of bending angle and asymptotic miss distance from Doppler shift, spacecraft position and spacecraft velocity measurement.

$$f_d = f_T \left[\frac{\left(\mathbf{V_T} \cdot \hat{\mathbf{e}}_T + \mathbf{V_R} \cdot \hat{\mathbf{e}}_R \right)}{c} + \frac{\left(\Phi_T - \Phi R \right)}{c^2} + \frac{\left(V_R{}^2 - V_T{}^2 \right)}{2c^2} + O\left(\frac{V}{c} \right)^3 \right] \qquad (2.2.8)$$

where $\mathbf{V_R}$ and $\mathbf{V_T}$ are receiver and transmitter velocity vectors, $\hat{\mathbf{e}}_R$ and $\hat{\mathbf{e}}_T$ are unit vectors representing the direction of the ray path at the receiver and transmitter, Φ_R and Φ_T are gravitational potentials at the receiver and transmitter, and c is the velocity of light. The first term represents the Doppler shift produced by the projection of the spacecraft velocities on to the ray paths, the second term is a general relativity correction for the different spacecraft gravitational potentials, and the third and fourth terms represent higher order relativistic corrections for spacecraft velocity.

The relativistic terms in eq. (2.2.8) are eliminated based on knowledge of orbital geometry and Earth's gravity field leaving the first term on the RHS of eq. (2.2.8) ($\equiv f_{nd}$). Resolving the velocity vectors into radial and tangential components gives the following expression for f_{nd}.

$$f_{nd} = - f_T / c \left(V_T^r \cos \phi_T + V_T^\theta \sin \phi_T + V_R^r \cos \phi_R - V_R^\theta \sin \phi_R \right) \qquad (2.2.9)$$

where the superscripts r and θ refer to radial and tangential velocity components, and ϕ_T and ϕ_R are the angles between the ray path and the spacecraft position vector at the transmitter and receiver (figure 2.5). From Bouguer's rule (eq. 2.2.5) and the geometry of figure 2.5, we have:

$$r_T \sin \phi_T = r_R \sin \phi_R = a \qquad (2.2.10)$$

$$\alpha = \phi_T + \phi_R + \theta - \pi \qquad (2.2.11)$$

where r_R and r_T are the distances of the receiver and transmitter from the center of mass of the Earth, and θ is the angle between the transmitter and receiver position vectors.

Given measurements of transmitter and receiver position and velocity vectors, transmitter frequency, and Doppler shift, $\alpha(a)$ can be derived iteratively from eqs. (2.2.9), (2.2.10), and (2.2.11) by eliminating ϕ_T and ϕ_R. However, because Bouguer's rule is invoked, this derivation is only exact under the assumption of local spherical symmetry. Errors in α and a due to this assumption have a small effect on retrieved refractive index profiles compared with the use of the Abel transform inversion. Both effects are included in the analysis of Section 3.

2.3 Derivation of Atmospheric Properties from Refractive Index Profiles

The dependence of refractive index on atmospheric properties.

In order to derive atmospheric properties from retrieved profiles of the real component of atmospheric refractive index, n, it is necessary to know how these properties influence n. At microwave wavelengths in Earth's atmosphere, n contains contributions from four main sources. These are, in order of importance, the dry neutral atmosphere, water vapor, free electrons in the ionosphere, and particulates (primarily liquid water). Their effects are given to first order by:

$$N = (n-1) \times 10^6 = 77.6 \frac{P}{T} + 3.73 \times 10^5 \frac{P_w}{T^2} - 4.03 \times 10^7 \frac{n_e}{f^2} + 1.4\,W \quad (2.3.1)$$

where N is refractivity, P is atmospheric pressure in millibars, T is atmospheric temperature in Kelvin, P_w is water vapor partial pressure in millibars, n_e is electron number density per cubic meter, f is transmitter frequency in Hertz, and W is liquid water content in grams per cubic meter. Throughout this paper, the four refractivity terms in eq. (2.3.1) are referred to as the dry, moist, ionospheric, and scattering terms.

The dry refractivity term is due to the polarizability of molecules in the atmosphere; that is, the ability of an incident electric field to induce an electric dipole in the molecules. Dry refractivity is proportional to molecular number density and is dominant below ~60-90 km. The moist refractivity term is due primarily to the large permanent dipole moment of water vapor and becomes significant in the lower troposphere. The form of the dry and moist terms in eq. (2.3.1) was given originally by Smith and Weintraub [1953] and their accuracy is discussed in Section 3.9. The ionospheric term in eq. (2.3.1) which is a first order approximation to the Appleton-Hartree eq. [Papas, 1965] is mainly due to free electrons in the ionosphere and becomes important above ~60-90 km. The scattering term given in eq. (2.3.1) is due to liquid water droplets suspended in the atmosphere. For ice crystals the scale factor is roughly 0.6 rather than 1.4 (Appendix A.1). For realistic suspensions of water or ice, the scattering term is small in comparison with the other terms and is therefore neglected (Appendix A.1).

A representative example of atmospheric bending as a function of ray path tangent height is shown in figure 2.6. The first curve is bending due to the vertical refractivity structure of the U.S. Standard Atmosphere. Bending decreases rapidly with tangent altitude, from 1 to 2 degrees near the surface to ~2×10^{-5} degrees at 80 km where the bending roughly equals the measurement precision. The small size of measurement error relative to bending angle magnitudes in the lower atmosphere demonstrates the extreme precision with which atmospheric delay can be determined.

The second and third curves in figure 2.6 include the bending contribution of the ionosphere based on Chapman layer representations of the E- and F-region electron density profiles. The second and third curves include the ionospheric bending under nighttime and daytime conditions respectively at the maximum of the solar cycle. Ionospheric bending depends weakly on altitude for tangent altitudes below 80 km because ionospheric refractivity gradients are greatest above 90 km, significantly above the ray path

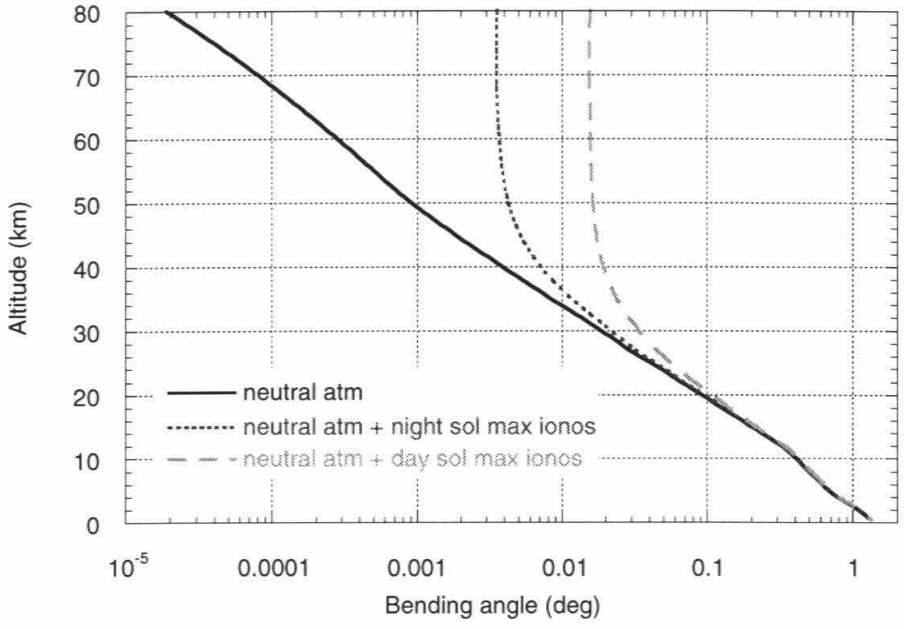

Figure 2.6. Atmospheric and ionospheric bending for limb ray paths plotted as a function of ray path tangent height. Solid line: bending calculated for U.S. Standard Atmosphere (Champion et al., 1985). Long dashed line: bending calculated for U.S. standard atmosphere plus typical mid-latitude ionosphere during daytime, solar maximum conditions. Short dashed line: bending calculated for U.S. standard atmosphere plus typical mid-latitude ionosphere during nighttime, solar maximum conditions.

tangent point. Under daytime conditions during the maximum of the solar cycle, ionospheric bending exceeds atmospheric bending above 30 km and therefore must be well calibrated to derive accurate profiles in the middle atmosphere.

It is clear from eq. (2.3.1) that there is ambiguity between the contributions of the dry, moist, and ionospheric terms to refractivity. Separation and removal of the dispersive ionospheric contribution in eq. (2.3.1) using the dual GPS transmitter frequencies L1 and L2 is described in detail in Section 3.7.

The derivation of pressure and temperature.

After correcting for the ionosphere, only the dry and moist atmospheric contributions to refractivity remain. In regions where the atmosphere is drier than a volume mixing ratio of 10^{-4}, the moist term can be neglected and eq. (2.3.1) reduces to $N = 77.6\ P/T$. Combining this with the equation of state (the ideal gas law) gives density as a function of refractivity:

$$\rho(r) = \frac{N(r)\, m}{b_1 R} = \frac{P(r)\, m}{T(r)\, R} \tag{2.3.2}$$

where R is the gas constant, m is the mean molecular mass of dry air, ρ (r) is the air density in kg/m^{-3}, and b_1= 77.6 N-unit K/mbar. The small effects of non-ideal behavior, ignored in eq. (2.3.2) for simplicity and clarity, are discussed in Section 3.9. In dry air below the homopause (~100 km altitude), m is a known constant, so that pressure can be obtained from density by integrating the equation of hydrostatic equilibrium:

$$\partial P/\partial r = -g\rho \tag{2.3.3}$$

where g is the acceleration due to gravity. The integration is started at a pressure level derived from a temperature guess high in the atmosphere. Initial pressure errors in the temperature guess decrease rapidly as the integration moves deeper into the atmosphere

(Section 3.10), and accurate temperature profiles can be derived from profiles of ρ and P throughout the stratosphere down to the mid-troposphere and below depending on latitude.

The presence of significant tropospheric water vapor complicates the interpretation of refractivity. However, in the colder tropospheric regions, where water vapor concentrations are small, accurate profiles of density, pressure and temperature can be retrieved given an estimate of humidity. From eq. (2.3.1) and the gas law:

$$\rho = \frac{m}{b_1 R}\left(N - \frac{b_2 P_w}{T^2}\right) \tag{2.3.4}$$

where $b_2 = 3.73\text{x}10^5$ K^2/mbar and m is now the mean molecular mass of moist air. Pressure and temperature profiles are derived from the density profile using eq. (2.3.3) and the new temperature and mean molecular mass profiles are then used in eq. (2.3.4) to update the estimate of ρ. This process is repeated iteratively a few times until the derived temperature profile converges. Water vapor concentration estimates can be obtained from independent observations, weather analyses or climatology. Sensitivities to water vapor uncertainties are characterized in Section 3.13.

The derivation of water vapor.

Deeper in the troposphere, water vapor concentrations increase contributing up to 30% to net refractivity near the surface in the tropics where they can dominate the vertical refractivity gradients and bending. Eq. (2.3.1) can be rearranged to solve for water vapor, so that:

$$P_w = \frac{N T^2 - b_1 P T}{b_2} \tag{2.3.5}$$

The recovery of water vapor from measurements of N, using eq. (2.3.5) requires knowledge of temperature derived from climatology, independent observations or weather analyses. The vertical scales of water vapor variability, generally much smaller than those

of the dry atmosphere, may also be used to constrain the wet and dry contributions to refractivity. In practice, T(r) is assumed and $P_w(r)$ and P(r) are derived iteratively from eqs. (2.3.3), (2.3.4), and (2.3.5).

2.4 Tracking Signals in the Lower Troposphere

In the lower troposphere, increasing pressure and humidity reduce the strength of the occulted signals while increasing their variability, a combination which could limit the altitude range of the occultation coverage. While these problems have, in fact, limited initial GPS-MET coverage in the lower troposphere, they are surmountable with improvements in instrumentation and inversion techniques within the resolution constraints defined in Section 2.5. In order to acquire an occulted signal, the signal-to-noise intensity ratio (SNR) at the receiver must exceed a critical value of about 10. If SNR falls below this value, phase lock and the signal are lost. SNR is determined by GPS transmitter power, receiver noise performance, observational geometry, and signal attenuation due to the atmosphere. In the lower troposphere, signal acquisition can be further complicated by atmospheric multipath where several weak occulted signals exist simultaneously. The effects of atmospheric multipath and atmospheric attenuation and the SNR performance expected from a GPS receiver in low Earth orbit are described below.

Atmospheric attenuation.

Attenuation mechanisms affecting propagation of GPS signals across the atmospheric limb are absorption by atmospheric constituents, extinction by particle scattering, and defocusing. Atmospheric absorption at L-band wavelengths is dominated by the far wings of pressure-broadened molecular oxygen and water vapor lines, with the contribution of oxygen being at least an order of magnitude greater than that of water vapor. For a limb path, net L-band optical depth varies roughly as the square of pressure at

the tangent height. Calculations using absorption spectra from Chahine *et al.* (1983) show that absorption by molecular oxygen reduces GPS signal intensities by a factor of 2, for a limb path with a tangent height near the surface. Extinction caused by realistic suspensions of water droplets or ice particles in the atmosphere is negligible at L-band wavelengths (Appendix A.1).

The dominant attenuation mechanism is defocusing which occurs when rapidly changing vertical refractivity gradients cause adjacent, nearly-parallel rays entering the atmosphere to bend differentially and diverge, reducing (or sometimes enhancing) signal intensity at the receiver. The intensity of an occulted signal after defocusing can be represented by the approximate expression (Haugstad, 1978):

$$I = I_0 M, \text{ where } M = (1 + D\,\partial\alpha/\partial a)^{-1} \approx (1 - D\alpha/H_\alpha)^{-1}, \qquad (2.4.1)$$

I is the observed signal intensity, I_0 is the signal intensity observed in the absence of defocusing, D is the distance from the receiver to the limb, α is bending angle, and a is the impact parameter (see figure 2.3). H_α is the local bending angle scale height defined by the approximation $\alpha = \alpha_0 e^{-(a-a_0)/H_\alpha}$, where $a \approx a_0$. The variation of the defocusing factor M with limb ray path tangent altitude, calculated for a vertical refractivity profile corresponding to a radiosonde ascent from Hilo, Hawaii, in July 1991, assuming spherical symmetry, is illustrated in figure 2.7a. Three general features of the curve should be noted. Firstly, a gradual, overall reduction in M with decreasing altitude is caused by the exponential increase of bending angle with decreasing altitude (figure 2.6, eq. (2.4.1)). Secondly, rapid vertical variations in M are associated with sharp vertical changes in the refractivity gradient caused primarily by vertical water vapor mixing ratio fluctuations. These variations provide a sensitive measure of the local refractivity scale height and humidity fluctuations. Finally, signals with ray path tangent heights that fall within certain vertical intervals are attenuated completely. In figure 2.7a, attenuation is complete at the top of the marine boundary layer where extremely large vertical refractivity gradients are

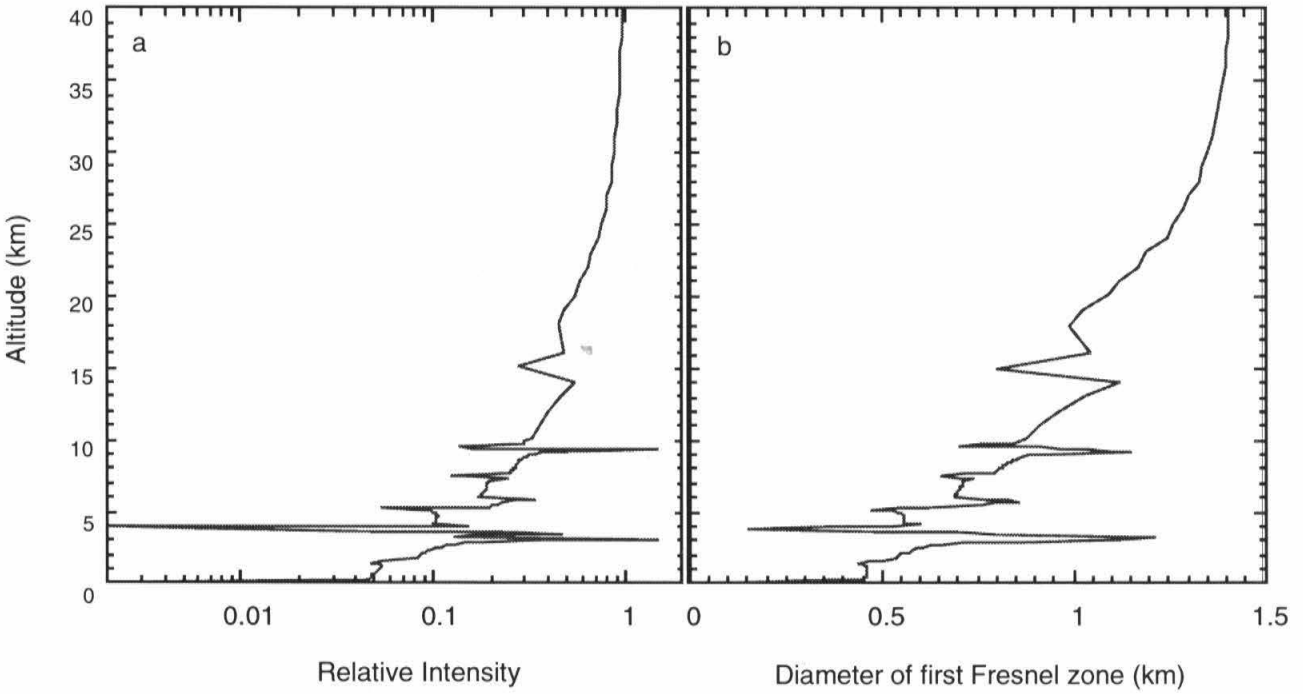

Figure 2.7 a. Variation of defocusing factor, M, with altitude [eq. (2.4.1)]. M is calculated for a refractivity profile derived from radiosonde temperature and humidity profiles, assuming spherical symmetry. b. The diameter of the first Fresnel zone defined by eq.'s (2.5.2) and (2.5.3) plotted as a function of ray path tangent height for a refractivity profile derived from radiosonde temperature and humidity profiles, assuming spherical symmetry. Radiosonde ascent 12:00 UT, 11 July 1991, Hilo, Hawaii.

produced by sudden changes in temperature and humidity associated with the transition between the boundary layer and overlying atmosphere.

Critical refractivity gradients

Complete attenuation occurs at ray path tangent altitudes where the vertical refractivity gradient becomes so large that the radius of curvature of the ray is smaller than the radius of curvature of the atmosphere causing the ray to curve down and intersect the surface. This happens when the refractivity gradient, $dN/dr \leq -10^6/R_c$, where R_c is the radius of curvature of the atmosphere; for a mean value of R_c, $dN/dr \sim -0.16$ m^{-1}. The vertical atmospheric gradients required to satisfy this inequality can be investigated by differentiating the dry and moist refractivity terms of eq. (2.3.1) to give:

$$\frac{\partial N}{\partial r} = -\frac{b_1 P}{H_P T} - \left(\frac{b_1 P}{T^2} + \frac{2 b_2 P_W}{T^3}\right)\frac{\partial T}{\partial r} + \frac{b_2}{T^2}\frac{\partial P_W}{\partial r} \qquad (2.4.2)$$

where H_P is the pressure scale height. The three terms in eq. (2.4.2) represent the contributions of the vertical pressure, temperature, and water vapor mixing ratio gradients to dN/dr. For T= 280 K, P = 850 mbar and P_w = 9 mbar (90% relative humidity), conditions representative of the top of the marine boundary layer, the first term in eq. (2.4.2) is -0.03 m^{-1}, demonstrating that realistic pressure gradients alone are incapable of generating critical refractivity gradients. The other terms, considered separately, require temperature gradients of $\geq$ +140 K/km or water vapor partial pressure gradients of $\leq$ -34 mbar/km to produce critical gradients. The mean vertical extent of the boundary layer transition region observed during the Atlantic Tradewind Experiment (ATEX) in 1969 was ~200 m [Augstein *et al.*, 1974]. The mean temperature and water vapor partial pressure immediately above this region are 5 K higher and 6 mbar lower than below it, so that contributions to dN/dr by the temperature and moisture gradients are about -0.03 m^{-1} and -0.12 m^{-1} respectively. While the mean temperature gradient contribution is considerably less than the critical value, the mean water vapor gradient is 75% of the critical value and

their sum nearly equals the critical value indicating a significant fraction of occultations will experience critical refraction in regions where a sharp inversion caps the marine boundary layer.

Region where critical gradients can be found

The maximum altitude of critical gradients can be constrained using a hypothetical layer with completely dry air above and 100% saturated air below and a vertical thickness of ~100 m. Notice that the layer must extend 200 km or more horizontally to produce the critical bending. The saturation vapor pressure yielding the critical refraction gradient of -34 mb/km is 3.4 mb which corresponds to a temperature of -5°C. The climatological -5°C temperature contour which envelops the region of possible critical refraction has a maximum altitude near 5 km at low latitudes and intercepts the surface near 60° latitude [Peixoto and Oort, 1992].

One can argue that still lower saturation vapor pressures will produce critical gradients by simply narrowing the vertical extent of the layer. Decreasing the layer thickness to 10m would raise the altitude at which critical refraction could occur by about 5 km. However, since most of the received signal is contributed by the region within the first Fresnel zone, the vertical extent of the critical refraction layer must span a significant fraction of the diameter of the first Fresnel zone to produce extinction. Otherwise, the layer will only remove a portion of the signal flux causing scintillations plus some reduction in intensity but not extinction. The relatively large Fresnel zone resulting from a vantage point in low Earth orbit reduces the occurrence of extinction by smoothing over small scale radio holes which would plague observations made closer to the holes and/or at shorter wavelengths. Note that the extreme gradients at the top of the tradewind inversion decrease the Fresnel zone diameter to ~150 m (figure 2.7b). Therefore, 100 m is a representative minimum interval thickness capable of producing extinction in spaceborne GPS observations and the

-5°C temperature contour defines the upper altitude limit of critical refraction for these observations.

This conclusion is confirmed by the initial GPS-MET results which even with a low antenna gain and receiver minimally modified from a ground-based geodetic receiver, have demonstrated an ability to track routinely down to ~5 km altitude. Initial GPS-MET signal tracking at higher latitudes systematically extends to lower altitudes, and frequent loss of lock near 5 km altitude at low latitudes in the initial GPS-MET occultations may coincide with the critical refraction estimate but more probably reflects the low SNR of the prototype receiver. Further, with some simple tracking modifications, the occulted signal can be tracked routinely to 1-3 km altitude as we discuss in Section 7. Future receivers launching over the next few years with higher antenna gain and better signal tracking will improve upon the GPS-MET limitations.

Vertical extent of intervals of critical refraction

The vertical extent of the critical refraction interval can be derived using Bouguer's rule where $n\ r\ \sin\phi = n_t\ r_t$ along the ray path (eq. (2.2.5)), where the subscript "t" refers to values at the ray path tangent where ϕ is $\pi/2$. For a ray that leaves the atmosphere, ϕ must be less than or equal to $\pi/2$ away from the tangent point, such that $nr \geq n_t r_t$. It follows that if the radii and refractive indices of the tangent points for the extreme ray paths that bound the interval of complete attenuation are r_1, n_1 above it and r_0, n_0 below it, then

$$r_1 n_1 = r_0 n_0 \tag{2.4.3}$$

The vertical extent of the interval is then given by:

$$\Delta r = r_1 - r_t = r_1\ (1 - n_1/n_t) = r_1\ \Delta n\ /\ n_t \approx r_1 \Delta n \tag{2.4.4}$$

where $\Delta n = n_t - n_1$. The large change in index of refraction, Δn, primarily occurs across the boundary layer transition region which has a vertical extent, Δz. Therefore,

$$\Delta r \approx r_1 \, (dn/dz)_{tr} \, \Delta z \; = \; r_1 \, \eta / R_c \, \Delta z \; \approx \; \eta \, \Delta z \qquad (2.4.5)$$

where η is the ratio between the large index of refraction gradient across the transition region, $(dn/dz)_{tr}$, and the critical refractivity gradient, $1/R_c$. The vertical interval of critical refraction therefore depends on both the magnitude and vertical extent of the large refractivity gradient. For example, a very large vertical moisture gradient observed by a radiosonde over Hilo, Hawaii (10/1/93 00:00UT), exceeded the critical gradient by a factor of 3 over a vertical interval, $\Delta z \sim 80$ m. Based upon eq. (2.4.5), an occultation of this region would have observed a vertical interval of critical refraction 240 m in extent.

Derivation of refractivity within and below radio holes

The existence of a critical refractivity interval raises the question of recovery of refractivity at altitudes within and below the interval. As the tangent height of a ray descends toward an interval where critical refraction occurs, bending increases and at the point of critical refraction, the bending angle becomes infinite. Analogously, the bending angle of a ray path whose tangent height lies at the bottom edge of the critical refraction interval is also infinite and bending immediately below the interval is finite. Since $r_1 n_1 = r_0 n_0$, $a_1 = a_0$ and the interval of critical refraction therefore collapses to a *singularity* in the α versus a profile. Accurate recovery of refractivity structure below the interval of critical refraction via the Abel transform requires first that the existence of the singularity be identified in the data using the characteristic disappearance of the signal and second that the numerical implementation of the Abel transform defined in eq. (2.2.2) be capable of integrating across the singularity.

The numerical implementation of the Abel transform in the JPL GPS occultation system (JGOS) does not at the present time meet these requirements. The JGOS numerical implementation assumes $d\alpha/da$ is continuous which is untrue at the singularity. Forcing a continuous derivative across the singularity results in a large error which is why most of

the simulation results in the horizontal refractivity do not extend below 2 km altitude (Sections 3.11 and 4.2). Finer and finer vertical sampling (which is not physically realizable without a diffraction correction) does reduce the magnitude of the refractivity error indicating a numerical integrator across the singularity will improve the results.

The Abel transform does not recover refractivity structure within the interval. Recovery of refractivity within the interval must be done via interpolation if the Abel transform approach is used. There are several constraints which can be applied to both the singularity and interpolation across the critical refraction interval. Constraints include n and r derived at the top of the interval via the transform of data taken above the interval plus the requirement that $\partial N/\partial r$ at the top of the interval equal the critical gradient, $-10^6 R_c$. At the bottom of the interval, constraints include the Abel transform results integrated across the singularity plus $a_0 = r_0 n_0 = a_1$. The vertical gradient of refractivity is constrained that its integral over the interval must equal $n_1 - n_0$ together with requirements of continuity at the boundaries and smoothness over the interval and that its structure explains the vertical extent of critical refraction. Given the small vertical extent of critical refraction intervals, interpolation across the interval should be accurate limited primarily by the accuracy of the Abel integration across the singularity.

Errors in derived refractivity due to misrepresentation of the singularity will be largest at the bottom of the critical refraction interval and will decrease rapidly at lower altitudes because the contribution of the critical refraction interval to total bending decreases dramatically as the tangent height descends (figure 2.8). The magnitude of the refractivity error is limited by the fact that the retrieved heights must be monotonic. A large refractivity error will result in a radius which is out of sequence and clearly in error.

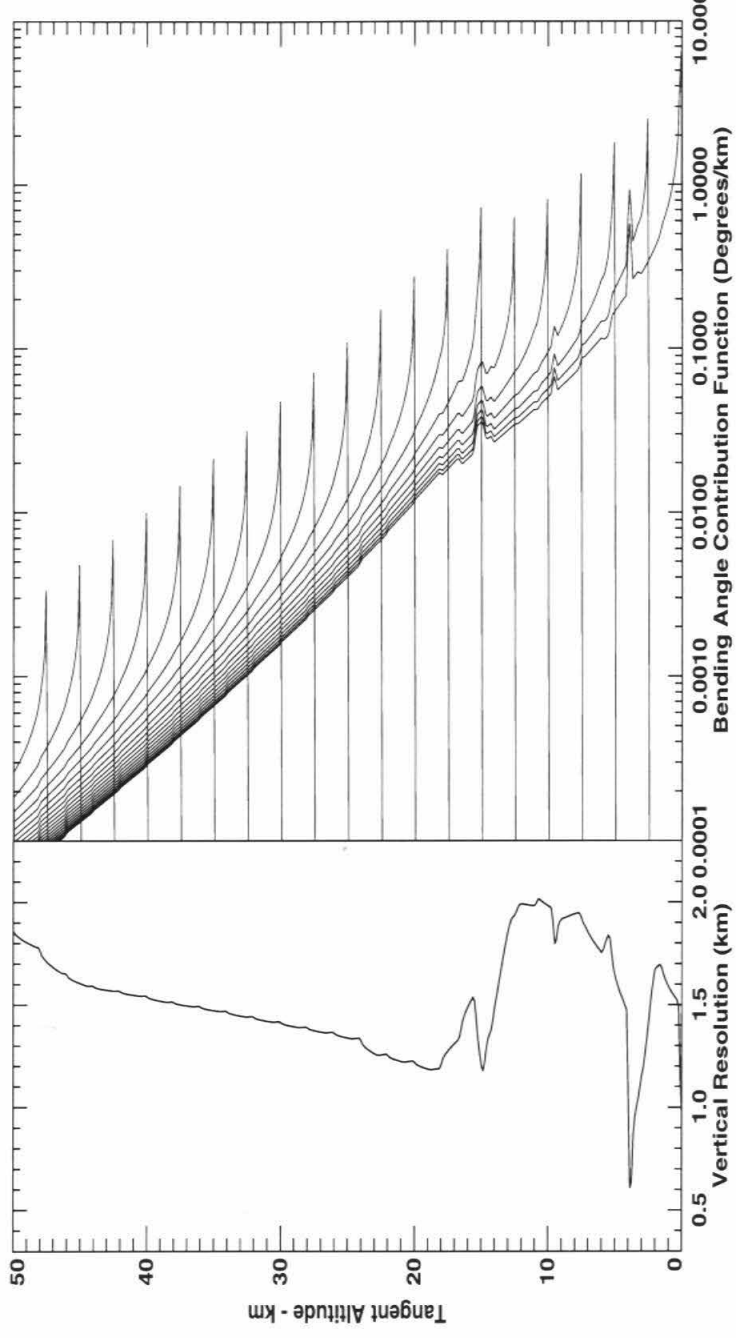

Figure 2.8. Bending contribution functions (solid lines) and vertical resolution profile (dotted line) for individual radio occultation bending angle measurements plotted as a function of altitude. Resolution is defined here as the vertical interval between the ray path tangent height and the height at which half of the total bending has been accumulated. These curves are calculated for a refractivity profile derived from radiosonde temperature and humidity profiles, assuming spherical symmetry. Radiosonde ascent 12:00 UT, 11 July 1991, Hilo, Hawaii.

Signal-to-noise performance.

The GPS network transmits relatively powerful signals designed for simple, hand-held receivers. In the absence of atmospheric attenuation, the prototype GPS-MET receiver with a small hemispherical field-of-view antenna in low Earth orbit has obtained SNR's of ~1-2x10^3 over the 20 msec integration period used at present for the radio-occultation observations. Future instrumentation will improve upon this SNR by perhaps an order of magnitude.

Allowing for the molecular oxygen attenuation factor of 2 near the surface, and the SNR of 10 required for receiver phase lock, attenuation of the amplitude by a factor of at least 100 (20 dB) in intensity due to atmospheric defocusing can be tolerated before loss of signal occurs. This margin is sufficient for signals to be tracked throughout the atmosphere with the exception of regions of very sharp refractivity gradients (figure 2.7a). GPS signal paths will in fact probe the entire atmosphere apart from occasional small critical refraction intervals just discussed. Implications of SNR on the accuracy of derived refractivity, pressure and temperature are discussed in Section 3.3.

Encryption of GPS signals for military purposes is also an issue. Two types of encryption exist: Selective Availability (SA), which alters the apparent GPS satellite positions, and Anti-Spoofing (AS), which eliminates jamming of the GPS receivers by false signals. Methods for eliminating SA are discussed in Section 3.8. The AS encryption affects the L1 and L2 P-code signals (Section 2.1) and therefore calibration of the ionosphere (Section 3.7). A classified receiver which knows the encryption can recover the full precision of the signals. An alternative method is a non-classified cross-correlation approach which produces little degradation when SNR's are high (Section 3.3). Unfortunately, the SNR of the cross-correlated signal decreases as the square of the atmospheric attenuation such that the factor of 40 attenuation due to molecular oxygen absorption and defocusing near the surface (figure 2.7a) will decrease the SNR of the

cross-correlated signal by 1600 (32 dB). The integration time used when the ray path tangent point is in the lower troposphere is of order 1 to 2 seconds as defined by the time for the ray path to descend or ascend a Fresnel diameter (Sections 2.5 and 3.11). Therefore, given the SNR of 10, the cross-correlated signal should have a 1 second SNR of 42 dB or greater in the absence of the atmosphere. While cross-correlation SNR's of the prototype GPS-MET instrumentation are not this high, future instrumentation will produce 1 second cross-correlation SNR's of order 47 dB indicating the ionosphere calibration can be maintained through most of the troposphere with unclassified instrumentation. Since errors due to incomplete calibration of the ionosphere are generally negligible in the lower troposphere (Sections 3.7 and 3.14), interpolation or extrapolation of the ionosphere calibration across the short vertical intervals in the troposphere where attenuation is too large to maintain cross-correlation tracking, should have negligible impact on retrieval accuracy.

Atmospheric multipath

Atmospheric multipath (AMP) refers to the situation where more than one ray path through the atmosphere connect the transmitter and receiver. AMP requires that the magnitude of the bending angle *increase* locally with altitude requiring that the magnitude of the refractivity gradient *increase* with altitude and $d^2N/dr^2 < 0$ in this local vertical interval. From eqs. (2.2.6) and (2.2.7), the requirement for multipath is the difference between bending angles, $\alpha(r_1)$ and $\alpha(r_2)$, for two signal paths simultaneously reaching the receiver must be

$$\Delta\alpha \simeq \Delta a/D \sim \Delta r/D \qquad (2.4.6)$$

where $\Delta r = r_1 - r_2$ and D is the distance from the receiver to the limb. Eq. (2.4.6) assumes the receiver is much closer to the limb than the transmitter as is the case for a LEO GPS receiver. A more restrictive definition of AMP requires that the multiple ray path tangent heights be separated by at least the diameter of the first Fresnel zone ($\Delta r \geq Z_F$) in order that

distinct geometrical rays exist. The $\alpha(a)$ profile derived from these rays will then be invertable via the Abel transform. With D = 3000 km and $Z_F \sim 1$ km, $\Delta\alpha \sim 3x10^{-4}$ radians = 0.018 degrees which is a small fraction of total bending angle in the troposphere (figure 2.6) implying multipath may be encountered frequently in the troposphere.

The most common cause of AMP at GPS wavelengths is sharp layers in moisture and sometimes temperature in the troposphere. A narrow layer of a high concentration of water will produce a dipole spike in refractivity gradient which is superimposed on the approximately exponential refractivity gradient of the background atmosphere. A situation frequently encountered in the subtropics is the transition from a dry region to a moist region below, such as the transition from the free troposphere to a convective boundary layer below. In this geometry, a single spike in refractivity gradient at the transition is superimposed on the approximately exponential gradient of the background atmosphere. The spike in refractivity gradient produces a sharp peak in bending angle at the altitude of the spike which decreases rapidly at lower tangent altitudes because of the very sharp weighting of the $[r^2n^2 - a^2]^{-1/2}$ term in eq. (2.2.1) near the tangent altitude. As a result, AMP is limited to an interval at and just below the high refractivity gradient layer.

Using simple scaling arguments one can determine approximately the effect of sharp transition layers on bending angle and the conditions of AMP. From eq. (2.4.6), the requirement for AMP can be written as

$$\alpha(z_t) - \alpha_0(z_t - \Delta z) \sim \Delta z / D$$

where z_t is the ray path tangent height in the transition layer with the high refractivity gradient and $z_t - \Delta z$ is some lower altitude where $\Delta z \geq Z_F$, the diameter of the first Fresnel zone. The bending at height, z_t, is the bending due to the background exponential structure of the atmosphere plus the contribution due to the high gradient in the transition layer:

$$\alpha(z_t) = \alpha_0(z_t) + \delta\alpha_t(z_t) \cong \alpha_0(z_t - \Delta z) + \Delta z / D \qquad (2.4.7)$$

The 0 subscript refers to the bending in the absence of the transition layer and $\delta\alpha_t(z_t)$ is the bending contribution of the transition gradient. The background bending is approximately exponential with altitude:

$$\alpha_0(z_t - \Delta z) \cong \alpha_0(z_t)\, e^{\Delta z / H_\alpha} \tag{2.4.8}$$

Combining eqs. (2.4.7) and (2.4.8) yields the extra bending from the transition layer required to produce AMP:

$$\delta\alpha_t(z_t) \cong \alpha_0(z_t - \Delta z) - \alpha_0(z_t) + \Delta z/D \cong \alpha_0(z_t)\left[e^{\Delta z / H_\alpha} - 1\right] + \Delta z/D \tag{2.4.9}$$

Half of the bending occurs within a vertical interval, $\Delta z_{1/2}$, of the ray tangent height (figure 2.8). Further, for a given refractivity gradient, the contribution of a given layer scales approximately as the square root of the layer thickness (analogous to discussion in section 3.1). Therefore, the total bending at z_t can be written as

$$\frac{\alpha_0(z_t) + \delta\alpha_t(z_t)}{\alpha_0(z_t)} \simeq 1 + \frac{\left(\Delta^{dN/dr}\right)_t}{(dN/dr)_0} \frac{1}{2}\left[\frac{\Delta z_t}{\Delta z_{1/2}}\right]^{1/2} \tag{2.4.10}$$

where Δz_t is the thickness of the transition layer, $(\Delta dN/dr)_t$ is the extra refractivity gradient contributed by the transition layer and $(dN/dr)_0$ is the refractivity gradient at z_t contributed by the exponential refractivity behavior. Combining eqs. (2.4.9) and (2.4.10) yields

$$e^{\Delta z / H_\alpha} - 1 \simeq \frac{\Delta z}{H_\alpha} \simeq \frac{\left(\Delta^{dN/dr}\right)_t}{(dN/dr)_0} \frac{1}{2}\left[\frac{\Delta z_t}{\Delta z_{1/2}}\right]^{1/2} - \frac{\Delta z}{D\,\alpha_0(z_t)} \tag{2.4.11}$$

and the requirement on the incremental increase in refractivity gradient at the transition layer is

$$\left(\Delta^{dN/dr}\right)_t \simeq 2\,(dN/dr)_0\,\Delta z\left[\frac{1}{H_\alpha} + \frac{1}{D\,\alpha_0(z_t)}\right]\left[\frac{\Delta z_{1/2}}{\Delta z_t}\right]^{1/2} \tag{2.4.12}$$

In the troposphere, bending is generally large enough that the $D\,\alpha_0$ term can be ignored to first order. With representative tropospheric values of $\Delta z = Z_F \sim 0.8$ km, $\Delta z_{1/2} = 1.6$ km (figure 2.8), $H_\alpha = 6$ km and $\Delta z_t = 0.1$ km, eq. (2.4.12) requires the additional refractivity gradient of the layer be at least 110% of the background atmosphere gradient to produce AMP with distinctly separated ray paths.

Atmospheric multipath due to the dry term : The large refractivity gradient needed to create AMP can be provided by a large positive temperature gradient at a thermal inversion. The density scale height, H_ρ, (= H_N, the refractivity scale height, in dry air) is given by

$$H_\rho = \left(\frac{g\,m}{R\,T} - \frac{dT}{Tdz}\right)^{-1} = \frac{RT}{mg}\left(1 + \frac{dT}{Tdz}\frac{RT}{gm}\right)^{-1} = \frac{H_P}{\left(1 + H_P\frac{dT}{Tdz}\right)} \qquad (2.4.13)$$

Since the refractivity gradient is $-N/H_N$, increasing the gradient requires decreasing H_N which in turn requires that the inversion temperature gradient be related to the background temperature gradient as

$$\left(\frac{dT}{dz}\right)_t = \frac{T}{H_P}\left[\frac{H_{N0}}{H_{Nt}} - 1\right] + \left(\frac{dT}{dz}\right)_0 \frac{H_{N0}}{H_{Nt}} \qquad (2.4.14)$$

The strong tradewind inversion across the subtropics provides a case where AMP will occur. For ATEX, the average inversion thickness was ~200m and Z_F is ~0.6 km in the lower troposphere where the inversion exists (figure 2.7b) so the requirement for atmospheric multipath according to eq. (2.4.12) is a refractivity gradient higher than that of the background atmosphere by 60% and the required temperature gradient at the inversion is +9 K/km. Since the average temperature gradient across the tradewind inversion observed during ATEX was +25 K/km, geometric multipath should occur systematically at the tradewind inversion. The humidity contribution will further enhance the refractivity gradient at the inversion. While easily met at the tradewind inversion, generally it will be difficult to find temperature gradients large enough to create AMP. Notice the conditions necessary for atmospheric multipath are easier to meet than the critical refraction conditions which are not met by average ATEX temperature inversion.

Atmospheric multipath due to humidity gradients: Creating AMP with a water vapor layer requires that the concentration of the vapor be large enough and vary rapidly enough with altitude to meet the requirement of eq. (2.4.12). The extreme situation of 100% humidity below the layer and 0% above establishes the minimum temperature at which geometric AMP can be caused by water vapor for a given layer thickness:

$$P_w^* \simeq 2\frac{T^2 N}{b_2 H_N}\Delta z\left[\frac{1}{H_N}+\frac{1}{D\,\alpha_0(z_t)}\right]\left[\Delta z_t\,\Delta z_{1/2}\right]^{1/2} \qquad (2.4.15)$$

where P_w* is the saturation vapor pressure. The height where annual climatological values of N and P_w* are related via eq. (2.4.15) for a 100 m thick layer and 1 km separation of the multipath rays varies with latitude as indicated in Table 2.1. The maximum height decreases by more than 5 km from the equator to high latitudes. The maximum height of AMP at the equator is ~5 km higher than the maximum height of critical refraction.

Table 2.1

Maximum height of atmospheric multipath

Latitude	Altitude (km)	P (mb)	T (C)	T (K)	N
60	4.5	550	-26	247	173
30	8.3	350	-30	243	112
0	9.8	287	-32.5	240.5	93

Notice that in the subtropics, the last signal observed in a setting occultation and the first signal observed during a rising occultation may be the ray path with a tangent height at the tradewind inversion.

AMP signal acquisition: All ray paths must be acquired during an occultation because each ray path has a unique tangent altitude and samples a unique value of a in the $\alpha(a)$ profile. The receiver must therefore simultaneously track multiple images of the same GPS signal. Notice that when AMP is produced by a single layer, there are three coexistent ray paths. Their tangent heights are such that one is just above and another is just below the maximum gradient of the layer, and the third lies farther below where the background

exponential atmospheric gradient contributes most of the bending. If multiple layers exist within the same AMP interval, the number of simultaneous ray paths may be 5 or 7 and will always be an odd number. However, if the critical refraction occurs, a particular ray path will not be detectable. A very important point is that even though the ray paths cross, each has a unique Doppler shift under the assumption of local spherical symmetry such that each signal and its corresponding α and a can be separated according to its frequency.

Sorting $\alpha(a)$: Another complication when AMP occurs is the ray path tangent height no longer decreases monotonically with time and the resulting sequence of $\alpha(a)$ will not be a monotonically descending order of a. Therefore, the $\alpha(a)$ sequence must be sorted into a monotonically ascending or descending order of a needed to calculate the Abel transform.

2.5 Spatial Resolution

Limits to vertical and horizontal resolution.

This section considers the fundamental limits to the spatial resolution of atmospheric profiles retrieved from GPS radio occultation measurements. The vertical resolution of a single, instantaneous bending measurement is determined by the contribution of individual atmospheric layers to net bending along the ray path (eq. (2.2.1)). Better vertical resolution can be retrieved from a profile of closely separated measurements, subject to limits imposed by diffraction. For the limb sounding geometry appropriate to GPS radio occultation, horizontal resolution may be defined by the distance traversed by the radio path as it enters and exits a layer having a vertical resolution of ΔZ; in this case the horizontal and vertical resolution are related by the approximate expression:

$$\Delta L = 2(2R\Delta Z)^{1/2} \qquad (2.5.1)$$

where ΔL is horizontal resolution, ΔZ is vertical resolution, and R is the radius of the atmosphere at the ray path tangent height.

The vertical resolution expected from GPS radio occultation measurements is described below, and the possibility of improving on the limits imposed by diffraction is discussed. The section concludes by emphasizing the need for measurements with high vertical resolution and with spatial resolution that is compatible with the vertical and horizontal scales of characteristic atmospheric structure. Limb sounding radio occultation measurements are shown to provide a good match to this structure.

Vertical resolution - Individual measurements of bending angle

The variation of the vertical resolution of an individual bending angle measurement (one in a sequence of measurements taken during an occultation) is illustrated in figure 2.8 as a function of ray path tangent altitude for the atmospheric profile used in figure 2.7. The curves on the right show the relative contribution of each altitude interval to net atmospheric bending for several tangent altitudes. These contribution functions are strongly peaked at the tangent altitude due to the limb sounding geometry and exponential variation of refractivity with altitude (eq. (2.2.1)). Vertical resolution, defined as the vertical interval above the tangent height contributing 50% to net bending, is plotted as a function of tangent altitude in the left-hand side of figure 2.8. For the atmospheric profile used in figure 2.8, the 50% bending interval varies from 0.5 to 2.0 km and corresponding via eq. (2.5.1) to a 50% horizontal interval from 160 to 320 km. The smallest 50% interval occurs at ray path tangent altitudes near 3.8 km where large vertical refractivity gradients associated with rapid changes in humidity across the top of the boundary layer concentrate the bending such that vertical resolution is improving in a region where it is needed most.

Although the vertical resolution of single occultation measurements is described well by figure 2.8, the vertical resolution of refractivity profiles retrieved from GPS radio occultation profiles can be significantly better given profiles of closely spaced bending angle measurements, high measurement precision, and an inversion scheme that correctly represents the contribution of different levels to bending angle. All of these conditions are

met for GPS occultations measured by a LEO satellite, provided the atmosphere is locally spherical in structure so that the Abel transform inversion of eq. (2.2.1) can be used. Under these conditions, the ultimate limits to vertical and horizontal resolution of retrieved profiles are set by diffraction.

Vertical resolution - Diffraction limit.

As with other remote sounding techniques, radio occultation resolution is limited by diffraction. The theory of Section 2.2 assumes that transmitter signals follow a dimensionless ray path to the receiver. This is the path of minimum phase delay relative to the delay associated with neighboring, physically realizable paths. However, a geometrical ray path is an approximation to Fresnel diffraction as wavelength approaches zero. Because signals originate at a point source and are detected by a receiver phase-locked to the minimum phase delay, the cross-beam sampling at the limb takes the form of concentric Fresnel zones of positive and negative response centered on the ray path. The contributions of adjacent Fresnel zones cancel so that the effective sampling can be represented by the first Fresnel zone [Born and Wolf, 1980].

In the absence of significant atmospheric bending, the diameter of the first Fresnel zone, Z_F, at the ray path tangent level is given by:

$$Z_F = 2\,[\lambda\, L_T L_R / (L_T + L_R)]^{1/2} \qquad \approx 2\,[\lambda\, L_R]^{1/2} \quad (\text{If } L_T >> L_R) \qquad (2.5.2)$$

where λ is the GPS signal wavelength, and L_T and L_R are the distance from the transmitter and receiver to the limb respectively. L_T is approximately 28,500 km for the GPS transmitters and, for a receiver in a 700 km altitude orbit, L_R is about 3,100 km. At the GPS L1 wavelength of 19 cm, eq. (2.5.2) yields a value for Z_F of 1.4 km. From eq. (2.5.1) the horizontal resolution corresponding to a vertical resolution of 1.4 km is about 270 km.

Modification in the Fresnel zone radius due to the atmosphere

Diffraction-limited vertical resolution is also influenced by the vertical refractivity gradient. Deep in the atmosphere, the radius of the first Fresnel zone will shrink because the path delay increases more rapidly with a change in path tangent height than in free space.

We will assume a spherically symmetric atmosphere such that the only dependence is on r, the radius. The variable, a, defined in eq. (2.2.5) is constant along a given ray path. In figure 2.9, three adjacent ray paths are shown for a fixed receiver location. The gray line represents the curve through the ray path tangent points as a changes. The goal is to find the point x_{12} such that the propagation time from the receiver through x_{12} via paths x1 and x1' is one half cycle greater than the propagation time along path 2. To accomplish this, we will estimate the difference in delay between points x_2 and x_1 and then points x_1 and x_{12} and then take the difference as the difference between the propagation delay to points x_{12} and x_2.

In changing from ray 2 to ray 1, the separation angle between the transmitter and receiver changes by

$$\Delta\theta = \int_{a_1}^{a_2} \frac{d\theta}{da}\, da \tag{2.5.3}$$

The equivalent change in propagation delay is

$$\Delta\tau = \int_{a_1}^{a_2} \frac{d\theta}{da}\frac{d\tau}{d\theta}\, da = \tau_2 - \tau_1$$

From Bullen (1953), $d\tau/d\theta = a/c$ where c is the speed of light in a vacuum so

$$\Delta\tau = \frac{1}{c}\int_{a_1}^{a_2} \frac{d\theta}{da}\, a\, da$$

The propagation time between x_1 and x_{12} along ray path 1 is

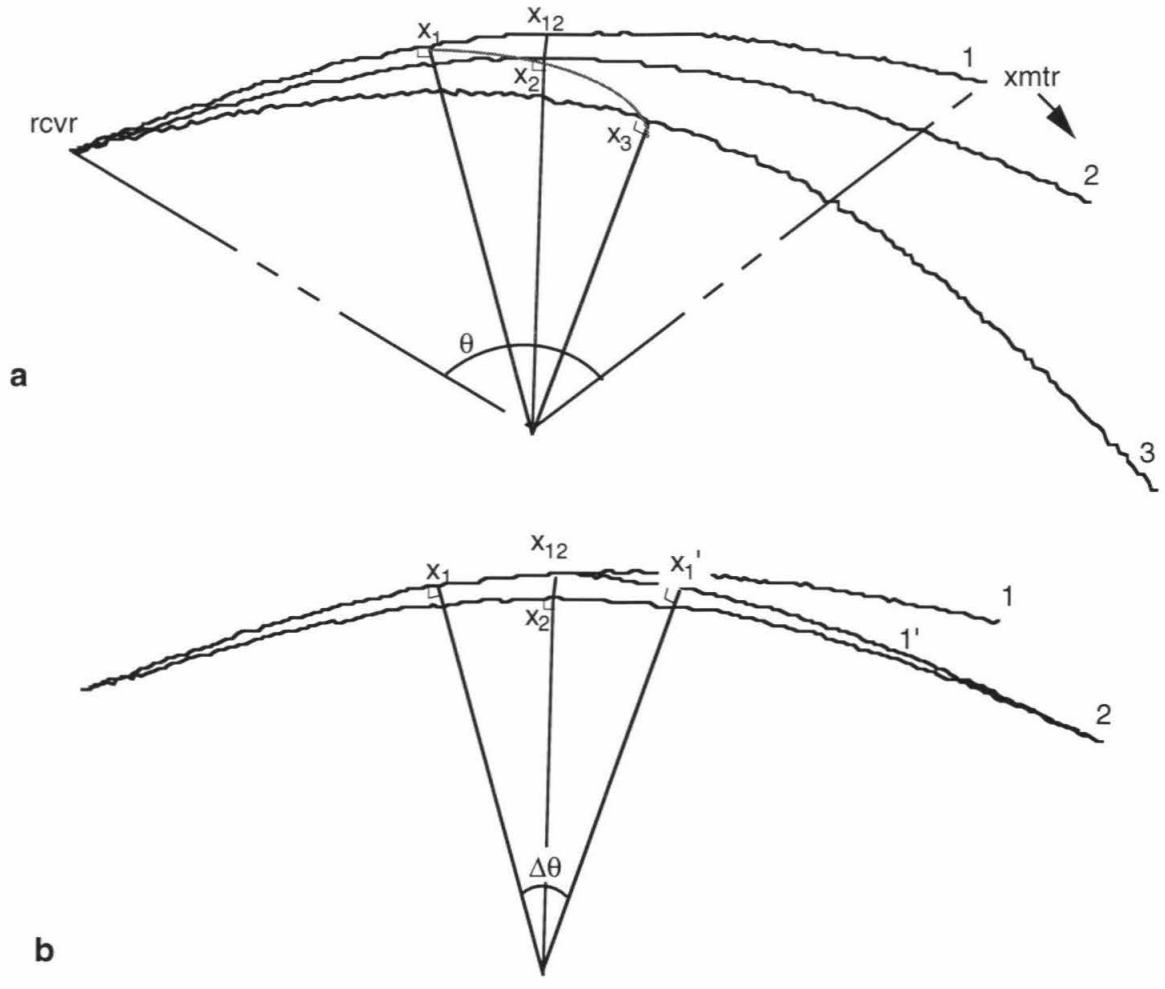

Figure 2.9 Geometry for deriving the modification to the diameter of the first Fresnel zone due to the atmosphere.

$$\Delta\tau = \int_{\theta_1}^{\theta_2} \frac{d\tau}{d\theta}\, d\theta = \int_{\theta_1}^{\theta_2} \frac{a_1}{c}\, d\theta = \frac{a_1}{c}\int_{\theta_1}^{\theta_2} d\theta = \frac{a_1}{c}\int_{a_1}^{a_2} \frac{d\theta}{da}\, da$$

where the last step is because the angle, $\Delta\theta$, is defined by eq. (2.5.3). The difference in propagation time is therefore

$$\Delta\tau = \frac{a_1}{c}\int_{a_2}^{a_1} \frac{d\theta}{da}\, da - \frac{1}{c}\int_{a_2}^{a_1} \frac{d\theta}{da}\, a\, da = \frac{1}{c}\int_{a_2}^{a_1} \left[a_1 - a\right] \frac{d\theta}{da}\, da$$

The radial distance separating a_1 and a_2 under the constraint that the extra delay equals $\lambda/2c$

$$\int_{a_2}^{a_1} \left[a_1 - a\right] \frac{d\theta}{da}\, da = \frac{\lambda}{2} \qquad (2.5.4)$$

is the radius of the first Fresnel zone and the diffraction resolution is roughly twice this length scale. Eq. (2.5.4) can be solved numerically given a discrete set of $\frac{d\theta}{da}$ vs. a derived via raytracing.

Approximations

Assuming $\frac{d\theta}{da}$ changes slowly over the interval from a_2 to a_1, eq. (2.5.4) can be approximated as

$$\frac{\left[a_1 - a_2\right]^2}{2} \overline{\frac{d\theta}{da}} = \frac{\lambda}{2}$$

where $\overline{\frac{d\theta}{da}}$ is the average value of $\frac{d\theta}{da}$ over the a_2 to a_1 interval. Since $a = nr$ at the ray path tangent point,

$$dr = \frac{da}{r\frac{dn}{dr} + n}$$

and

$$\left[r_1 - r_2\right] \cong \sqrt{\frac{\lambda}{\overline{\frac{d\theta}{da}}}}\, \frac{1}{\left[r\frac{dn}{dr} + n\right]}$$

which can also be written as

$$\left[r_1 - r_2\right] \cong \sqrt{\frac{\lambda}{\frac{d\theta}{dr}\frac{dr}{da}}}\,\frac{dr}{da} = \sqrt{\frac{\lambda}{\frac{d\theta}{dr}\frac{da}{dr}}} \qquad (2.5.5)$$

figure 2.7b shows the results of eq. (2.5.5) using raytracing results through a spherically symmetric atmosphere based on a radiosonde profile from Hilo, Hawaii, in July 1991. These results were checked approximately using the expression

$$Z_F = Z_{F\,(free\ space)}\ \ M^{1/2}$$

from Fjeldbo and Eshleman [1969] where M is defined in eq. (2.4.1). The variation of M with altitude for the temperature and humidity profiles acquired during a radiosonde ascent from Hilo, Hawaii, in July 1991 is shown in figure 2.7a.

The gradual decrease in Z_F from 1.4 km in the stratosphere to less than 0.5 km near the surface is due to the exponential increase of refractivity gradient with decreasing altitude. Within the troposphere, rapid variations in Z_F are due primarily to sharp changes in vertical refractivity gradient caused by changes in humidity. This effect is particularly strong within a few hundred meters of the surface and at the top of the boundary layer near 3.8 km, where diffraction-limited vertical resolution lies in the 100 to 200 m range.

From the examples given above, it is clear that diffraction-limited vertical resolution of GPS radio occultation measurements is comparable to or better than the vertical resolution imposed by atmospheric structure on a single bending angle measurement. Furthermore, vertical resolution is best in the troposphere and in regions with large vertical refractivity gradients where it is most needed. The diffraction limited horizontal resolution of occultation measurements in the cross-beam direction is approximately 1.4 km at all levels, because horizontal refractivity gradients are generally small.

Improved vertical resolution.

Vertical resolution superior to that imposed by Fresnel diffraction should be possible under certain conditions. Atmospheric structures smaller than the Fresnel scale produce diffraction patterns in the signal measured by the receiver, known as phase and amplitude scintillations. Because occulted signals are monochromatic with well defined phase, scintillation measurements obtained during an occultation can in principle be inverted to remove the effects of diffraction and reconstruct vertical profiles at resolutions finer than the Fresnel scale, provided atmospheric structure varies primarily in the vertical rather than horizontal direction such that the 1-D diffraction pattern acquired in the occultation recovers sufficient information to allow the inversion.

Techniques for inverting and removing the effects of diffraction in Voyager occultation observations of the Saturnian and Uranian ring systems were developed by researchers at Stanford University which dramatically improved radial resolution relative to diffraction-limited scales [Marouf *et al.*, 1986]. The removal of diffraction effects within an atmosphere is more complicated because the atmosphere is a three-dimensional structure. The idea has been extended to atmospheres, using the thin phase screen approximation, to detect and characterize planetary waves in the Uranian stratosphere [Hinson and Magalhaes, 1991]. Development of a fourier optics approach to account for and remove the effects of diffraction from observations of thin atmospheres is under development [Karayel and Hinson, 1996]. Another idea is to "back-propagate" the signal from the actual reception location to a location closer to the Earth's limb [Gurvich and Sokolovskiy, 1996] which reduces the Fresnel zone size and improves vertical resolution. Since the resulting improvement in resolution scales as the ratio of the original receiver altitude to the new reception altitude to the one-fourth power, lowering the reception altitude from 700 km to 100 km (roughly the height of layers in the E-region of the ionosphere) via back-propagation, would improve vertical resolution by $7^{1/4}$ or 40%. If

the altitude were reduced to 30 km (where there is little bending), vertical resolution would increase by a factor of 2.2, improving near-surface resolution to ~200 m (figure 2.7b), a significant increase given the small vertical extent of the boundary layer.

The questions concerning diffraction correction lie in how to perform the back-propagation, and how well back-propagating across the refractivity structure of the ionosphere and the upper atmosphere will work. Further development is also needed to extend diffraction removal into the relatively high refractivity environment of Earth's troposphere.

The importance of consistent vertical and horizontal resolution

From eq. (2.5.1), the ratio of horizontal to vertical resolution for GPS radio occultation measurements is approximately $\Delta L/\Delta Z \sim 2(2R/\Delta Z)^{1/2}$, and ranges from 190 when ΔZ = 1.4 km to 700 when ΔZ = 0.1 km, with typical tropospheric values near 270. These magnitudes are similar to the ratio of consistent horizontal and vertical resolution discussed by Lindzen and Fox-Rabinowitz [1989] for quasi-geostrophic flow and internal gravity waves and to the resolution ratio of regional and global atmospheric models. For quasi-geostrophic flow the ratio is the Rossby ratio between vertical and horizontal scales given by the expression:

$$\Delta L/\Delta Z = N_B/f_o \qquad (2.5.6)$$

where f_o is the coriolis parameter and N_B is the Brunt-Väisälä frequency. The similarity between the GPS values and those derived from eq. (2.5.6) demonstrates the relationship between vertical and horizontal resolution for GPS radio occultation measurements is consistent with most atmospheric structure and that occultation measurements are well suited for assimilation into numerical weather prediction models, particularly regional and

global models. Further, the ability to routinely probe the troposphere in a limb sounding geometry complements nadir-viewing observations with passive sensors which provide high horizontal resolution, but more limited vertical resolution.

3. Refractivity Profile Inversion - Sources of Error and Expected Accuracy

3.1. Introduction

In this section, the accuracy with which refractivity, density, pressure and temperature can be derived from occultation observations is discussed and assessed. The section is structured as follows. Section 3.2 begins with a general overview of error propagation through the profile retrieval process working sequentially through derivations of refractivity, density, pressure and temperature. The overview is followed by a sequence of subsections briefly describing each of the errors identified below. Using the approach described in Section 3.2, the impact of each error on derived refractivity, pressure and temperature is characterized. When possible, analytic methods are used. Finally, the individual error contributions are drawn together into an overall, vertical accuracy budget for refractivity, density, pressure and temperature which also defines the limiting sources of error. A reader primarily interested in the error summary may wish to move immediately to Section 3.14.

Individual error sources can be grouped into three categories: measurement errors, calibration errors and retrieval errors. The more significant sources are summarized below and are discussed in this order in the rest of this section.

Measurement errors:

- Signal to noise ratio performance
- Clocks instability
- Local multipath

Calibration errors:

- Residual ionospheric effects
- Orbit determination accuracy

Differencing schemes to remove clock error

Retrieval errors:

Refractivity constant uncertainties

Uncertainties in Abel and hydrostatic integral boundary conditions

Horizontal refractivity structure

The geometrical optics approximation

Water vapor/density ambiguity

3.2 Error Propagation through GPS Profile Retrievals

To characterize the individual sources of error which follow, a general description of the propagation of errors through the retrieval process is needed. For errors related to the phase observations, the first step is to determine the response of the transformation from measured atmospheric phase delay to refractivity to a single phase perturbation. This impulse response includes the sequential conversion from phase to frequency, then bending angle and finally refractivity (via the Abel transform). Knowledge of the impulse response allows a complete description of the impact of phase measurement errors on refractivity as well as density which is proportional to refractivity in the absence of significant water vapor. The second step describes the transformation of refractivity errors into errors in pressure through the hydrostatic integral. The third step describes how errors in refractivity and pressure combine to produce temperature errors via the ideal gas law.

Impact of phase error on refractivity

Imperfect phase measurements degrade derived refractivity accuracy. Derivation of refractivity from measurements of phase can be divided into three steps. First, phase is differentiated to produce frequency. Second, frequency is transformed into bending angle

and finally bending angle is converted into refractivity via the Abel integral. These steps are linear, at least at the error magnitudes discussed in this section, and the error in refractivity due to errors in phase can therefore be described using linear system theory as [Papoulis, 1984]:

$$\delta N(t) = \delta\phi(t) \otimes h_D(t) \otimes h_A(t) \tag{3.2.1}$$

where δN and $\delta\phi$ are refractivity and phase errors, h_D is the impulse response of the differentiation step, h_A is the impulse response of the combined second (frequency to bending angle) and third (Abel transform) steps and $\otimes$ signifies convolution. Describing the transformation of noise in terms of an impulse response also allows the refractivity retrieval process to be characterized in the fourier domain, described by Appendix A.2 which will be used in Sections 3.4 and 3.5 to describe the effects of clock instability and local multipath error.

The impulse response of the differentiator depends on the particular numerical implementation chosen. For simplicity, a simple discrete differentiator which differences consecutive phase samples is used here:

$$h_D(t) = \frac{\delta(t + \frac{\tau}{2}) - \delta(t - \frac{\tau}{2})}{\tau} \tag{3.2.2}$$

where $\delta(t)$ represents the Dirac delta function and τ is the time interval between phase measurements which is set to the time required for the ray path to cross the diameter of the first Fresnel zone, the practical resolution limit in the geometric optics approximation (eq. 2.5.2). The impulse response, $h_A(t)$, will be characterized numerically by inputting a set of Doppler frequencies containing a single frequency error into the bending angle and Abel transformation steps. Integration of the Abel transform requires numerical integration, the details of which affect response to errors. The specific implementation used here approximates a continuous representation of $\alpha(a)$ for the Abel integration by fitting a cubic

polynomial to $ln(\alpha)$ as a function of a where the natural logarithm is used because, to first order, the dependence of bending angle on miss distance is exponential. Figure 3.1 shows the response of the Abel transform to a single 1 mm/sec error equivalent to a frequency error of $5x10^{-3}$ Hz for a radio wavelength of 20 cm. The response includes errors in derived radius and height.

From eq. (3.2.2), it is clear that a single phase error produces a double pulsed frequency error which, when convolved with h_A, produces the vertically focused refractivity error shown in figure 3.2, defining the refractivity error response to phase error. The altitude of the largest refractivity error in figure 3.2 occurs at the altitude of the first of the two frequency peaks. The second frequency pulse cancels much of the remaining contribution of h_A's convolution with the first frequency pulse.

Relationship between height and time

Because phase noise is typically a function of time, eq. (3.2.1) relates measurement error to refractivity error as a function of time whereas atmospheric refractivity is a function of height to first approximation. Time and height are related by the rate at which the ray path tangent height descends (or ascends in a rising occultation) through the atmosphere. At altitudes near the tropopause and above, where bending is not too large, this rate is roughly $V_\perp$, the component of the LEO orbital velocity orthogonal to the limb, typically between 2 and 3 km/sec. For consideration of measurement noise, the region most affected is the stratosphere because, as will be shown, sources of error other than measurement noise dominate in the lower stratosphere and below.

Scaling the impulse response to shorter and longer integration times

Frequency error is proportional to phase error divided by the phase estimation time interval, τ_{est}, so refractivity error must be proportional to τ_{est}^{-1}. Since vertical resolution is

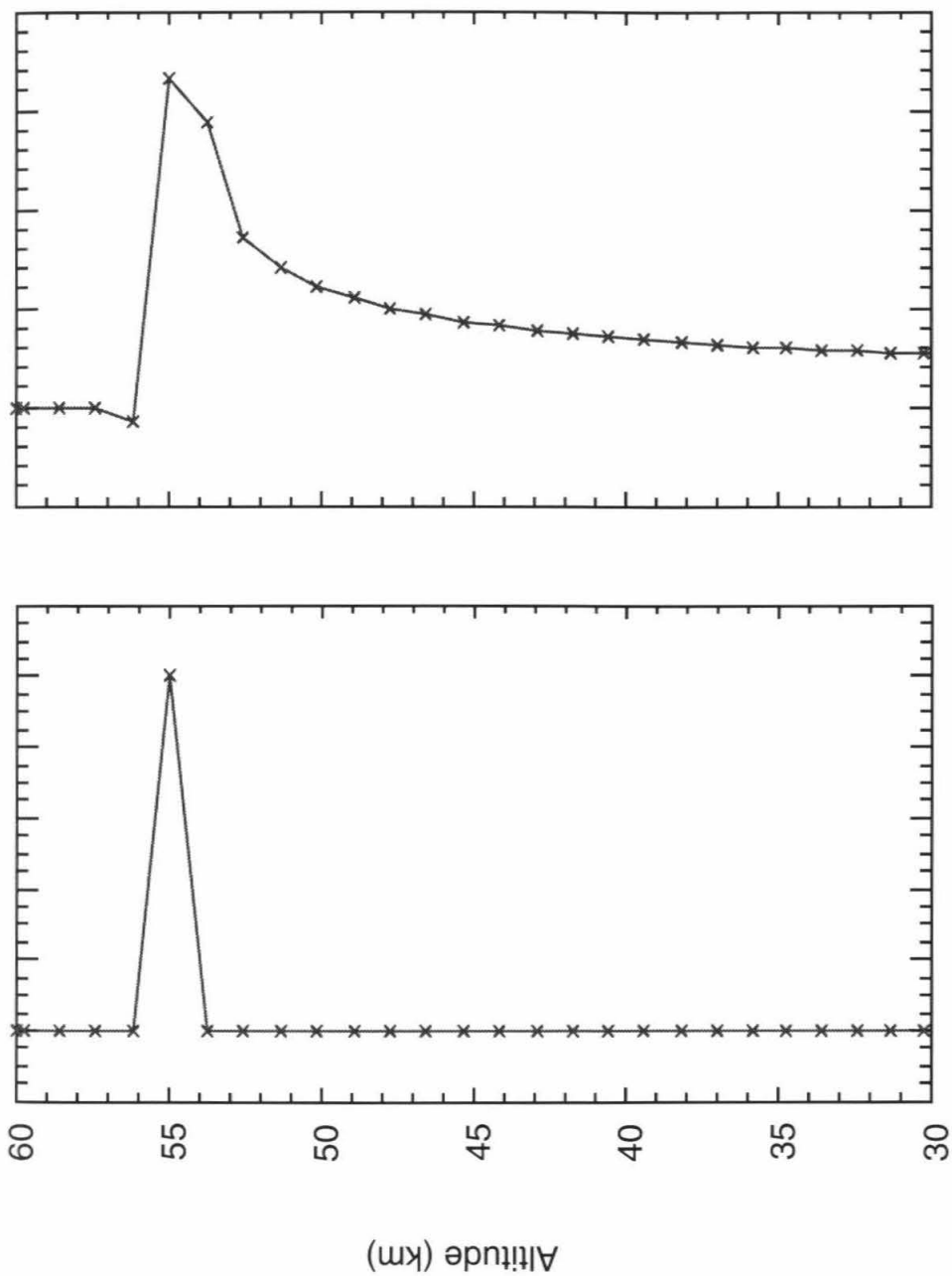

Figure 3.1 (a) Single frequency measurement error. (b) Refractivity error response to frequency error in (a).

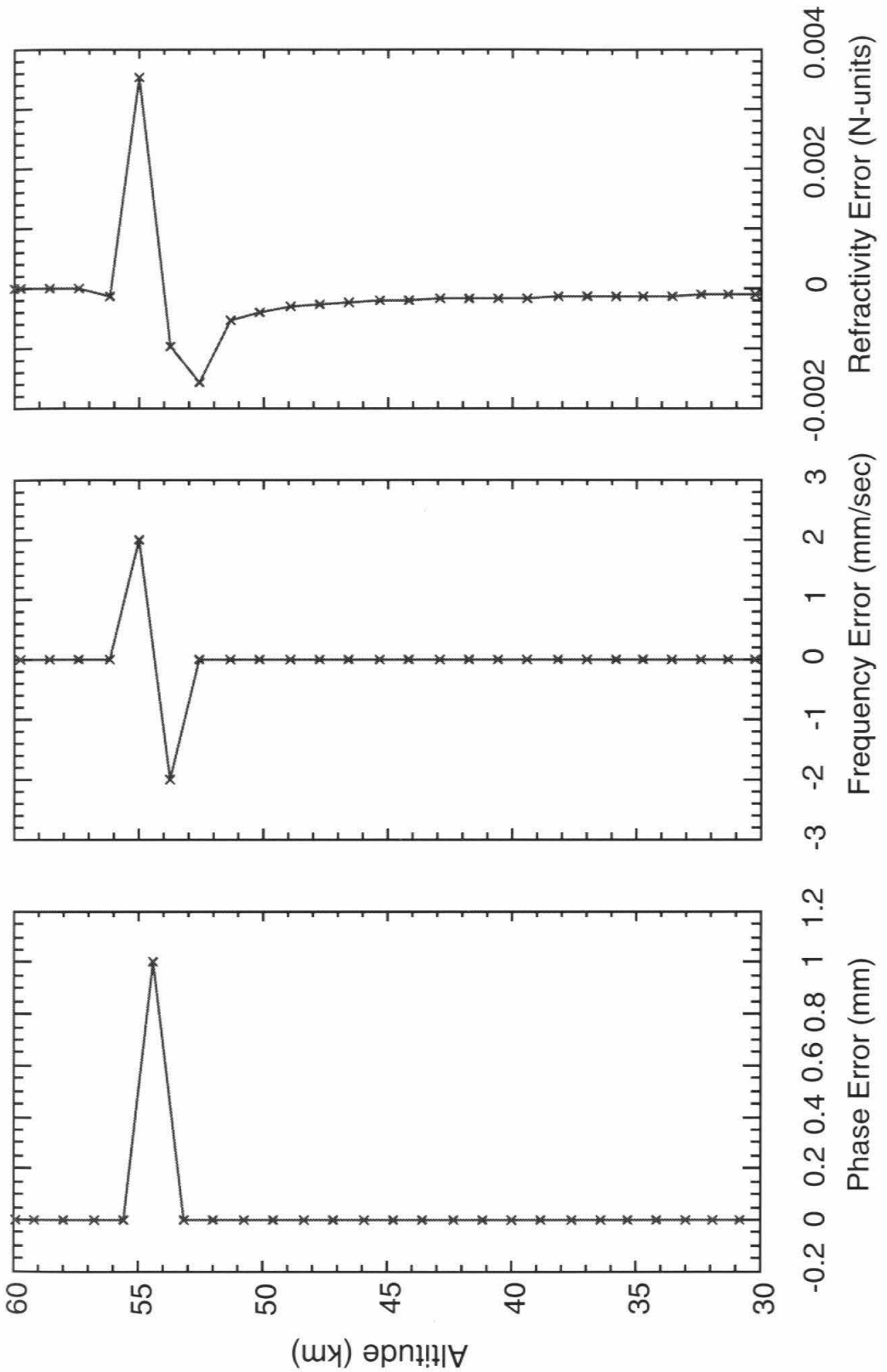

Figure 3.2 (a) single phase error. (b) Frequency error resulting from (a). (c) Refractivity error response to frequency error in (b).

also tied to τ_{est}, an improvement in refractivity error will come at the expense of vertical resolution. Of the various sources of error, thermal noise, described in Section 3.3, exhibits the greatest sensitivity to τ_{est}. For thermal noise, rms phase error is proportional to $\tau_{est}^{-1/2}$, and frequency and therefore refractivity error scale in proportion to $\tau_{est}^{-3/2}$. The $\tau_{est}^{-3/2}$ scaling can be used at high altitudes where thermal noise can dominate other sources of error in a tradeoff of resolution for accuracy. For instance, increasing τ_{est} by a factor of 2 will reduce the resolution by the same factor of 2, but, will reduce thermal error by a factor of 2.8, which is equivalent to raising the altitude at which a level of accuracy is achieved by a scale height.

The magnitude of the discrete frequency error impulse response, $h_A(t)$, is proportional to the sampling interval because the frequency error is not a true delta function due to the finite width of the input error pulse. The numerical implementation of the Abel integral interpolates between samples and the actual error in the Abel transform is somewhat triangular in shape where the width at the mid-height of the triangle is the time interval between samples fed into the Abel integral. Therefore, when the interval separating the discrete values which make up the impulse response is changed, the magnitude of each of the discrete values scales as the square root of the ratio of the new and original intervals.

Effect of errors in refractivity on density and pressure

Under dry conditions, density is proportional to refractivity via eq. (2.3.2) and so density errors are proportional to refractivity errors and their fractional errors are equal. The effect of water vapor on retrievals of density, pressure and temperature becomes important in the warm regions of the troposphere as is described in Section 3.13. In the extremely dry air found in the cold (< 250K) regions of the troposphere and stratosphere, where density is directly proportional to refractivity, pressure is derived from density via the hydrostatic integral

$$P(z) = P(z_u) + \int_z^{z_u} g\,\rho\,dz'$$

where z_u is the height at which the integral is initiated and $P(z_u)$ is the pressure at that height. The choice of z_u is discussed in Section 3.10 based upon a tradeoff between noise characteristics described later. The fractional pressure error is then

$$\frac{\delta P(z)}{P(z)} = \frac{\delta P(z_u)}{P(z)} + \frac{\int_z^{z_u} g\,\delta\rho(z')\,dz'}{P(z)} \qquad (3.2.3)$$

The contribution to the fractional pressure error at z by the error in the initial guess of pressure at z_u, $\delta P(z_u)$, decreases exponentially as z decreases. As temperature is less variable than pressure at middle atmospheric heights, $P(z_u)$ is probably best estimated from estimates of $T(z_u)$ and $\rho(z_u)$ with $T(z_u)$ estimated from a combination of climatology, models and other observations and $\rho(z_u)$ derived directly from the occultation observations themselves. The pressure error can be written as

$$\frac{\delta P(z)}{P(z)} = \left(\frac{\delta T(z_u)}{T(z_u)} + \frac{\delta\rho(z_u)}{\rho(z_u)}\right)\frac{P(z_u)}{P(z)} + \frac{\int_z^{z_u} g\,\delta\rho(z')\,dz'}{P(z)} \qquad (3.2.4)$$

Effects of refractivity errors on geopotential height

Atmospheric fluid dynamics equations are generally written with pressure rather than altitude as the vertical coordinate such that the geopotential associated with a pressure level becomes a dependent variable of fundamental interest. Since the change in pressure with height is defined via the hydrostatic and ideal gas law as

$$\frac{\delta P}{P} = -\frac{m\,g}{R\,T}\,\delta z$$

errors in pressure height, δz, and geopotential, $\delta\Phi$, are related to the fractional pressure error as follows:

$$\frac{\delta P}{P} = -\frac{\delta z}{H} = -\frac{\delta \Phi}{g\,H} \tag{3.2.5}$$

Therefore the approximate geopotential height error can be found by multiplying the fractional pressure error at any derived pressure level by the local pressure scale height.

Effect of Density and Pressure Errors on Temperature

Temperature is derived via the ideal gas law, $T = P/(R\rho)$, and fractional temperature error is therefore related to fractional errors in density and pressure by

$$\frac{\delta T}{T} = \frac{\delta P}{P} - \frac{\delta \rho}{\rho} \tag{3.2.6}$$

The fractional mean square error in temperature is given by

$$\left\langle \left(\frac{\delta T}{T}\right)^2 \right\rangle = \left\langle \left(\frac{\delta P}{P}\right)^2 \right\rangle + \left\langle \left(\frac{\delta \rho}{\rho}\right)^2 \right\rangle - 2\left\langle \frac{\delta \rho}{\rho}\,\frac{\delta P}{P} \right\rangle \tag{3.2.7}$$

The magnitude of the temperature error therefore depends on the vertical correlation of the density error. If the density error correlation length is small such that density error varies rapidly relative to a scale height, its integrated effect in the pressure error (eq. (3.2.3)) is small causing the rapid density errors to map almost directly into temperature errors. When density errors vary slowly relative to a scale height, the fractional pressure error will be greater than the fractional density error and the pressure error contribution to temperature error will dominate. When the density error correlation length is roughly a scale height, fractional density and pressure errors will be comparable and the fractional temperature error will be smaller than either of them. Note that when density error systematically increases as $(z_u\text{-}z)^n$, the temperature error will pass through zero at a height, n scale heights below z_u, marking the transition between the density error dominated region above and the pressure error dominated region below.

3.3 Thermal Noise

The finite GPS signal strength and the receiver amplifier noise combine to produce a random phase error, referred to as thermal error, which is independent from sample to sample. Because thermal errors vary rapidly with height, they contribute significantly to refractivity and temperature error at high altitudes but produce little pressure error.

The mean square error of thermal phase noise is given by

$$\left\langle \delta\phi(\tau)^2 \right\rangle^{1/2} = \frac{\lambda\, \sigma_n(\tau)}{2\pi A} = \frac{\lambda}{2\pi}\left(2\, SNR_0\, \tau\right)^{-1/2}$$

where $<\delta\phi^2>^{1/2}$ is rms phase error in units of length, τ is the integration time in seconds, λ is the GPS signal wavelength, A is the signal amplitude, σ_n is the root-mean-square (rms) of the thermal noise and SNR_0 is the signal to noise ratio for an integration time of 1 second. The signal strength is determined by GPS transmitted power, antenna gain and the atmospheric attenuation mechanisms discussed in Section 2.4. σ_n is determined by the receiver amplifier and blackbody radiation received from the Earth. Table 1 shows SNR's and rms thermal phase errors representative of the prototype GPS-MET and future receiving systems. Using future generation LEO receiver hardware, the rms phase measurement errors will be ~0.05 mm for an integration time of 0.5 seconds, the approximate time for a ray path tangent point to descend a Fresnel diameter, z_F, in the middle atmosphere. Subsequent amplification of the phase error by a factor of 3 due to ionosphere calibration (Section 3.7) and 2 due to clock double-differencing (Section 3.8) yields a calibrated rms phase error of ~0.3 mm in 0.5 seconds for the GPS receiver being developed within JPL. When the AS encryption is on (Section 2.4), use of an unclassified receiver with a small receiving antenna will roughly double the thermal error (column 3, Table 1).

Refractivity error

Using the impulse response described in Section 3.2 (eq. (3.2.1)), the refractivity noise variance due to thermal noise is

$$\left\langle \delta N_i^2 \right\rangle = \left\langle \left(\sum_{j=1}^{j\,\max} \delta\phi_j \, h_{DA_{i-j}} \right)^2 \right\rangle = \left\langle \delta\phi^2 \right\rangle \sum_{j=1}^{j\,\max} h_{DA_{i-j}}^2 \tag{3.3.1}$$

where j is the sample number counted from the first sample used in the Abel transform and $j_{max} = (z_a\text{-}z)/z_F$ where z_a is the height at which the Abel transform is initiated. The second step of eq. (3.3.1) can be made because the sample to sample correlation is zero and $\langle\delta\phi^2\rangle$ is independent of altitude at least for ray paths at or above the tropopause where thermal noise can be a significant contributor to refractivity error. More than 95% of the total variance of the impulse response given in figure 3.2 is contributed by the phase error in the measurement made at the altitude of the refractivity retrieval such that the impact of white phase noise on derived refractivity is quite focused vertically. The rms refractivity error of eq. (3.3.1) is also independent of altitude. An rms refractivity error of 0.0012 N-units produced by an rms phase error of 0.3 mm in 0.5 seconds corresponds to fractional refractivity errors of ~0.6% and 2% at 50 and 60 km altitude depending on latitude and season which are significant at high altitudes.

Pressure errors

Given a sequence of independent refractivity errors, the pressure error via eq. (3.2.3) is:

$$\delta P(z) = \sum_{j=1}^{j\,\max} \int_{z_i}^{z_i + z_F} g \, \delta\rho(z') \, dz'$$

where $j_{max} = (z_h\text{-}z)/z_F$ and z_h ($\leq z_a$) is the height at which the hydrostatic integral is started. Since the $\delta\rho$ errors are independent of one another, the rms pressure error is random walk in form described as

$$\left\langle \delta P(z)^2 \right\rangle^{1/2} = \left\langle \delta \rho^2 \right\rangle^{1/2} g\, z_F^{1/2} \left(z_h - z\right)^{1/2} \tag{3.3.2}$$

The square root dependence on altitude in eq. (3.3.2) represents the weakest height dependence of any pressure error considered here, and the fractional pressure error therefore decreases rapidly as z decreases because of the exponential increase in pressure with decreasing height.

Temperature errors

Combining eqs. (3.3.1) and (3.3.2) into (3.2.7) yields

$$\frac{\left\langle \delta T(z)^2 \right\rangle^{1/2}}{T(z)} \cong \frac{\left\langle \delta \rho^2 \right\rangle^{1/2}}{\rho(z)} \left[\frac{z_F\left(z_h - z\right)}{H(z)^2} - 2\frac{z_F}{H(z)} + 1 \right]^{1/2} \tag{3.3.3}$$

where H(z) is the pressure scale height at the altitude, z. Beginning at z_h, the density error contribution (the third term in the brackets) dominates that of the pressure error (the first term in the brackets) until $(z_h\text{-}z) \sim H^2/z_F$ which, for representative values of z_F = 1.4 km and H = 7 km, is about 5 scale heights or 35 km. Therefore, the density error contribution to temperature error generally dominates that contributed by pressure error in altitude regions where the contribution of thermal noise is significant.

3.4 Clock and Troposphere Errors

Since signal phase is the principle observable in occultation observations, the phase stability of the reference clocks of the GPS transmitters and receivers is important. The phase error associated with fractional frequency stabilities of 10^{-13} available from ultra-stable quartz oscillators does not limit accuracy. GPS satellite cesium clock instabilities are approximately an order of magnitude larger over the time scales of interest (0.01 to 60 seconds) and could be important at high altitudes. The most significant clock related error at present is the intentional variation of the GPS signal frequencies, referred to as Selective

Availability (SA) which is applied for security reasons. Removal of the SA modulation, which would otherwise ruin the observations, and clock instability via differencing is discussed in Section 3.8.

Even though clock error can largely be removed through differencing, it will be characterized here for several reasons. First, a tropospherically induced error is introduced on the GPS links to ground-based receivers used in differencing (Section 3.8) with stability characteristics very similar to those of high quality crystal oscillators. Second, if the SA modulation is turned off as the National Research Council [1995] has recommended, it can be advantageous to not difference, depending on the magnitude of clock error relative to the penalties incurred in increased thermal and local multipath errors. Finally, since this work represents a broad and systematic assessment of radio occultation technique accuracy, the results developed here are also relevant to planetary occultation experiments where differencing is not an option.

Clock noise is best described in the Fourier frequency domain and the transformation of clock phase error into refractivity, density, pressure and temperature will therefore be characterized in terms of the fourier domain relationships developed in Appendix A.2.

Refractivity density, pressure and temperature error

The phase noise spectra of reference clocks is generally power law in form [Barnes *et al.*, 1971]:

$$S_{\phi}(f) = \sum_{k=1}^{k\,\max} A_k \left| f \right|^{-b_k} \tag{3.4.1}$$

where f is the fourier frequency. For noise characterized by $b \leq 2$ which includes the cesium and rubidium clocks used on the GPS satellite transmitters and hydrogen maser frequency references used on the ground, eq. (3.4.1) can be plugged into the fourier

domain equations for mean square errors in refractivity, density, pressure and temperature respectively in Appendix A.2 to estimate errors in these same quantities due to clock instability. For $b \geq 3$, which includes quartz oscillators used in many space-borne applications, filtering to attenuate the low frequency phase noise is required to make the refractivity variance finite. Differencing the phase of each occultation measurement with the phase of measurements taken just prior to the occultation can provide the necessary attenuation. In this case, the filtered phase noise power spectrum, $S_{\phi\ filt}(f)$, is defined as

$$S_{\phi\ filt}(f) = 4 \sin^2(\pi f t_{occ})\, S_{\phi}(f) \qquad (3.4.2)$$

where t_{occ} is the length of an occultation. For clock phase noise characterized by $3 \leq b \leq 4$, eq. (3.4.2) can be plugged into the mean square error equations of Appendix A.2 to estimate errors in refractivity, density, pressure and temperature.

Discussion

Figure 3.3 shows fractional rms refractivity, pressure and temperature errors for two types of noise. The first represents GPS transmitter cesium clock noise (b=2 or white frequency [WF] noise with a 1 second Allan deviation of 10^{-12}). The second is representative of ultra stable quartz clocks noise such as the TOPEX and Mars Observer oscillators and tropospheric noise introduced on the differencing links described previously (b=3 or flicker frequency [FF] noise with a 1 second Allan deviation of 10^{-13}). In both cases the Abel transform is initiated at 60 km. The difference between the cesium and quartz type errors is indicative of the improvement which results if the GPS transmitter clock instability is differenced out. While the 1% refractivity error near 60 km due to the cesium will have some influence on refractivity accuracy, the small magnitude of these errors indicates clock errors should not dominate retrieval error in general.

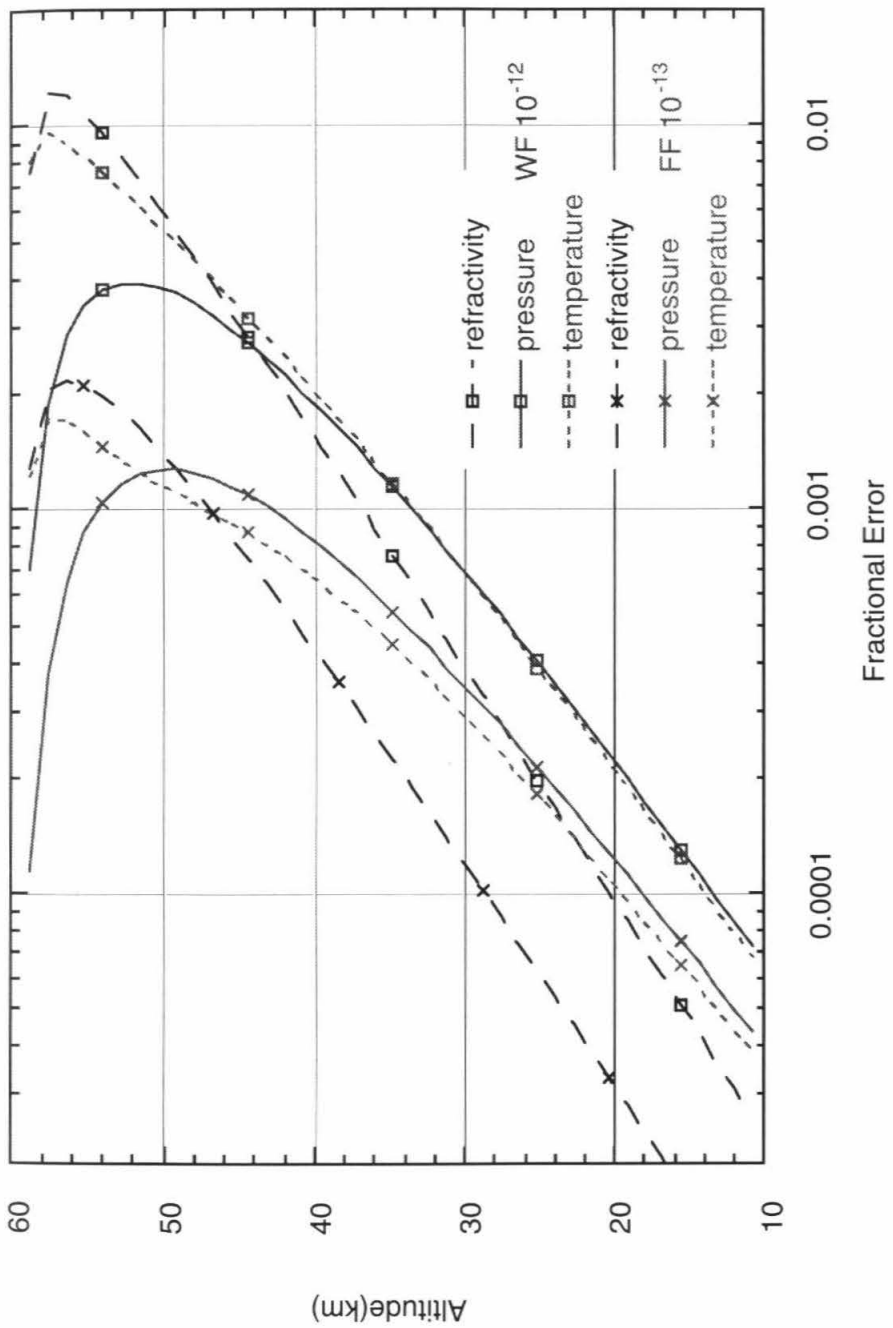

Figure 3.3 Fractional refractivity, pressure and temperature errors resulting from clock-type instabilities.
(a) $\sigma_y(1\ \text{sec}) = 10^{-12}$, b=3 representative of ultra-stable quartz oscillators or tropospheric propagation noise.
(b) $\sigma_y(1\ \text{sec}) = 10^{-12}$, b=2 representative of GPS satellite cesium oscillator instabilities

3.5 Local Multipath

The broad-beam receiver antennas which significantly simplify GPS instrumentation are unfortunately sensitive to local multipath which occurs when multiple images of the signal arrive at the antenna after scattering off structures in the vicinity of the antenna. In contrast, *atmospheric* multipath occurs when sharp vertical variations in atmospheric refractivity structure create multiple, simultaneous signal paths through the atmosphere. As discussed in more detail in Appendix A.3, local multipath creates slowly varying phase errors, a fraction of a wavelength in magnitude which affect accuracy, particularly pressure and geopotential accuracy, in the middle atmosphere where it may be the limiting error.

To evaluate the impact of local multipath on the retrieved refractivity profiles, the fourier representation of the refractivity retrieval scheme developed in Appendix A.3 is used. The multipath phase noise spectrum depends on the specific antenna and spacecraft configuration. Therefore, in order to make further progress, some spectral characteristics must be assumed. A reasonable first order approximation is to take a representative local multipath phase noise variance, σ_ϕ^2, and distribute it uniformly across a frequency interval, $\pm\Delta f$ in width. A frequency cut off of 0.01 Hz is used, representative of the high frequency limit for multipath on a 1 meter sized spacecraft [Appendix A.3]. The rms phase error, σ_ϕ, will be taken as 10 mm, representative of expected values for broad beam antennas [Appendix A.3].

Refractivity, pressure and temperature error

Plugging the local multipath noise power spectral density of $\sigma_\phi^2/(2\Delta f)$ into the fourier domain equations for mean square errors in refractivity, density, pressure and temperature respectively in Appendix A.2 yields estimates of rms errors in these same quantities. Refractivity error grows as the square root of Δz, the altitude interval between

the retrieval altitude and the Abel integral initialization altitude, z_u. The pressure error grows as $\Delta z^{3/2}$, consistent with eq. (3.2.3). At 10 mm rms and z_u= 60 km, the refractivity error is of order 1% between 50 and 60 km and decreases rapidly at lower altitudes. The resulting pressure error is significant throughout the stratosphere with an rms geopotential height error ranging from about 60 m near the stratopause to 6 m at 20 km altitude indicating multipath may limit geopotential accuracy through much of the stratosphere, particularly under nighttime or solar minimum conditions when the error contribution from the ionosphere is relatively small.

The fractional temperature error is the difference between the fractional refractivity and pressure errors. The maximum temperature error occurs about 3 km below z_u and for z_u = 60 km the error is ~2.5 K. Temperature error also exhibits a distinct minimum at 1.5 scale heights below z_u, consistent with the refractivity and pressure error scaling with Δz as discussed in Section 3.2. The secondary maximum error near 40 km is about 0.8 K. In the future, more directional antennas may reduce this error at a cost of instrumental complexity, cost and size. Also, scattering in the vicinity of the receiving antenna may be modeled such that multipath error is reduced during processing.

3.6 Orbit Determination

As discussed in Section 2.2, knowledge of the time-evolving viewing geometry is fundamental to interpreting the occultation observations and errors in geometry will degrade retrieved accuracy. The present level of orbital accuracy which can be achieved using GPS is sufficient to make their contribution to retrieval errors very small.

Position errors

Satellite position primarily affects retrieved refractivity altitudes. GPS transmitter and LEO receiver orbits are known to the meter and several cm levels respectively [Bertiger

et al., 1994]. For a receiver orbiting at 700 km altitude, the distance from the transmitter to the limb is roughly 9 times that of the receiver to the limb, such that the uncertainty in the ray path height at the limb is ~10 cm. Since other errors limit the geopotential height of pressure surfaces to at least the meter level, the contribution of a position error is not significant.

Refractivity , pressure and temperature errors due to velocity error

Orbital velocity error affects the retrieval process primarily through the component of the velocity error projected along the ray path. This component contributes directly as an erroneous atmospheric Doppler frequency. An error in velocity can therefore be treated as a frequency estimation error using the impulse response approach described in Section 3.2. The velocity of the receiver can be determined to ~0.05 mm/sec [Bertiger and Wu, 1996]. Figure 3.4 shows the effect of a 0.05 mm/sec velocity error on refractivity, pressure and temperature. All errors are small with maximum rms errors in geopotential height and temperature errors of ~10 m and 0.3 K at ~50 km and 57 km respectively for z_u = 60 km. The errors decrease rapidly at lower altitudes indicating orbital velocity error should not limit retrieval accuracy at any altitude.

Orbital acceleration errors over the short occultation duration are very small and therefore negligible.

3.7 Incomplete Calibration of the Ionosphere

The signal path is bent during propagation through the ionosphere on the way into and out of the neutral atmosphere. This bending effect must be removed to derive an accurate representation of neutral atmospheric structure, particularly at stratospheric altitudes and above. Although the presently-used calibration scheme of Vorob'ev and Krasil'nikova [1993] removes most of the ionospheric effect, some residual remains. Simple simulations indicate that the nighttime ionospheric structure during the maximum of

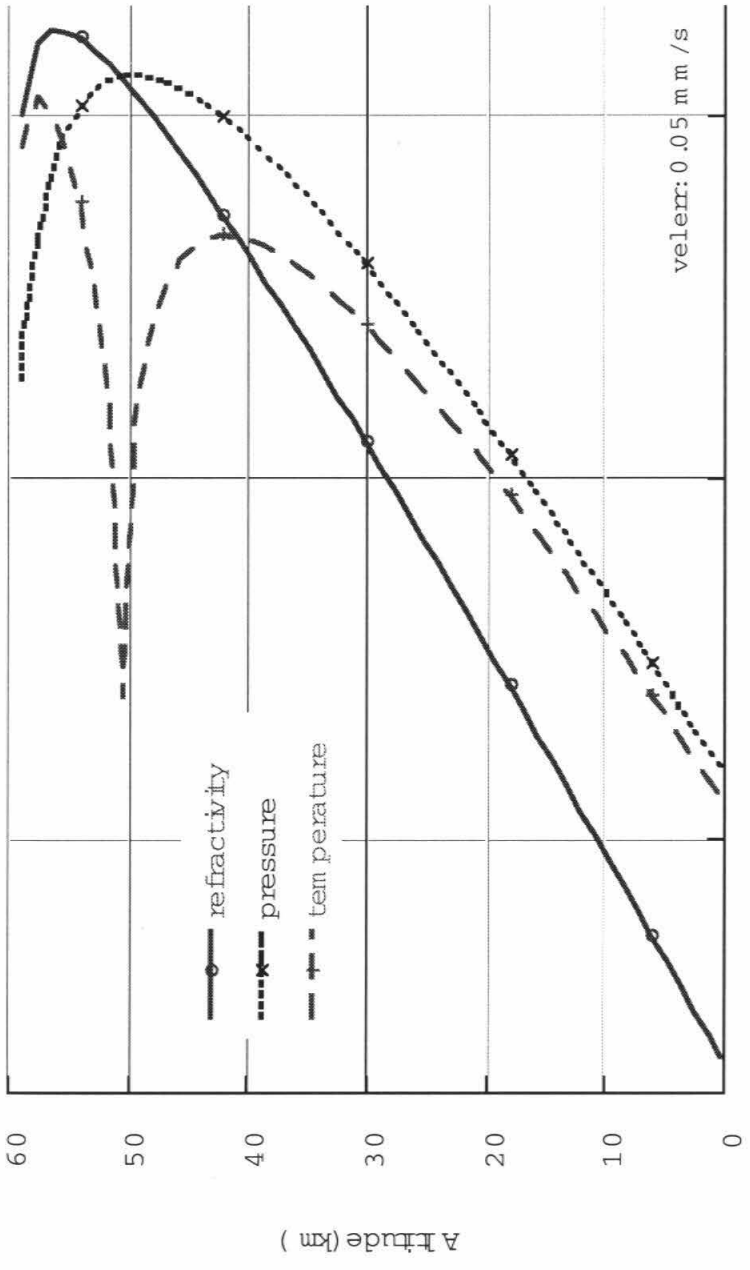

Figure 3.4 Fractional refractivity, pressure and temperature errors resulting from a 0.05 mm/sec velocity error along the line of site between the transmitter and receiver.

the solar cycle and daytime ionosphere at solar minimum will be removed to a level below that of other significant errors. Simulations also indicate that residual daytime ionosphere during solar maximum will be a limiting source of error in the stratosphere with the presently-used calibration scheme. Small scale ionospheric structure not eliminated entirely in the calibration process will appear as bursts of noise in the retrievals.

Ionospheric refractivity is dispersive and represented to first order (in MKS units) as $40.3 \times 10^6 \, n_e/f^2$ where n_e is the electron density and f is the radio signal frequency. The full Appleton-Hartree expression for ionospheric refractivity is given in Papas [1965], and a more complete power series expansion can be found for example in Bassiri and Hajj [1993]. Each GPS satellite continually radiates signals at two frequencies such that the dispersive ionospheric effect observed on links to ground-based receivers can be removed to first order using a linear combination of the two GPS signal frequencies

$$c_i \, L_1 - (c_i - 1) \, L_2 = L$$

where L is the optical path length along a signal path defined as

$$L_{opt} = \int_{path} n \, dl = \int_{path} \left(1 + N\right) dl$$

L_1 and L_2 are the optical path lengths at the two GPS signal frequencies, 1.57542 and 1.2276 GHz respectively and the coefficient, $c_i = \left(1 - f_2^2/f_1^2\right)^{-1}$ equals 2.5457...

In an occultation geometry, dispersive separation of the two GPS signal paths necessitates a different approach such as the one used here which was developed by Vorob'ev and Krasil'nikova [1993]. Substituting for the index of refraction given in eq. (2.3.1) in eq. (2.2.1) and expending in powers of (1/f) keeping only the first two leading terms, the total bending of the radio signal (assuming a spherically symmetric index of refraction) is given by

$$\alpha(a) = 2a \int_{r_0}^{\infty} \frac{\frac{dn_{neut}}{dr}\, dr}{n_{neut}\sqrt{\left(n_{neut} r\right)^2 - a^2}} + 2a \frac{40.3}{f^2} \int_{r_0}^{\infty} \frac{\frac{dn_e}{dr}\, dr}{n_{neut}\sqrt{\left(n_{neut} r\right)^2 - a^2}} \qquad (3.7.1)$$

The first and second terms on the RHS of eq. (3.7.1) correspond to the bending induced by the neutral atmosphere and ionosphere respectively. The ionosphere-free bending is then:

$$\alpha_{neutral}(a) = c_i\, \alpha_1(a) - \left(c_i - 1\right) \alpha_2(a) \qquad (3.7.2)$$

where α_1 and α_2 are the total bending angles for the two GPS signal frequencies. Application of eq. (3.7.2) requires, in general, that the discrete set of derived values of $\alpha_2(a_2)$ be interpolated to the set of a_1 values.

A penalty in applying this calibration is the amplification of thermal and local multipath errors. Assuming thermal noise at the two signal frequencies is comparable, the rms thermal error after application of the calibration is increased by a factor of ~3. The effect on local multipath noise scales similarly. This amplification factor of 3 is included in the discussions of thermal and local multipath noise in Sections 3.3 and 3.5.

To quantify to first order the performance of eq. (3.7.2) at removing the effects of the ionosphere, simulated occultation measurements have been generated by raytracing through a model consisting of Chapman layer representations of the F and E layers of the ionosphere and a representative neutral atmosphere. The four extremes of the diurnal and solar cycles have been represented based upon the typical mid-latitude ionosphere electron densities [Rich and Basu, 1985; Kelley, 1989]. Table 2 shows the residual bending angle error at 60 km altitude left after calibration for each of the four cases.

Pressure and temperature

The nighttime solar minimum error is very small, and the nighttime solar maximum and daytime solar minimum errors are sufficiently similar that only daytime solar maximum and nighttime solar maximum errors will be considered further. The impact of the daytime, solar maximum ionosphere on derived refractivity, pressure and temperature is shown in figures 3.5-3.7. Analogous results for nighttime solar maximum ionosphere conditions are shown in figures 3.8-3.10. The Abel and hydrostatic integrals in figures 3.5-3.10 were initiated with occultation data at 60 km. Refractivity errors increase approximately as the square root of Δz and fractional refractivity error therefore grows approximately exponentially with height. The geopotential height error in figure 3.6 indicates that residual ionosphere during daytime solar maximum conditions will limit accuracy above 25 km altitude. The peak temperature error of 6.5 K for daytime, solar maximum conditions occurs one Fresnel diameter below 60 km and a secondary peak of -1.5K occurs near 43 km (figure 3.7). The temperature error for all cases goes through zero near 50 km, consistent with the square root dependence of refractivity error on altitude (Section 3.2). The other three ionosphere cases do not appear to limit accuracy.

Discussion

The fractional errors in density, pressure and temperature all decrease rapidly with decreasing height because the density, pressure and temperature errors vary slowly relative to the exponential dependence of pressure and density on height. The negative bias apparent in the refractivity errors in Table 2 indicates that the calibration scheme defined by eq. (3.7.2) overcorrects slightly for the effects of the ionosphere and may be the one source of error which leaves a systematic bias signature in the retrievals. The temperature error behavior will be somewhat more complicated as indicated in figures 3.7 and 3.10.

The residual ionosphere errors in figures 3.5-3.10 can be improved using a better calibration scheme which leaves a residual zero-mean error. Higher order corrections to upward-looking ground-based GPS receiver observations have been developed to reduce

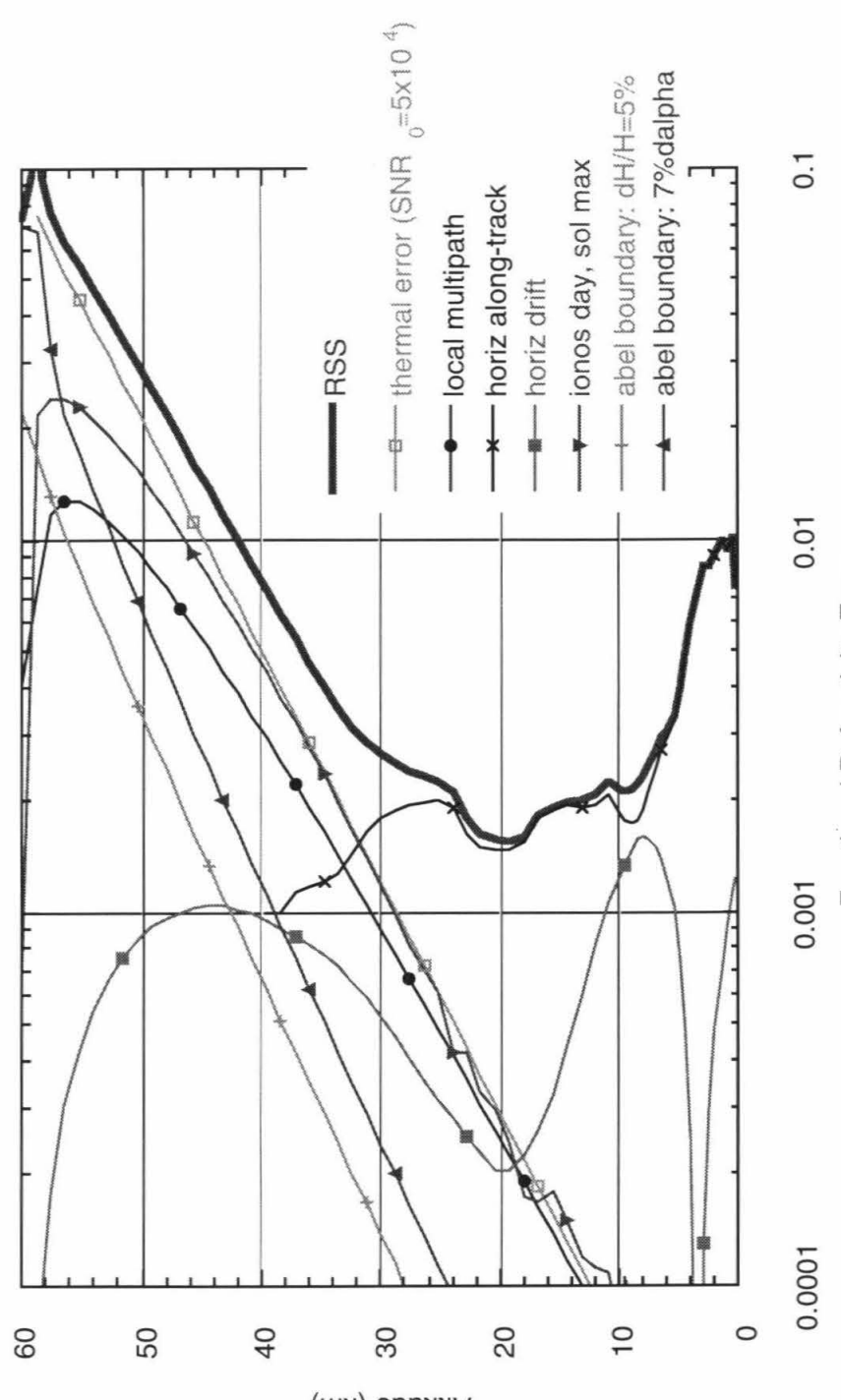

Figure 3.5 Summary of refractivity errors versus height. thermal error: 1 second SNR=$5x10^4$; local multipath: 10 mm rms spread over 0.01 Hz; horizontal refractivity structure: along-track from simulation and horizontal motion of ray path tangent point from tropospheric and stratospheric climatologies near 30°S for June-July-August; ionosphere error: daytime, solar maximum conditions; abel boundary errors: 7% in α, 5% in H_α

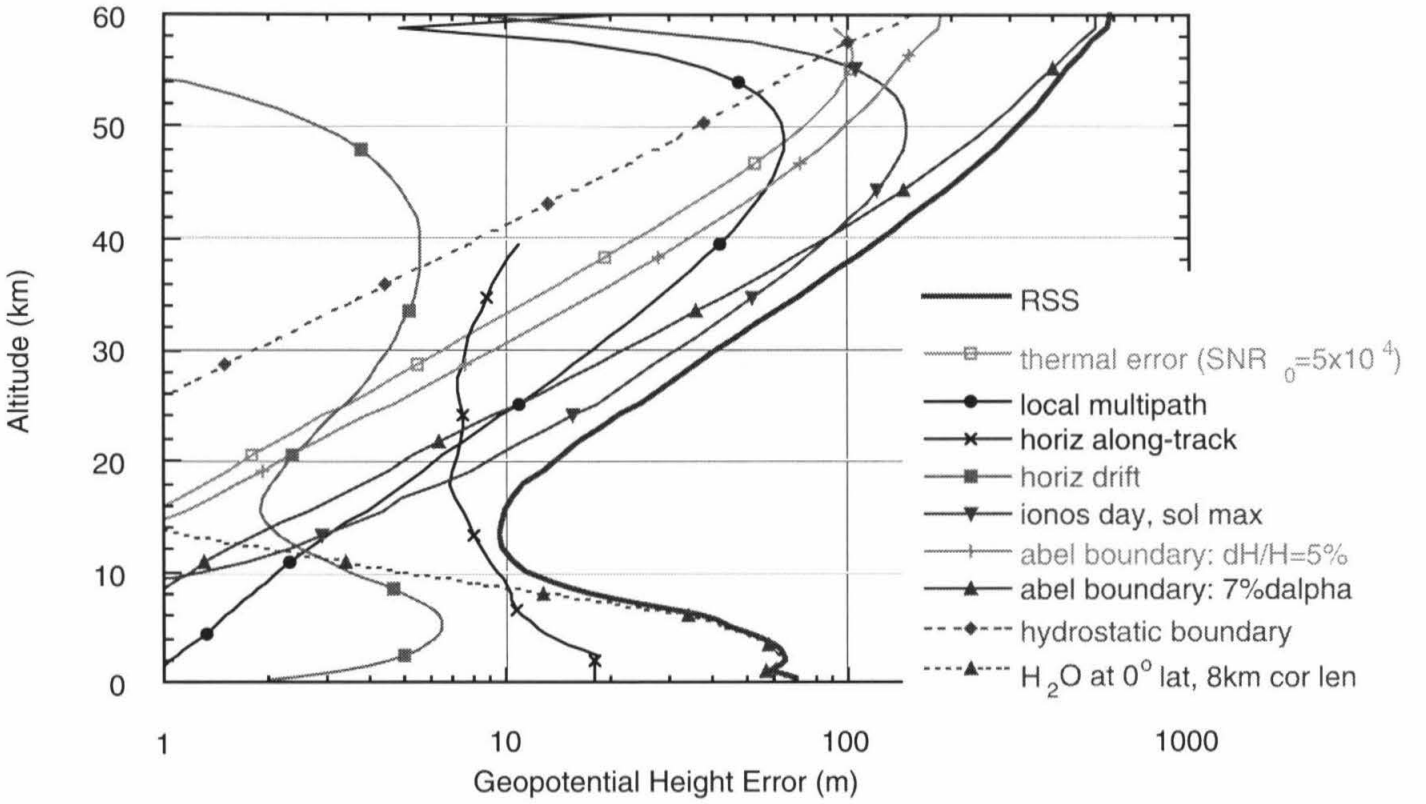

Figure 3.6 Summary of geopotential height errors versus height. thermal error: 1 second SNR=$5x10^4$; local multipath: 10 mm rms spread over 0.01 Hz; horizontal refractivity structure: along-track from simulation and horizontal motion of ray path tangent point from tropospheric and stratospheric climatologies near 30°S for June-July-August; ionosphere error: daytime, solar maximum conditions; abel boundary error: 7% in α, 5% in H_α; hydrostatic boundary error: 5 K; Tropospheric water vapor: 0° latitude with 8 km vertical correlation length based on figure 3.18

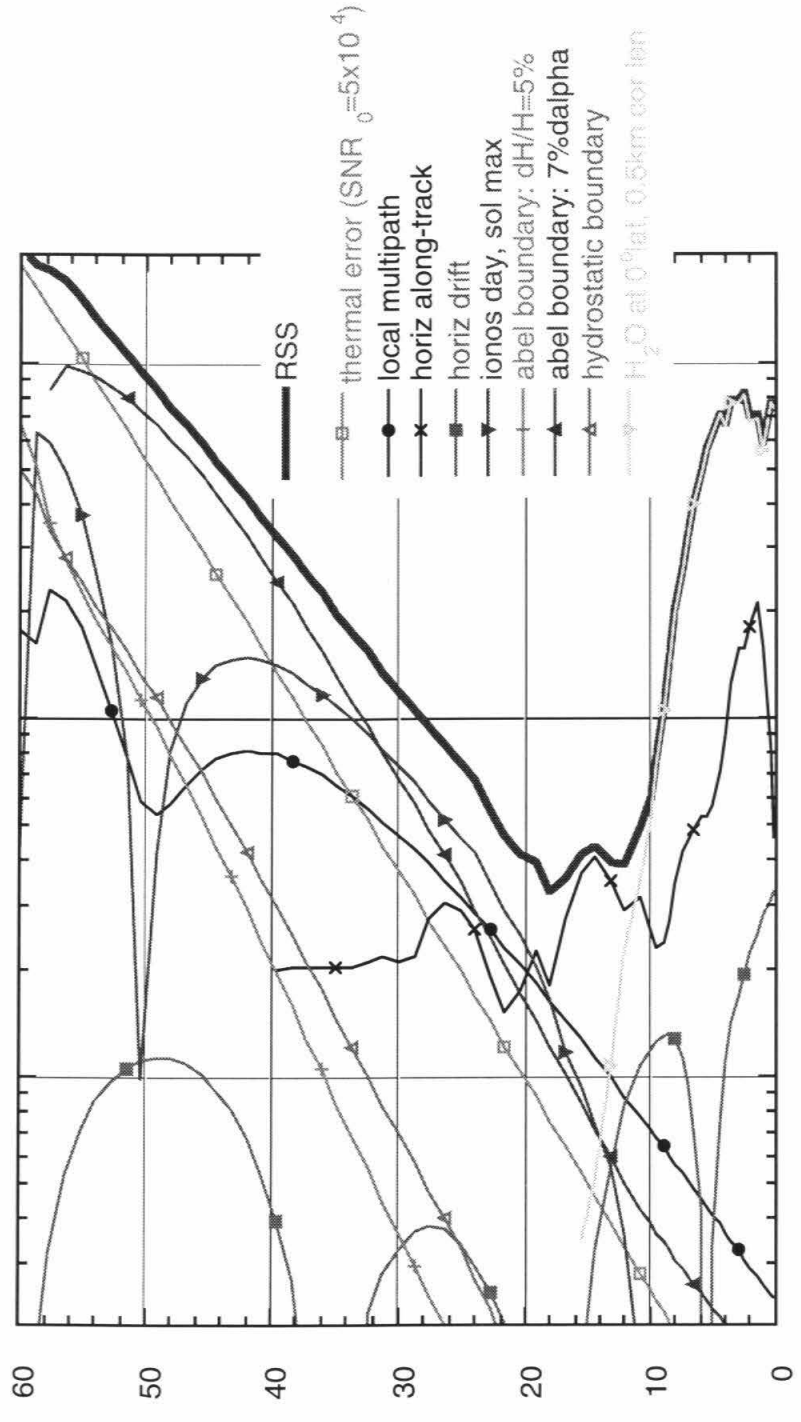

Figure 3.7 Summary of temperature errors versus height. Conditions same as figure 3.6.

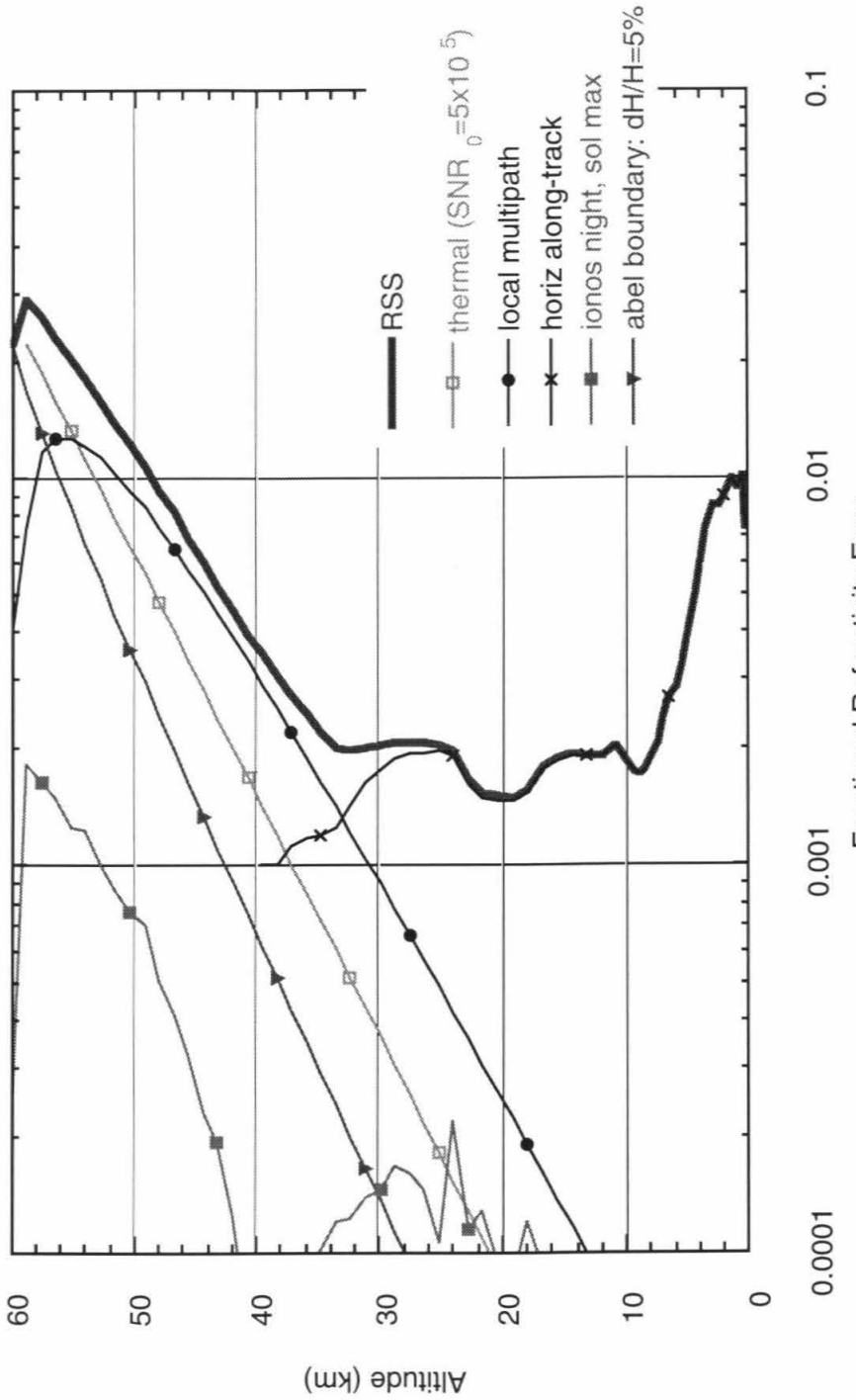

Figure 3.8 Summary of refractivity errors versus height. thermal error: 1 second SNR=$5x10^5$; local multipath: 10 mm rms spread over 0.01 Hz; horizontal refractivity structure: along-track from simulation; ionosphere error: nighttime, solar maximum conditions; abel boundary error: 5% in H_α

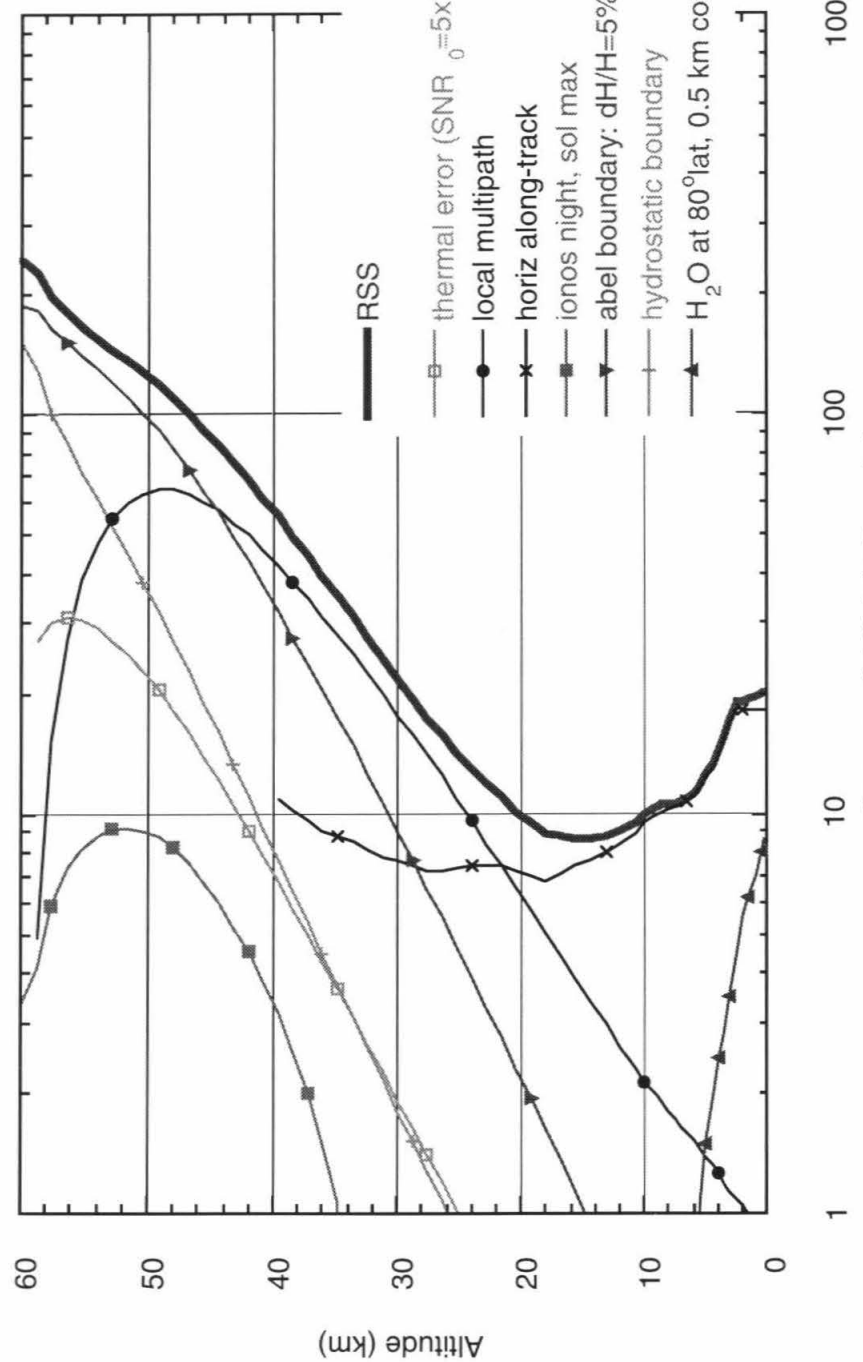

Figure 3.9 Summary of geopotential height errors versus height. thermal error: 1 second SNR=$5x10^5$; local multipath: 10 mm rms spread over 0.01 Hz; horizontal refractivity structure: along-track from simulation; ionosphere error: nighttime, solar maximum conditions; abel boundary error: 5% in H_α; hydrostatic boundary error: 5 K; Tropospheric water vapor: 80°latitude with 0.5 km vertical correlation length based on figure 3.18

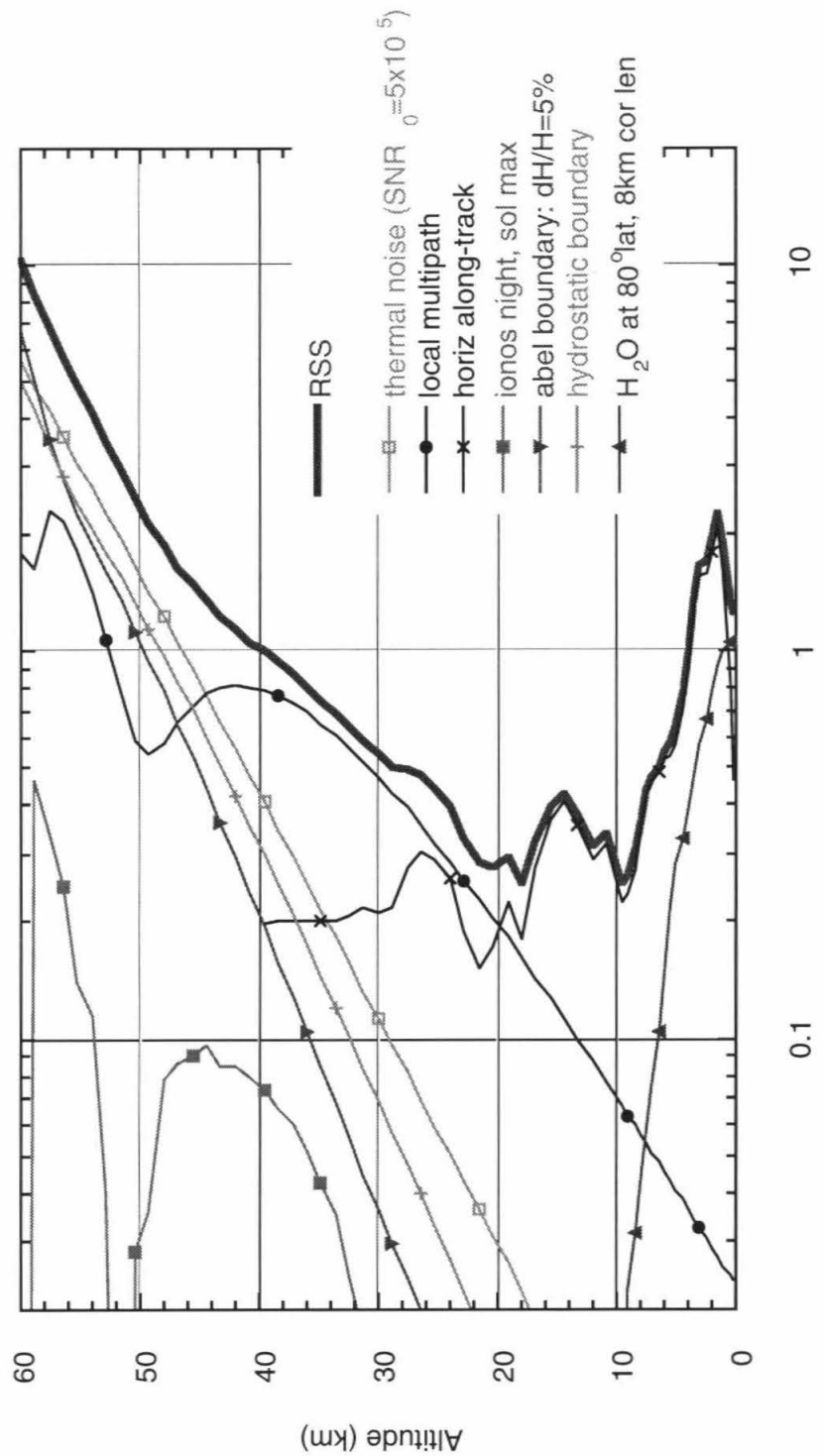

Figure 3.10 Summary of temperature errors versus height. Conditions same as figure 3.9.

errors by an order of magnitude [Bassiri and Hajj, 1993]. A similar higher order correction scheme should reduce ionospheric residual errors by at least a factor of 3. Nighttime observations at stratospheric altitudes should be very accurate for climate studies using the first order correction.

Caveats

The double Chapman layer representation of the F and E regions of the ionosphere used in the simulations presented here is realistic in a first order climatological sense. However, it contains neither horizontal nor small scale variations in structure. In the presence of horizontal variations, another source of error not considered here enters in the inversion process because the total bending is derived based on the assumption of spherical symmetry of the total index of refraction. This assumption is far less accurate in the ionosphere than in the neutral atmosphere due to the fact that the ray probes the ionosphere with a much longer scale than the neutral atmosphere. However, the error introduced by the non-spherical symmetry of the ionosphere is of order $(1/f^2)$ or higher (simply because the total ionospheric bending is of order $(1/f^2)$ or higher); therefore, the linear combination of eq. (3.7.2) will cancel the $(1/f^2)$ term leaving residual higher order terms which are comparable to the terms considered in the simulation.

Calibration of small scale variations in plasma density may be more problematic. Because the ray paths of the two GPS signal frequencies do not overlap for their entire length, the effect of plasma structure which varies over the scale of the separation between the two paths (of order 10-500 meter depending on the ionosphere and the tangent height) does not cancel completely when the present calibration is applied. In addition, structural variations on vertical scales smaller than the Fresnel zone diameter will produce scintillation (diffraction) effects which differ at the two wavelengths because Fresnel zone size depends on wavelength (eq. 2.5.2). These effects will be relatively rapid, somewhat similar to

thermal noise, and will be present during portions of occultations. The E-region, which exhibits sharp, small-scale structure and can vary dramatically over time scales of minutes [Kelley, 1989], is a likely culprit. Further, ionospheric scintillations produced at low latitudes by occasional rapid evening recombination events or magnetic storm activity will produce some residual effect because of sub-Fresnel structure. The overall conclusion is that the residual errors simulated here are probably representative at least in a climatological sense but a more complete assessment of the impact of the ionosphere is warranted as is the development of a higher order calibration scheme.

3.8 Differencing Schemes to Remove Clock Instabilities

The instability of the transmitter and receiver reference clocks and more importantly the SA encryption on the transmitter signal can be reduced to insignificant levels by differencing. Use of the double differencing scheme was in fact a crucial factor enabling the prototype GPS-MET mission through substantial reduction of both cost and development schedule. While the SA modulation can be eliminated with a classified facility, a more attractive, unclassified approach is to eliminate the SA modulation and inherent clock instabilities by differencing [Hajj and Kursinski, 1991, Hardy *et al.* 1994]. The general receiver and transmitter geometry required for differencing is shown in figure 3.11. Three differencing schemes are summarized in Table 3.

The differencing concept can be understood using scheme 1 in Table 3. In scheme 1, the carrier phase of link A (the occultation link in figure 3.11) is subtracted from that of link B (the link between GPS_1 and a ground receiver), which differences out the GPS_1 SA modulation and clock instabilities effectively replacing them with the instabilities of the ground receiver reference clock. Reduction or removal of the SA modulation without a classified facility requires either scheme 1 or 3. With either of these schemes and 1 second sampling and cubic interpolation, phase errors due to the SA modulation can be reduced to less than 0.1 mm [Wu *et al.*, 1990]. The reduction of receiver clock instabilities with

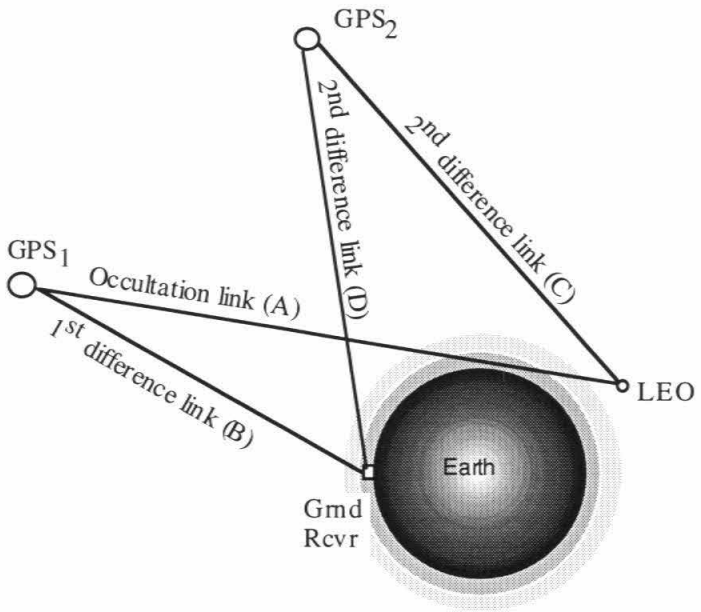

Figure 3.11 Geometry required for differencing to remove clock instabilities.

schemes 2 and 3 eliminates the need for a high quality receiver clock which significantly reduces the size, power and cost of LEO receiver instrumentation.

The penalty of a differencing scheme is the additional thermal and local multipath noise on each differencing link. Since thermal noise is uncorrelated between links, the rms thermal error increases as the square root of the number of links, assuming the thermal noise on all links is comparable. Increases in local multipath error scale similarly. In the future, substantial improvements are expected in the SNR of the orbiting receiver measurements (Section 3.3). It is therefore important in differencing schemes 1 or 3 that similar substantial improvements be made to the set of differencing receivers such that their SNR's are comparable to the SNR's of the flight receiver measurements or else the improved flight receiver SNR will have been wasted.

As mentioned in Section 3.4, another penalty is introduced if the second receiver in schemes 1 and 3 is ground-based, caused by phase variations incurred during signal propagation through the troposphere on links B and D. This tropospheric noise is roughly equivalent to that of a quartz oscillator with 10^{-13} fractional frequency stability at 1 second integration time [Treuheft and Lanyi, 1987; Linfield, 1996] and is generally small compared to other sources of noise (Section 3.4).

3.9 Refractivity Constant Uncertainties and Non-ideal Behavior

The accuracy of the relationships between atmospheric refractivity, density, pressure, temperature and water vapor as defined in eq. (2.3.1) and the ideal gas law will be examined. Compressibility and non-ideal behavior in the equations of refractivity and state lead to corrections up to 0.09% at the surface which decrease roughly in proportion to pressure at higher altitudes. The corrections should be included to achieve sub-Kelvin temperature accuracies expected in the upper troposphere (figures 3.7 and 3.10). In the

future, with a more general nonspherical retrieval process, some improvement in the refractivity equation and constants may be desirable.

Equation of refractivity

The accuracy of variables derived from refractivity depends on the accuracy to which the relationship between refractivity, pressure, temperature and water vapor is known. Eq. (3.9.1) gives a more general expression for atmospheric refractivity than eq. (2.3.1). It contains three terms representing dry gas polarizability, water vapor polarizability and the permanent dipole moment of water vapor, and includes effects of compressibility [Thayer, 1974]:

$$N = 77.60\left(P_a/T\right)Z_a^{-1} + 64.8\left(P_w/T\right)Z_w^{-1} + 3.776x10^5\left(P_w/T^2\right)Z_w^{-1} \qquad (3.9.1)$$

where P_a and P_w are the partial pressure of the dry gas and water vapor in mbar, T is temperature in Kelvin and Z_a and Z_w are compressibility factors given by Owens [1967] as

$$Z_a^{-1} = 1 + P_a\left[5.79x10^{-7}\left(1 + 0.52/T\right) - 9.4611x10^{-4}t/T^2\right]$$

$$Z_w^{-1} = 1 + 1650\left(P_w/T^3\right)\left[1 - 0.01317\,t + 1.75x10^{-4}\,t^2 + 1.44x10^{-6}\,t^3\right]$$

where t is temperature in Celsius.

The uncertainty of the dry refractivity term in eq. (3.9.1) is 0.02% [Thayer, 1974]. The *additional* uncertainty from the approximations in eq. (2.3.1) for dry air is <0.06% (<0.03% above 500 mbar), due to neglect of the compressibility factor.

The uncertainty of the wet refractivity terms is approximately 1%, due to the difficulty of measuring water vapor density. Use of eq. (2.3.1) introduces a fractional error in wet refractivity of 1% (T-359)/300 due to the simplification of two terms and < 0.4% due to the neglect of the compressibility factor. The error in retrieved humidity will

be < 2% due to errors in the refractivity formula that is used. Since occultations may ultimately yield near-surface, tropical humidities accurate to ~1% [Kursinski *et al.*, 1995], an improved knowledge of the wet refractivity constants in eq. (3.9.1) may become desirable.

Equation of state

The van der Waals' equation of state which includes a first order representation of non-ideal behavior is given as

$$\left(P + n^2 a\right)\left(1 - nb\right) = n R T \tag{3.9.2}$$

where n is number density in moles/m^3, a = 0.0014 m^6 mbar/$mole^2$ and b = 3.75x10^{-5} m^3/mole. The values of a and b represent weighted averages of the important atmospheric constituents, nitrogen, oxygen, argon and water.

Derivation of density, pressure and temperature from refractivity

Solving for n as a function of P and T (ignoring the cubic term in eq. (3.9.2)) yields

$$n = \frac{P}{R T}\left[1 + \frac{P}{R T}\left(\frac{a}{R T} - b\right)\right] \tag{3.9.3}$$

Neglecting the effects of moisture, which are discussed in Section 3.13, eqs. (3.9.1) and (3.9.3) can be combined to relate n as a function of N and T

$$n = \frac{N}{R b_1}\left[1 + \frac{N}{R b_1}\left(\frac{a}{R T} - b - b_a R T\right)\right] \tag{3.9.4}$$

where b_a is 5.79x10^{-7} $mbar^{-1}$ and b_1 is 77.60 N-unit K/mbar. The fractional difference in the densities defined by eqs. (3.9.4) and (2.3.2) is

$$\frac{\delta n}{n} = \frac{N}{R\,b_1}\left(\frac{a}{R\,T} - b - b_a R\,T\right) \tag{3.9.5}$$

To evaluate the error defined in eq. (3.9.5), a linear temperature gradient, $\dot{T}$, will be assumed with

$$T = T_0 + \dot{T}\,(z\text{-}z_0), \qquad \frac{P}{P_0} = \left(\frac{T}{T_0}\right)^{-mg/R\dot{T}} = \left(\frac{T}{T_0}\right)^{-c_0}, \qquad n \approx \frac{P_0}{R}\,T_0^{\,c_0}\,T^{-c_0-1} = c_1\,T^{-c_0-1}$$

such that the density error is

$$\delta n = n^2\left(\frac{a}{R\,T} - b - b_a R\,T\right) = c_1^2\,T^{-2c_0-2}\left(\frac{a}{R\,T} - b - b_a R\,T\right) \tag{3.9.6}$$

Hydrostatically integrating the density error in eq. (3.9.6) determines the pressure error, δP. The change in pressure error between two heights, z_1 and z_2, is therefore given as

$$\delta P_2 - \delta P_1 = R\,c_0\,c_1^2\,T^{-2c_0-2}\left(\frac{a}{R\left(2\,c_0+2\right)} - \frac{b\,T}{\left(2\,c_0+1\right)} - \frac{b_a R\,T^2}{2\,c_0}\right)\Bigg|_{T_1}^{T_2} \tag{3.9.7}$$

The temperature error, δT, is the difference between the true temperature, T, and the temperature, $\hat{T}_{\text{ideal}}$, derived from the ideal gas law. $\hat{T}_{\text{ideal}}$ contains errors because the density and pressure estimates are imperfect, and errors in converting pressure and density to temperature. Therefore the temperature error is

$$\delta T = T - \hat{T}_{ideal} = \frac{P}{nR}\left(1 + \frac{n^2 a}{P} - nb\right) - \frac{\hat{P}}{\hat{n}\,R} = \left(\frac{P}{nR} - \frac{\hat{P}}{\hat{n}\,R}\right) + \frac{P}{nR}\left(\frac{n^2 a}{P} - nb\right)$$

The first term in brackets on the RHS is the temperature error resulting from errors in pressure and density. The second bracketed term is the contribution from the non-ideal portion of the van der Waals equation of state. From eq. (3.2.6), the fractional error of $\hat{T}_{\text{ideal}}$ is the difference between the fractional pressure and fractional density errors.

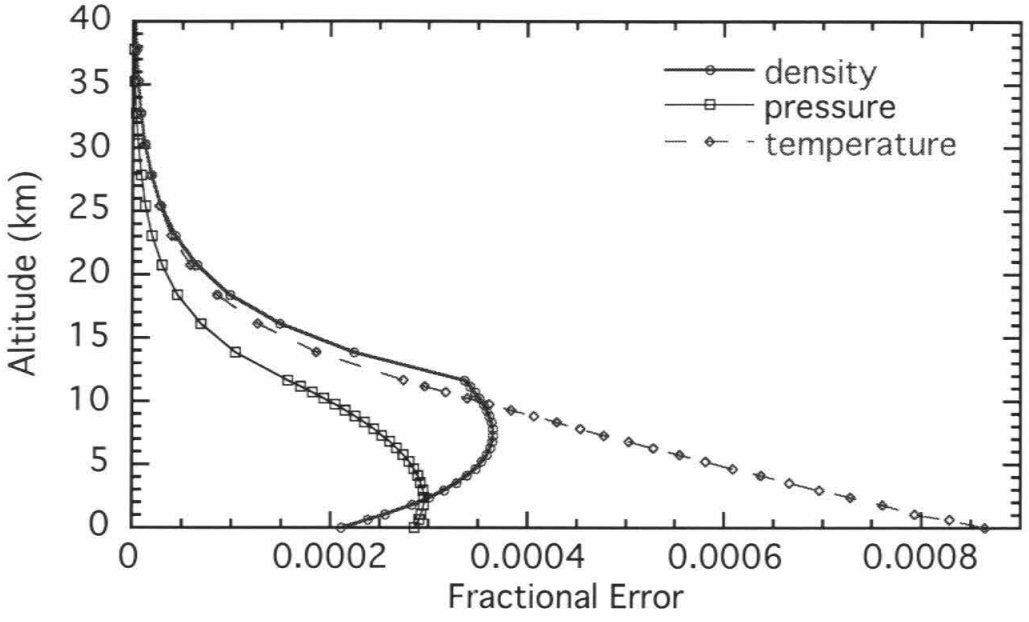

Figure 3.12 Fractional density, pressure and temperature errors due to approximations in refractivity equation [eq. (2.3.1)] and ideal gas law for conditions representative of mid-latitudes.

Therefore

$$\delta T = \frac{P}{nR}\left(\frac{\delta P}{P} - \frac{\delta n}{n}\right) + \frac{P}{nR}\left(\frac{n^2 a}{P} - nb\right) \tag{3.9.8}$$

Combining eqs. (3.9.8) and (3.9.5) gives

$$\delta T = \frac{P}{nR}\left(\frac{\delta P}{P} + b_a\, nRT\right) = \frac{P}{nR}\left(\frac{\delta P}{P} + b_a\, P\right) \tag{3.9.9}$$

figure 3.12 shows the density, pressure and temperature errors for a simple mid-latitude atmosphere model with a surface temperature of 288 K, a constant lapse rate between the surface and tropopause near 12 km ($T_{tropopause}$ ~ 215 K) and a constant lapse rate between the tropopause and stratopause at 50 km ($T_{stratopause}$ ~ 260 K). The density and pressure errors are less than 0.04%. The maximum temperature error of ~0.09% occurs at the surface and decreases rapidly to <0.04% at 10 km altitude. Errors under tropical conditions are similar. The errors in figure 3.12 are generally small in comparison with other sources of error. However, to derive temperatures accurate to tenths of a Kelvin in the upper troposphere (figures 3.7 and 3.10), the effects of compressibility and non-ideal behavior should be included to reduce errors in figure 3.12 by at least an order of magnitude.

3.10 Upper Altitude Boundary Errors

The upper integration limits of the Abel and hydrostatic integrals extend to an infinite altitude whereas the occultation observations do not. Therefore, high altitude density and bending angle structure used in these integrals must be estimated above some maximum altitude, z_u, below which the occultation data is used directly in the integrals. Incorrectly estimated atmospheric structure above z_u will produce errors in refractivity, density, pressure and temperature at altitudes below z_u which will decrease with decreasing altitude and can limit accuracy over altitudes several scale heights below z_u. It is therefore

desirable to start these integrals as high as possible. The tradeoff of balancing these errors and other errors described in this section determines the altitude at which to start the Abel and hydrostatic integrals and is discussed in the Section 3.14. High altitude climatological data will be used to characterize the variability and therefore the expected accuracy of estimated high altitude atmospheric structure.

Abel Transform Upper Boundary

The Abel transform integrates bending angle times a weighting factor as a function of asymptotic miss distance with an upper integration limit of infinity (eq. 2.2.2). To estimate atmospheric bending above the altitude, z_u, where the data becomes too noisy or biased to be used directly in the Abel integral, a simple, climatologically-based model can be used in which a few free parameters describe typical high altitude atmospheric structure for that latitude and season. The parameters can be estimated from a combination of measurements including the occultation observations themselves, models and climatology.

For the purpose of this discussion, a simple exponentially decaying model of $\alpha(a)$ above z_u will be used to characterize to first order the sensitivity of the refractivity and other retrieved variables to the high altitude atmospheric model. This model has two free parameters, bending angle at z_u ($\alpha(z_u)$) and bending angle scale height (H_α). Sensitivity of the retrieved variables such as refractivity can be characterized by varying $\alpha(z_u)$ and H_α. $\alpha(z_a)$ can be estimated accurately from the occultation data if errors are small relative to the atmospheric bending. From eqs. (2.2.6) and (2.2.7), it can be shown that bending angle error due to frequency measurement error is approximately

$$\delta\alpha = \delta f \, \lambda / V_\perp \qquad (3.10.1)$$

where δf is the frequency measurement error, λ is the GPS signal wavelength and $V_\perp$ is the ray path descent velocity in the absence of the atmosphere. The error in $\alpha(z_u)$ for several types of error is given in Table 4.

Based upon the U.S. Standard Atmosphere, a nominal 60 km altitude bending angle of 5×10^{-6} radians is used in Table 4. The magnitudes of local multipath and orbital velocity errors have been discussed previously in Sections 3.5 and 3.6. The thermal noise error in Table 4 corresponds to a one second SNR (signal power divided by thermal noise power) of 10^6 and an integration time of 2 seconds, roughly the time for the ray path to descend one scale height in the mesosphere, and includes amplification factors of 3 for the ionosphere correction and $2^{1/2}$ for clock differencing.

Under nighttime solar maximum conditions, where ionospheric electron densities are relatively small, the orbital velocity, local multipath, thermal and nighttime solar maximum errors are comparable yielding fractional bending angle and refractivity errors of order 1% or less at 60 km. Under daytime conditions during solar cycle maximum, incomplete removal of the ionosphere dominates the bending angle error at 60 km. The influence of the ionosphere error can be reduced by decreasing the altitude, z_u, at which the Abel integral is initiated with occultation data although as indicated in Section 3.14, 60 km is not far from optimum for minimizing overall errors near the tropopause. With the present ionosphere calibration scheme, this residual ionosphere error is also systematic and will bias derived refractivity.

The accuracy of the bending angle scale height estimate, H_α, depends on knowledge of atmospheric structure primarily over the 3 scale heights above z_u. Since bending angle scales in proportion to the vertical density gradient, the bending angle scale height is approximately the density scale height. The density scale height reflects the average temperature structure over several scale heights immediately above z_u. While the scale height can be estimated at z_u from the occultation measurements, the variations in the temperature structure above z_u are more difficult to estimate. Based on the 60 to 90 km climatology of Champion *et al.* [1985], the largest variations in the middle stratosphere to the mesopause regime occur during winter months in the Arctic and sub-Arctic regions

associated with sudden warming events. The variations in temperature between 60 and 90 km can be as large as +15/-30K (+7/-14%). Since variations at other latitudes and seasons are smaller, a value of 10 K will be used here as globally representative of rms temperature variations with respect to mean seasonal climatology at a given height and location. 10 K corresponds to a fractional scale height uncertainty of ~5% which causes a refractivity error at z_u of roughly 2.5%.

Based on the estimated errors in $\alpha(z_u)$ and H_α for z_u=60 km under solar maximum nighttime conditions, the scale height error will primarily determine the refractivity error associated with the upper boundary of the Abel transform. Under daytime solar maximum conditions, the uncertainty in $\alpha(z_u)$ due to residual ionosphere will exceed the scale height uncertainty, and the refractivity error at z_u=60 km will be ~7%.

Pressure and temperature errors

Hydrostatic equilibrium, integrated over a finite vertical interval, provides the *change* in pressure over the interval. Obtaining total pressure requires knowledge of the pressure at the upper boundary, z_u, of the hydrostatic integral. Fractional variations in pressure at mesospheric altitudes can be large because pressure at a given height depends exponentially on the temperature at all altitudes below. Since fractional temperature variations at heights near 60 km are generally smaller than the pressure variations, the pressure at z_u can be estimated using the density retrieved via the Abel transform and an estimate of temperature at z_u. The resulting fractional pressure error at z_u is

$$\frac{\delta P}{P}(z_u) = \frac{\delta T}{T}(z_u) + \frac{\delta \rho}{\rho}(z_u)$$

The temperature error term at z_u represents the *a priori* uncertainty in $T(z_u)$ and will be discussed shortly. The fractional density error at z_u (which equals the fractional refractivity error) is the error due to the Abel upper boundary error just discussed. As shown in figure

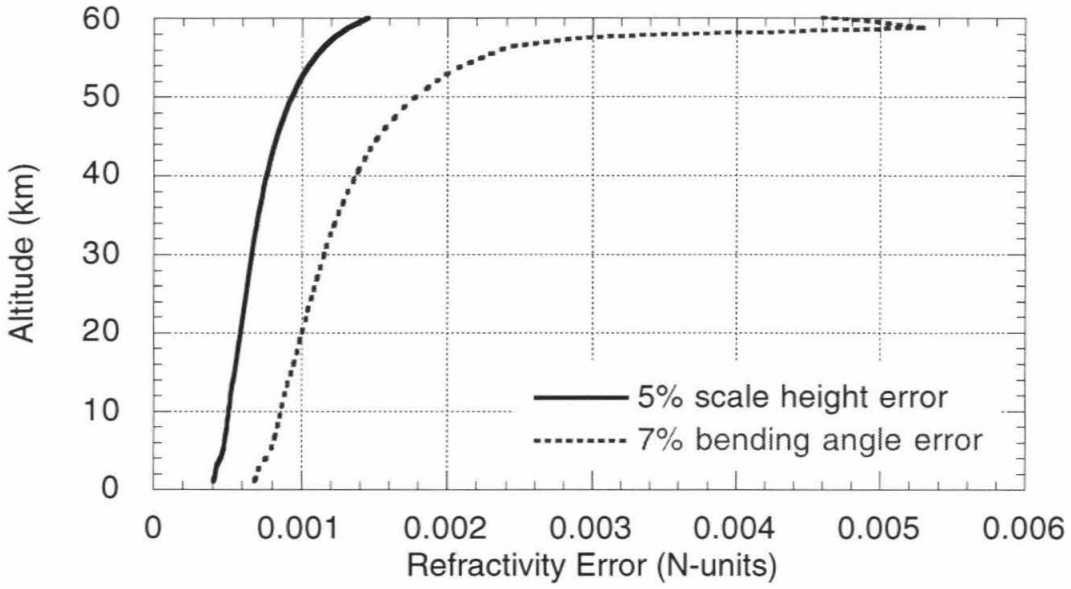

Figure 3.13 Refractivity boundary error in abel transform due to error in simple exponential bending angle structure extrapolated above 60 km, the height below which occultation data is utilized in the abel transform. Solid line: 5% in bending angle scale height, H_α. Dashed line: 7% error in α(60 km).

3.13, refractivity error due to the extrapolated bending angle model decreases with decreasing altitude below z_u and the pressure error therefore increases slowly over this same range. The fractional refractivity and pressure errors therefore decrease almost exponentially with decreasing height (figures 3.5, 3.6, 3.8, 3.9). The fractional temperature error contribution begins at zero at z_u and increases until it nearly equals the fractional pressure error.

Pressure and temperature errors due to initial temperature guess

The error in the temperature guess at z_u used to initiate the hydrostatic integral produces a constant pressure error. Therefore, the fractional pressure error caused by this boundary error decreases exponentially with decreasing height. Since there is no refractivity error associated with the error in the initial temperature guess, the corresponding fractional temperature error equals the fractional pressure error such that

$$\frac{\delta T(z)}{T(z)} = \frac{\delta P(z)}{P(z)} = \frac{\delta T(z_u)}{T(z_u)} \frac{P(z_u)}{P(z)} \cong \frac{\delta T(z_u)}{T(z_u)} \exp\left(-\frac{(z_u - z)}{H}\right)$$

where H is the average pressure scale height over the altitude interval, z to z_u. For example, a 2% or ~5 K temperature guess error at 0.2 mbar (~60 km) will produce a 0.004% pressure and temperature error at 100 mbar (~16 km) equivalent to a 0.25 m error in geopotential height and 0.01 K error in temperature. The impact of the initial temperature guess is therefore far more significant at high altitudes and it is desirable to raise z_u as high as possible to minimize its contribution at a given altitude.

3.11 Horizontal Refractivity Structure

Conversion of observed atmospheric Doppler shift into bending angle versus asymptotic miss distance and the subsequent derivation into refractivity versus height assume spherical symmetry in the vicinity of the ray path. This assumption works quite

well because the contribution to bending is focused around the ray path tangent point (figure 8). To assess the impact of horizontal refractivity variations, simulations using a 51 level, 40 km resolution, regional model have been performed. The results indicate horizontal variations in refractivity will limit retrieval accuracy from the lower stratosphere down to the lower troposphere The limiting rms errors in temperature and geopotential height are estimated to be a few tenths Kelvin and 7 to 8 m. Ultimately more general retrieval approaches will be developed which account for horizontal structure and improve upon the accuracy in the lower atmosphere.

Before discussing the simulation results, some intuition can be developed about the effect of horizontal refractivity variations on retrieval accuracy. Contours of constant atmospheric density would follow the geoid (defining a level of constant geopotential) in the absence of atmospheric dynamics. The geoid in turn can be approximated by an ellipsoid to an accuracy better than 100 m. The local center of curvature defined at the tangent to the ellipsoid can be used to interpret the occultation measurements. Therefore, it is informative to consider a radius of curvature error where the atmosphere is spherically symmetric but the assumed and actual centers of curvature differ. It can be shown that under these conditions, the values of a used in the Abel transform are incorrect resulting in an error in refractivity, δN, such that

$$\frac{\delta N}{N} \cong -\frac{\delta a}{2a}\left(1 + \frac{aN \times 10^{-6}}{H_N}\right) \qquad (3.11.1)$$

where H_N is the refractivity scale height [Appendix A.4]. From eq. (3.11.1), a fractional refractivity error of 10^{-3} requires a center of curvature error of order 10 km. Eq. (3.11.1) indicates that the difference between the equatorial and polar radii of the Earth would produce systematic fractional refractivity errors of order 2×10^{-3}. Of course, since the shape of the geoid is known to the meter level, the non-spherical shape can be accounted for. In

fact, an ellipsoid-based, local radius of curvature reduces this error by a factor of 100 or more.

Atmospheric dynamics alter the shape of constant density surfaces relative to the geoid. These distortions are generally small (but non-negligible) over the 200 km and larger horizontal scales relevant to occultation measurements. The largest horizontal refractivity variations and therefore refractivity errors occur in the warm, lower troposphere caused by water vapor whose horizontal variations are not constrained by hydrostatic equilibrium in the same manner as the bulk gas.

Error due to along-track horizontal refractivity structure

To characterize refractivity errors associated with realistic horizontal refractivity variations, retrievals have been simulated using both temperature and water fields obtained from a 40 km resolution NMC regional forecast ETA model [Janjic, 1990; Mesinger *et al.* 1988], modified by M. Zupanski for AIRS (Atmospheric Infrared Sounder) data simulation activities. The area covered by the model spanned roughly from Hawaii to the Great Lakes. Simulated measurements were created by raytracing signal paths from GPS satellites through the model atmospheric structure to a receiver in low Earth orbit. Profiles of refractivity were then derived from the measurements using the Abel transform. The retrieved refractivity error statistics of figures 3.5 and 3.8 show that rms errors are of order 0.2% above 500 mbar and increase rapidly at lower levels to approximately 1% associated with increasing horizontal humidity gradients. At the top of the tradewind inversion (~800 mbar), a layer roughly 200 meters thick can contain vertical refractivity gradients sufficiently large for critical refraction to occur. As discussed in Section 2.4, signal paths whose tangent heights lie within such a layer cannot emerge from the atmosphere. The current implementation of the retrieval scheme does not account for the bending angle singularity at the critical layer and does not derive refractivity accurately below critically refracting layers, so that only profiles without critical refraction are included in figures 3.5

and 3.8 below 800 mbar. As discussed in Section 2.4, the refractivity profile within and below the critically refracting layer will be accurately derived with an improved Abel transform implementation.

Pressure and temperature accuracy

figures 3.6 and 3.9 show the rms error in retrieved geopotential height. Again, the increase in rms error in the lower troposphere is associated with horizontal humidity gradients. RMS temperature error contributions are typically around 0.25 K above 400 mb (~7 km) increasing to ~2 K within 3 km of the surface (figures 3.7 and 3.10). The mean temperature error above 4 km is generally less than 0.05 K.

Horizontal motion during an occultation

Occultations are seldom diametric because the transmitter and receiver orbits are generally not co-planar causing the ray path tangent point to drift horizontally as it descends through the atmosphere. The horizontal drift of the ray path tangent point leads to an error in refractivity because the bending angle derived at higher altitudes is not the bending angle directly above the present tangent point. The ray path drift is generally of order the horizontal averaging interval or less and the resulting errors are therefore relatively small. The ratio of horizontal drift to vertical descent is greatest near the surface and the largest errors can therefore be expected there.

The bending angle error, $\delta\alpha(z,z_1,y)$, caused by horizontal drift of the ray path tangent point is given to first order as

$$\delta\alpha(z,z_1,y) = \Delta y(z,z_1,y)\,\frac{\partial\alpha}{\partial y}(z,y) \qquad (3.11.2)$$

where z is the height of the error, z_1 is the height of the lower integration limit of the Abel transform and the height of the refractivity retrieval, and Δy is the horizontal distance the ray path tangent point has drifted between heights z and z_1.

Horizontal drift

The motion of the tangent point is determined by the geometric motion between the transmitter and receiver relative to the Earth. For occultations where the direction to the transmitter is within $\pm 45^o$ of the velocity vector, the ratio of horizontal to vertical motion of the ray path tangent point relative to the limb of the Earth ranges from 0 to 3 at the top of the atmosphere. The number of daily occultations falling within $\pm 45^o$ of the receiver velocity vector for a high inclination orbit is 500 defining the number typically quoted as the daily occultations per orbiting receiver (another ~200 daily occultations occur at larger azimuths). As the ray path descends through the lower stratosphere, the vertical component of tangent point velocity slows as bending increases while the horizontal component of motion remains basically unchanged. Near the surface, the vertical descent velocity has decreased by an order of magnitude or more relative to the high altitude descent velocity. For example, in the isothermal case of figure 3.14, the ratio of horizontal to vertical displacement of the ray path tangent point ($\equiv$ dy/dz) is 2 at the top of the atmosphere, whereas at the surface, dy/dz ~ 30. In this example, the 10 km altitude ray path tangent point has drifted 140 km horizontally relative to its location at 60 km altitude, whereas over the last 10 km of descent to the surface, the ray path drifts an additional 160 km for a total drift of order 300 km, demonstrating the significant increase in horizontal drift near the surface.

Horizontal gradient of α

To simplify the discussion of the horizontal variation in α, α will be taken as proportional to the vertical density gradient such that $\alpha \sim c\ \rho/H_\rho$ where c is a scale factor, ρ

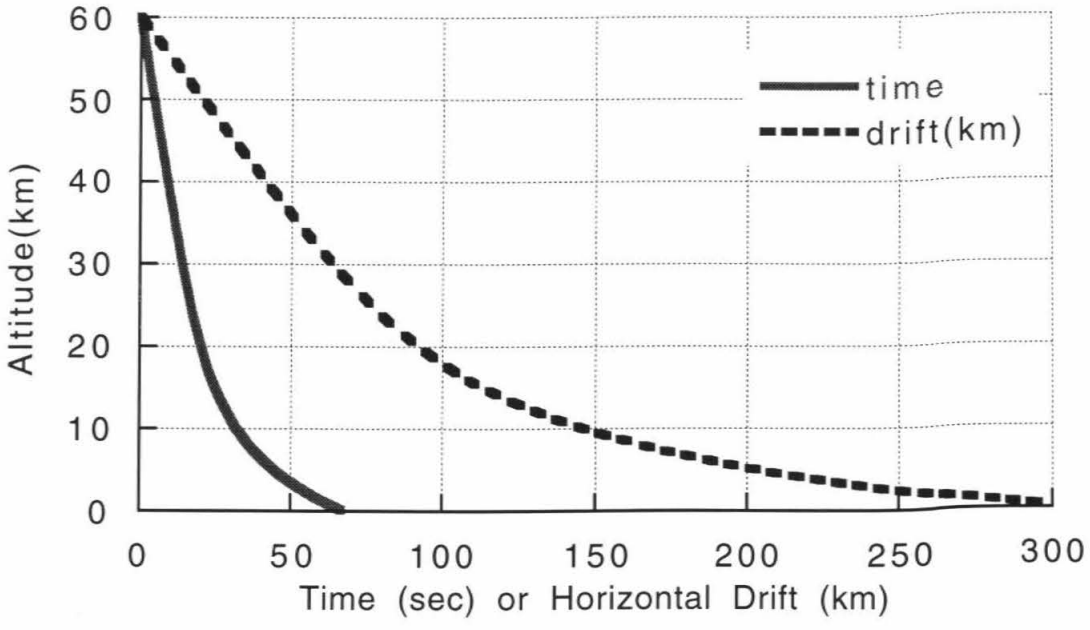

Figure 3.14 Ray path tangent height versus time (solid line) and versus horizontal location (dotted line) for an occultation located 30° in azimuth off the LEO velocity vector.

is mass density and H_ρ is the density scale height. This is a reasonable first approximation because the contribution to bending is very focused toward the ray path tangent region (figure 2.8). With this approximation,

$$\frac{\partial \alpha}{\alpha\, \partial y}(z,y) \simeq \frac{\partial \rho}{\rho\, \partial y}(z,y) - \frac{\partial H_\rho}{H_\rho\, \partial y}(z,y) \simeq \frac{\partial \rho}{\rho\, \partial y}(z,y) - \frac{\partial T}{T\, \partial y}(z,y) \qquad (3.11.3)$$

The density gradient can be characterized in terms of geostrophic and thermal wind balance. Taking the vertical gradient of the zonal geostrophic wind balance equation gives

$$\frac{\partial}{\partial z}(\rho u_g) = -\frac{1}{f}\frac{\partial}{\partial y}\frac{\partial p}{\partial z} = \frac{g}{f}\frac{\partial \rho}{\partial y}$$

from which the fractional meridional gradient of density can be related to the zonal geostrophic wind as

$$\frac{\partial \rho}{\rho\, \partial y} = \frac{f}{g}\left(\frac{\partial u_g}{\partial z} + \frac{u_g}{\rho}\frac{\partial \rho}{\partial z}\right) = \frac{f}{g}\left(\frac{\partial u_g}{\partial z} - \frac{u_g}{H_\rho}\right) \qquad (3.11.4)$$

Further, the meridional temperature gradient in eq. (3.11.3) is related to the vertical gradient of the geostrophic wind such that

$$\frac{\partial T}{T\, \partial y} = -\frac{H}{R}\frac{f}{T}\frac{\partial u_g}{\partial z} = -\frac{f}{g}\frac{\partial u_g}{\partial z} \qquad (3.11.5)$$

Eqs. (3.11.3) and (3.11.5) relate the horizontal gradient of bending angle to the geostrophic wind such that

$$\frac{\partial \alpha}{\alpha\, \partial y} \simeq \frac{f}{g}\left(2\frac{\partial u_g}{\partial z} - \frac{u_g}{H_\rho}\right) \qquad (3.11.6)$$

To estimate realistic and representative retrieval errors, the gradients have been derived from the climatology of the 25° to 40° south latitude band for the June-July-August period. This region contains some of the largest horizontal gradients and jets in both the troposphere and stratosphere. The climatological gradients of this region should exceed

typical atmospheric gradients and therefore yield generally conservative estimates of retrieved refractivity, pressure and temperature accuracy. On the other hand, these gradients may yield error estimates within a factor of two or three of severe weather conditions.

Between 25°S and 40°S in the troposphere, the peak zonal wind is ~35 m/s near 200 mbar and the vertical gradient of the zonal wind between the surface and 200 mbar ranges from 2 to 4×10^{-3} s^{-1} [figures 7.15 and 7.16 of Peixoto and Oort, 1992]. Of the two terms in the brackets of eq. (3.11.6), the vertical gradient dominates in the lower half of the troposphere and the scale height term becomes larger near the jet. The net magnitude of the fractional bending angle gradient is $\sim 5 \times 10^{-5}$ km^{-1}. In the stratosphere, the gradient grows from a value near 0 at 30 km to a value of 1.4×10^{-4} km^{-1} due primarily to the scale height term in eq. (3.11.6) and 100 m/s jet near the stratopause.

Shown in figure 3.15 are the refractivity, pressure and temperature errors which result from the ~30°S climatological conditions for a 2:1 horizontal to vertical displacement ratio at the top of the atmosphere. All fractional error magnitudes are of order 0.1% or smaller. The two jets at 10 and 50 km have magnitudes of 35 and 100 m/s. The stratopause jet, through the scale height term in eq. (3.11.6), dominates the error structure above 30 km. The peak in refractivity error and rapid reversal in sign of its slope near 8 km is associated with the change in sign of the gradient term in eq. (3.11.6) above and below the jet at 10 km. The rapid increase in the near surface error is associated with the relatively large horizontal drift in the ray path location there.

Pressure and temperature errors

The pressure error lags the refractivity error with a peak geopotential height error of about 6 or 7 m near 6 km altitude. In the 10 to 20 km altitude range, where the occultation observations are generally most accurate, the geopotential error goes through a relative

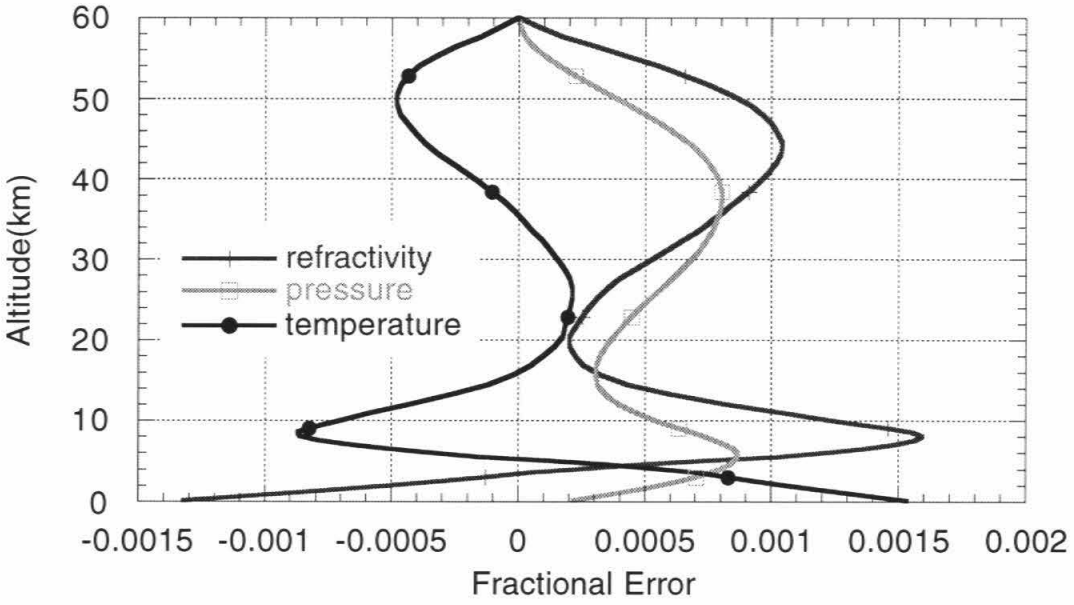

Figure 3.15 Fractional refractivity, pressure and temperature error resulting from horizontal drift of ray path tangent point during the occultation based on June-July-August climatologies near 30°S for the troposphere from Peixoto & Oort [1992], and stratosphere from Andrews, Holton & Leovy [1987].

minimum of 2 to 3 m and should not limit overall geopotential accuracy. Geopotential height errors in the stratosphere range from 0 at 60 km to a peak of ~7 m near 40 km. Peak tropospheric temperature errors of about 0.2 K and 0.5 K occur near 8 km and the surface respectively. Temperature errors in the stratosphere are generally less than 0.1 K.

Discussion

Since the dry horizontal gradients used to estimate the effect of horizontal motion of the ray path are representative of the largest climatological values found in the lower and middle atmosphere, the accuracy estimates should be generally conservative for dry regions. Errors under extreme weather conditions can be expected to exceed these estimates at least near the surface.

The horizontal structure of atmospheric water vapor will be important at low altitudes with a strong dependence on temperature. During warm tropospheric conditions in the summer hemisphere and the tropics, baroclinicity tends to be small on average and water vapor abundances large such that horizontal refractivity structure in the warmer regions of the lower troposphere will typically be dominated by water vapor. During winter when baroclinicity is often large, the temperature dependence of water vapor abundance tends to reduce refractivity contrast across winter cold fronts [Hardy *et al.*, 1994]. The refractivity of air on the cold side of the front has a higher dry density contribution due to the low temperature, whereas on the warm side, the dry density contribution is slightly lower due to warmer temperatures, but it is offset by the contribution of higher specific humidity. A more complete assessment of the influence of water is left as an area of future research as this and other data sets improve upon the presently limited knowledge of the distribution of atmospheric water vapor.

A final comment is that eventually, when multiple occultations are used in the retrieval process, the horizontal refractivity structure will be part of the solution and will

improve upon the accuracies estimated here to the extent that the density of occultations is sufficient to observe the horizontal structure. Similar statements can also be made regarding use of models in the retrieval/assimilation process and the horizontal resolution of the models.

3.12 Accuracy of the Geometric Optics Approximation

Except for a brief mention of diffraction during the discussion on resolution in Section 2.5, the geometric optics approximation has been assumed almost throughout. A new approach to representing diffraction in a planetary atmosphere is discussed and used to assess the error in the geometric optics approximation. The results shown here are based upon on a series of simulations representative of the sharp changes in temperature lapse rate which may be found around the tropical tropopause. The results of a comparison between diffraction and geometric optics approaches indicate that the geometric optics approach is accurate to several tenths of a Kelvin near the tropopause and above.

Generalized Huygens-Fresnel integral for EM propagation within planetary atmospheres

The Abel transform concept is based upon the concept of geometric rays which is very successful at describing propagation characteristics above a certain diffraction limit. However, the general case must deal with the wave nature of light and the effect of sharp structures on the propagation characteristics. Further, the accuracy limitations of the geometric optics approximation must be assessed.

Propagation of electromagnetic waves can be described via the Huygens-Fresnel concept where the signal at each location is a summation of spherical wavefronts at each point in space. For a monochromatic source, the complex signal at the point of reception is

$$U(r) = \int_S A \frac{e^{iks'}}{s'} \frac{e^{iks}}{s} K \, dS \qquad (3.12.1)$$

where A is signal amplitude, k is the wavenumber, s is the distance from the source to a point on the surface, S, and s' is the distance from that point on the surface to the reception location. K represents the directionality of the secondary signal emanating from the point on the surface. Eq. (3.12.1) describes propagation in an isotropic media. For the case of a planetary atmosphere, a more general formulation is needed which retains the conceptual framework of eq. (3.12.1) but generalizes the features of propagation. In eq. (3.12.1), s represents the optical path length between the source and a location on the surface which, when multiplied by k, represents the phase shift in radians along the path. s in the denominator represents the reduction in signal strength associated with a spherically-propagating wave. While straight lines satisfy Fermat's principle in an isotropic medium, geometric rays are the solution to Fermat's principle in an inhomogeneous planetary atmosphere. The general solution therefore replaces s and s' in the exponentials in eq. (3.12.1) with optical path lengths defined by raytracing from the source to locations on the surface, S, and from the surface to the reception point respectively. Attenuation of the signal amplitude within a planetary atmosphere must include defocusing due to differential bending estimated via raytracing as inversely proportional to the distance between adjacent rays.

$$U(r_r) = \int_S A\, G(r',r_r)\, e^{ikL(r',r_r)}\, G(r',r_t)\, e^{ikL(r_t,r')}\, K(r',r_r)\, dS \qquad (3.12.2)$$

where r_r, r_t and r' are the locations of the receiver, transmitter and a point on the surface, S, L(a,b) is the optical path length between points a and b, and G represents the modification of signal strength during propagation.

By symmetry, the obvious choice for a surface is a plane within the atmosphere which includes the center of curvature and the tangent point of the ray which is the solution for the given transmitter-receiver geometry. A more general approach includes additional surfaces between the source and reception point. The need for additional planes and the

accuracy of eq. (3.12.2) can be assessed by testing for convergence of the results as new planes are added.

Inclusion of ray optics within the Huygens-Fresnel framework provides a very general solution to propagation within planetary atmospheres and allows a wide range of phenomena to be explored. Unfortunately, it is numerically intensive because it requires both raytracing to establish atmospheric optical path lengths and attenuation and their subsequent integration in eq. (3.12.2).

Simulations near the tropopause

Roger Linfield at JPL has implemented eq. (3.12.2) complete with raytracing for spherically symmetric atmospheres generalizing previous work interpreting the occultation of a compact radio source by Venus. His implementation was used to investigate the effects of diffraction in the vicinity of the tropopause because of the sharp change in density scale height there. The dry density (and refractivity) scale height is given as

$$H_\rho = H\left(1 + \frac{H}{T}\frac{dT}{dz}\right)^{-1}$$

where H is the pressure scale height (RT/mg). The change in temperature gradient across the tropopause from a negative value in the troposphere to a positive one in the lower stratosphere above causes the density scale height in the overlying region to be smaller than that in the underlying region.

Forward modeling with diffraction was done with a series of intermediate 'planes' (with values a function of only one coordinate along these planes because of spherical symmetry) between the transmitter and receiver. Geometric ray tracing including defocusing determined the amplitude and phase at the first plane. Propagation between planes was accomplished in two steps. The first step consisted of geometric ray tracing from all points on a grid in one plane to all points on a grid in the second plane (i.e., a two-

dimensional set of traces). This ray tracing established the phase and amplitude on the first plane, as seen from the second plane. A Fresnel integration then determined the amplitude and phase at the second plane. At the height of the tropopause, use of one intermediate plane (near the true impact plane) was found adequate, based on its agreement with calculations for 3, 5 and 7 planes. For thicker atmospheric conditions (e.g., at lower altitudes in the troposphere), one plane may not suffice.

To characterize the error in the geometric optics approximation, the diffraction pattern in signal phase arising from sharp tropopause structures was simulated. The simulated phase was then passed through the retrieval process described in Sections 2.2 and 2.3. The retrieved structure was then compared to the input (model) temperature profile. The results of four sample cases are given in figure 3.16. Cases a and b each contain a single change in temperature lapse rate, with case b having the larger change. Cases c and d contain two and three lapse rate changes respectively with a separation of 0.5 km between the changes. The geometrically retrieved minimum temperature is systematically larger than the true minimum temperature because the sharp corner in the temperature structure is not resolved. The height of the minimum temperature in cases a and b is also slightly higher than the actual tropopause by 200 to 300 m. The temperature differences in the lower set of panels in the figure show the retrieval errors, which in all cases are $\leq$ 0.8 K and, except for the peak of the central error spike, are generally much less than that. The vertical extent of the error is mostly contained within $\pm$0.7 km (the radius of the first Fresnel zone) of the sharp structures. In case d, which is perhaps the closest to typical tropopause structure, errors are 0.2 K or less. The rms error over the 15 and 16.5 km altitude interval is roughly 0.1 K. A small positive bias less than 0.05 K also exists over this interval. The short vertical extent and small mean behavior of the effects indicate they will not produce an appreciable error in pressure.

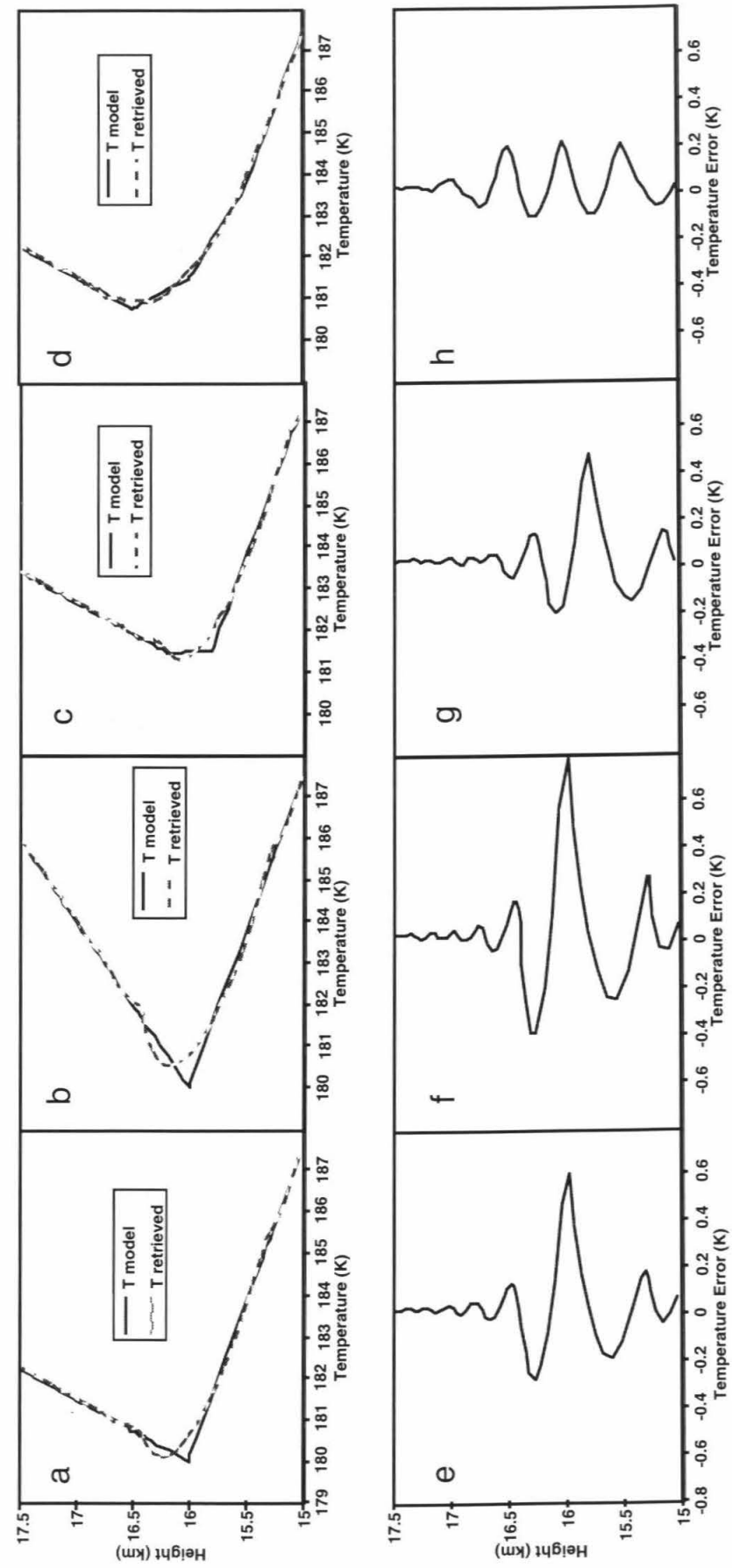

Figure 3.16 Errors in the geometric optics approximation under tropical tropopause-like conditions. Upper panels: 4 cases of model temperature (solid line) and geometrically retrieved temperature (dashed line) versus height. Lower panels: Temperature difference (retrieved minus model) versus height

The magnitude of the error when using the geometric optics approximation will depend on the sharpness of the tropopause. The results of figure 3.12 indicate errors due to sharp structure in the upper troposphere and above can be expected to be of order several tenths of a Kelvin or less. An investigation of lower troposphere diffraction effects, which will probably be larger due to water vapor structures, is planned. It is worth noting that the sharpness of changes in vertical thermal gradients at the tropopause will be apparent because it will produce a diffraction spike in the received signal amplitude as the signal path crosses the region. Such spikes are clearly present in some of the GPS/MET data indicating the presence of sharply varying vertical structure. These represent obvious regions where modeling of the effects of diffraction as mentioned in Section 2.5 would be warranted to retrieve the sub-Fresnel scale structure which apparently exists there.

3.13 Water Vapor Ambiguity

In the warm regions of the lower troposphere, water vapor contributes a large fraction of atmospheric refractivity at microwave wavelengths. Water vapor can therefore be derived from occultation observations, the accuracy of which was summarized in Section 2.3. The high specific humidity in these regions, also means uncertainties in humidity will adversely affect the accuracy of retrieved density, pressure and temperature. The effect of water vapor is a strong function of latitude and height such that low latitude derivations of temperature will be limited by uncertainties in water below 9 km whereas high latitude temperature accuracy will on average be better than 2 K down to the surface.

The influence of water vapor on the GPS occultation observations was realized almost with the conception of the idea [Yunck *et al.*, 1988]. When water vapor contributions to refractivity are significant, there are three equations (state, hydrostatic equilibrium and refractivity) and four unknowns (density, pressure, temperature and humidity). Therefore, additional information is required to separate the wet and dry contributions to refractivity.

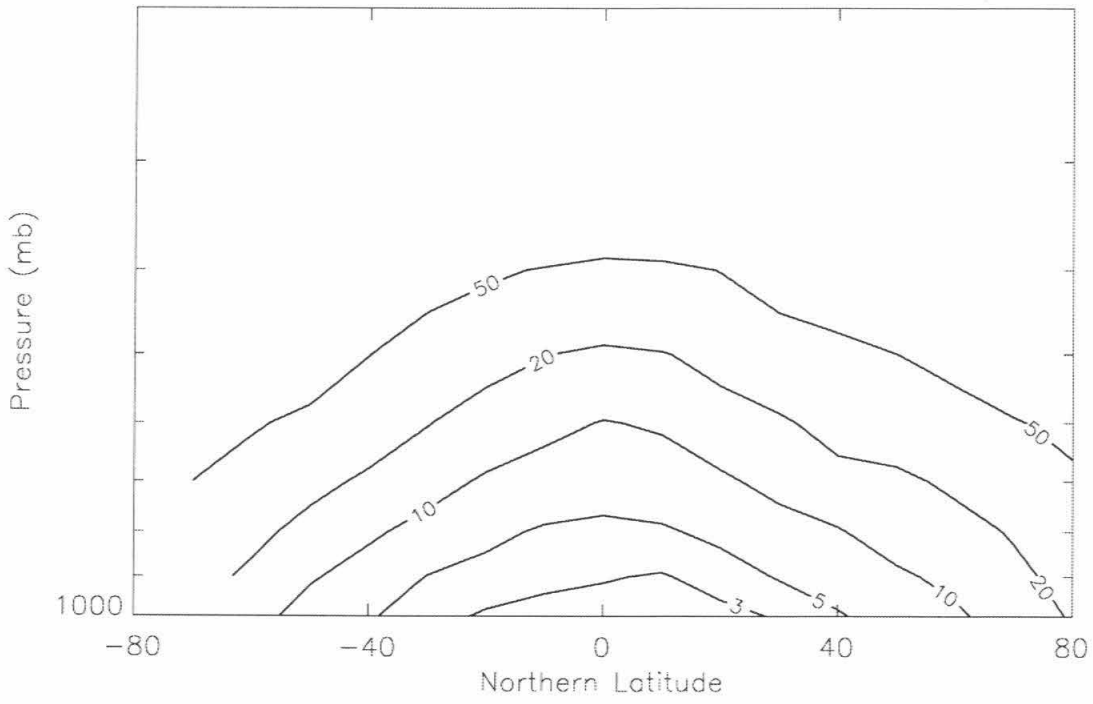

Figure 3.17 Dependence of B ($=b_1PT/b_2P_w$) on height and latitude based on the annual water vapor and temperature climatologies of Peixoto and Oort [1992].

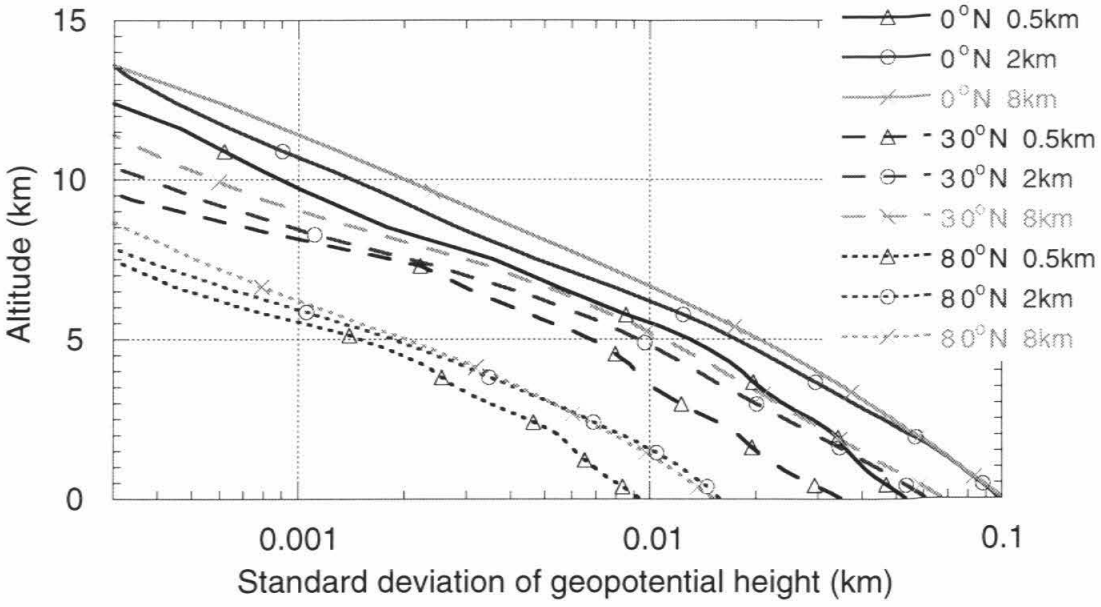

Figure 3.18 Geopotential height error due to 20% uncertainty in water vapor abundance based on annual climatologies for temperature and water vapor from Peixoto and Oort. Results are shown as a function of three latitudes, equatorial (solid line), 30°N (long dash) and 80°N (short dash) and three vertical water vapor uncertainty correlation lengths, 0.5 km (Δ), 2 km (o) and 8 km (×).

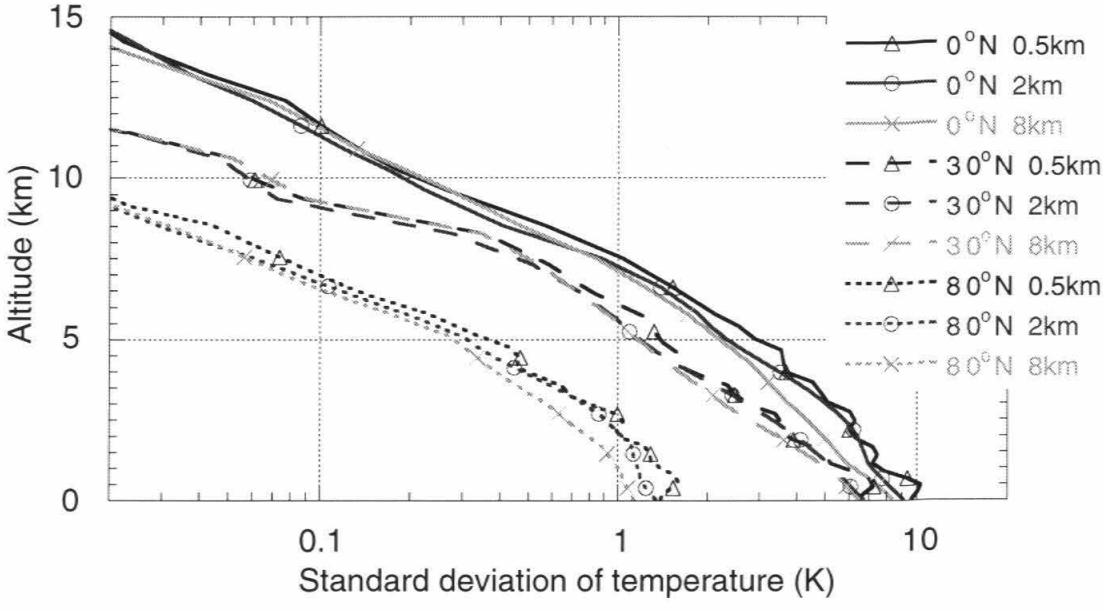

Figure 3.19 Temperature error due to 20% uncertainty in water vapor abundance. Conditions and legend same as figure 3.18.

Density, Pressure and Temperature Accuracy

The refractivity equation, eq. (2.3.5), can be differentiated to relate uncertainties in humidity, temperature and pressure as follows:

$$\frac{dP_w}{P_w} = (B+2)\frac{dT}{T} - B\frac{dP}{P} = 2\frac{dT}{T} - B\frac{d\rho}{\rho}$$

where $B = \frac{b_1 TP}{b_2 P_{wv}}$, and P_w is the partial pressure of water vapor. For the purposes of this section, refractivity is assumed to be known perfectly. B varies from a minimum value of ~2.5 near the surface in the tropics to of order 10^4 at the tropical tropopause and above because of the strong dependence of the water vapor mixing ratio on temperature. figure 3.17 shows the distribution of B with latitude and pressure up to 400 mbar based on the annual temperature and water vapor climatologies of Peixoto and Oort [1992]. The impact of water vapor is nonlinear, in fact quite nonlinear for small B, because of the coupling of density, pressure and temperature via the equations of hydrostatic equilibrium and state. Therefore, a Monte Carlo approach has been used to characterize water's impact. The uncertainty in water vapor is taken to be a percentage of the climatological water vapor abundance. Another important variable necessary to characterize water's impact on geopotential is the vertical correlation length of the water vapor uncertainty.

Figures 3.18 and 3.19 show the resulting uncertainties in geopotential height and temperature for latitudes of 0°, 30° and 80° spanning equator to pole. A 20% fractional uncertainty in water vapor is used, representative of uncertainties in the lower troposphere, but probably too small in the middle troposphere where 50% or even larger uncertainties may be more representative of present knowledge. Therefore, low latitude accuracies in figures 3.18 and 3.19 between 5 and 10 km altitude should be increased by a factor of 2 to 3 to be more representative of present uncertainties. Vertical correlation lengths in the uncertainty of the water vapor of 8, 2 and 0.5 km are shown, crudely corresponding to the

scale height of the background atmosphere, the water vapor scale height and small scale vapor layering related perhaps to clouds. Significant changes in the vertical distribution of water vapor can occur over still smaller intervals which can be extrapolated from the results in figures 3.18 and 3.19.

Discussion and Interpretation

The latitudinal and vertical dependence of accuracy in figures 3.18 and 3.19 reflects the strong temperature dependence of atmospheric humidity associated with the Clausius-Clapeyron relation. Low latitude errors are roughly an order of magnitude larger than those at high latitudes because low latitude air is moister than high latitude air by roughly this same factor. The results show that near-surface temperature errors at low latitudes will be 8 to 10 K for a 20% uncertainty in water vapor whereas at high latitudes errors will be of order 1.5 K or better.

The ~2 km scale height of atmospheric water (H_w), 4 to 5 times smaller than that of the bulk atmosphere (H), largely determines the vertical structure in figures 3.18 and 3.19. The small water scale height means that the impact of water increases dramatically at lower altitudes. Uncertainties in temperature and geopotential increase by a factor ranging from 3 to 10 over the bottom 5 km of the atmosphere. Assuming 50% uncertainty in middle tropospheric water vapor, 1K temperature accuracy will be limited to altitudes above about 9 km at low latitudes. Another result of the small size of H_w is the increase in fractional pressure error at lower altitudes lags the increase in fractional density error such that the ratio of the two is of order H_w/H or less. Therefore, from eq. (3.2.6), the fractional temperature error must nearly equal the density error and be much larger than the fractional pressure error. The fractional accuracy of pressure will therefore be better than that of temperature near the surface in warm, moist regions. For instance, at low latitudes, a surface error of 9 K (figure 3.19) is an error of ~3% error whereas a 100m geopotential

height error represents a ~1% error in pressure (figure 3.18). The difference is even more pronounced several kilometers above the surface.

figure 3.19 also demonstrates that fractional pressure and therefore geopotential accuracy depend on the correlation length of the water vapor uncertainty for lengths less than the water vapor scale height. Reducing the correlation length by a factor of 4 from 2 to 0.5 km reduces the surface geopotential error by roughly a factor of 2, consistent with a square root scaling expected from integrating a randomly varying error (i.e., a random walk process). The height below which the geopotential error exceeds 20m ranges from between 3.5 and 5 km at the equator and between 1.5 and 3.5 km at 30° depending on water vapor correlation length. RMS geopotential errors at 80° latitude are less than 20 m below 30 km altitude.

3.14 Summary of Refractivity, Density, Pressure and Temperature Accuracy

The refractivity, pressure and temperature errors discussed earlier have been assembled to establish an overall estimate of accuracy as a function of height, latitude, season and solar cycle. The predicted temperature accuracy is similar to the original prediction of G. Lindal [Yunck *et al.*, 1988] and subsequent update [Kursinski, 1994]. The analysis here also establishes the relative contributions and importance of individual sources of error as well as the vertical range of the retrievals. Because of the relevance to characterizing climate, biases and long term variations in these errors are discussed later in the section.

Three pairs of figures are shown, a pair each for refractivity, geopotential and temperature accuracy as a function of height. The accuracy represents the rms accuracy of individual retrieved profiles. Each pair of figures represents two extremes in conditions. The lower accuracy figure includes conditions representative of daytime solar maximum

conditions where the impact of the ionosphere is maximum, relatively low SNR equivalent to that achieved by the prototype GPS/MET mission ($SNR_0=5x10^4$) and low latitude tropospheric conditions for derived profiles of temperature and pressure where the contribution of water vapor to refractivity is maximum. The more optimal case represents nighttime solar maximum conditions, a relatively high, 1 second SNR of $5x10^5$ expected of future instrumentation and high latitude tropospheric conditions where the influence of water vapor is minimum. In each figure, the square root of the sum of the squares (RSS) of all error contributions is shown at each altitude level representing overall estimated accuracy. The orbit error and contribution of the troposphere on the differencing link discussed in Sections 3.4 and 3.6 are small and not shown in this set of figures in order to reduce clutter. Errors in the geometric optics approximation due to sharp vertical refractivity structure discussed in Section 3.12 are not shown because the frequency and statistical distribution of these structures is needed to create representative curves for figures 3.5-3.10. In fact, a global occultation data set may be the only data set capable of quantifying this distribution. A brief description of each figure summarizing its important features follows.

Figure 3.5, Refractivity: Daytime solar maximum, low SNR

The thermal noise at the level of the GPS/MET prototype receiver limits accuracy above 30 km altitude. As discussed in Section 3.3, the impact of thermal noise depends strongly on integration time and can be reduced at the expense of vertical resolution. The integration time used in figure 3.5 is the time for the ray path to descend a Fresnel diameter (~1.4 km). Increasing the integration time to descend through a layer 2.8 km thick would reduce the thermal contribution to the error by a factor of almost three bringing the thermal error down to a level comparable to the ionosphere errors. Horizontal structure limits accuracy below 30 km. The increase in refractivity error below 8 km is due to horizontal refractivity structure associated with higher water vapor concentrations at lower altitudes.

The altitude range over which the error is less than 0.4% (crudely equivalent to a 1 K error in temperature) ranges from roughly 5 km to 30 km.

Figure 3.8, Refractivity: Nighttime solar maximum, high SNR

Thermal noise and the 5% uncertainty in the high altitude scale height limit accuracy in the top 5 km. Local multipath limits accuracy between 35 and 55 km. Horizontal refractivity structure limits accuracy below 30 km. Fractional accuracy of 0.4% or better is predicted between 5 and 40 km.

Figure 3.6, Geopotential: Daytime solar maximum, low SNR, low latitude water vapor

Residual ionosphere error limits geopotential accuracy above 20 km. This error includes both the impact on the atmosphere model used above 60 km in the Abel integral discussed in Section 3.10 and the impact on derived refractivity below 60 km discussed in Section 3.7. The region between 8 and 18 km around the tropopause will be limited by horizontal refractivity structure to an accuracy of 8 to 10 m. At low latitudes, in the middle to lower troposphere, the uncertainty in water vapor will limit accuracy causing errors to increase to ~50 to 80 m near the surface. The fractional water vapor uncertainty used in this figure rises from 10% at the surface to 50% at 6 km altitude and above.

Figure 3.9, Geopotential: Nighttime solar maximum, high SNR, high latitude water vapor

The 10 K climatological uncertainty in temperature above 60 km and its impact on the Abel transform (Section 3.10) limits geopotential accuracy above 45 km. Local multipath limits accuracy over the altitude interval from 20 to 45 km. Horizontal refractivity structure limits accuracy below 20 km. The water vapor uncertainty represents the best case from figure 3.18 (Section 3.13) of high latitude concentrations with a small vertical correlation length of 0.5 km so the hydrostatic integral does not accumulate much pressure error. Under these conditions, geopotential height will be accurate to 20 m or

better below about 30 km. The highest accuracies of 8 to 10 m rms are predicted between 10 and 20 km altitude.

Figure 3.7, Temperature: Daytime solar maximum, low SNR, low latitude water vapor

As in figure 3.6, the ionosphere causes an erroneous high altitude atmosphere model impacting Abel transform accuracy at high altitudes. This error and the low SNR limit temperature accuracy above 40 km. As in the figure 3.5 discussion, doubling the integration time (2.8 km vertical layer) would reduce the thermal contribution to the error by a factor of almost three bringing the thermal error down to a level comparable to the hydrostatic and local multipath errors. Between 17 and 40 km, accuracy is limited by contributions from thermal noise, local multipath and residual ionosphere which has the largest influence. Between 9 and 17 km, horizontal refractivity structure limits retrieval accuracy. At equatorial latitudes, the uncertainty in water abundance limits temperature accuracy below 9 km altitude such that below 6 or 7 km, the occultation data is better used to derive water vapor than temperature (Section 4). As in figure 3.6, fractional water vapor uncertainty rises from 10% at the surface to 50% at 6 km altitude and above. Temperature accuracies better than 1 K are predicted between 8 and 30 km with accuracies of 0.3 to 0.4 K accuracy expected between 10 and 20 km.

Figure 20, Temperature: Nighttime solar maximum, high SNR, high latitude water vapor

Above 45 km altitude, several comparable sources of error influence derived temperature accuracy including uncertainties in high altitude structure which influence the Abel and hydrostatic integrals, thermal noise and local multipath. Local multipath dominates temperature accuracy between 30 and 45 km. Horizontal refractivity structure limits retrieval accuracy between 6 and 20 km. Below 6 km, uncertainty in water abundance dominates the error. Sub-Kelvin accuracy is predicted to extend from 40 km altitude to within 1 or 2 km of the surface.

Diurnal, seasonal and latitudinal dependence

A few additional comments can be made concerning the dependence of accuracy. In the lower troposphere, accuracy at high latitudes will exhibit a strong seasonal dependence associated with the large variations in temperature and therefore water vapor. The high latitude results shown in figures 3.8-3.10 represent an annual average so results during winter will be better, whereas results during summer will look more similar to the mid-latitude results in figures 3.8 and 3.9. The magnitude of the horizontal drift error in the lower troposphere should increase some at lower latitudes because of the increase in water vapor abundance and horizontal variations in water vapor abundance. Residual ionosphere error can be expected to exhibit some latitudinal dependence due to systematic latitudinal structure such as the equatorial and auroral jets.

Biases and slowly varying errors

To this point, the error analyses have been concerned primarily with rms errors of individual retrieved profiles. However, for climatic observations, biases and slow systematic variations in the errors are of concern. Sources of error potentially important to biases in profiles of refractivity, pressure and temperature retrieved in the upper troposphere are local multipath, horizontal refractivity structure, residual ionosphere and assumed water vapor for retrievals in the troposphere.

Local multipath: Errors due local multipath depend systematically on the spacecraft geometry and electrical properties in the vicinity of the receiver antenna as well as the occultation-viewing geometry. To the extent that the viewing geometry changes from occultation to occultation, the local multipath errors will have zero mean and therefore average down when regional and temporal averages are calculated. As mentioned in Section 3.5, directional antennas and modeling will reduce the multipath effect. Therefore

errors due to local multipath in spatially and temporally averaged occultation profiles acquired by a GPS receiving system designed to make climate observations should be at least an order of magnitude smaller than the errors in individual profiles estimated in figures 3.5-3.10.

Unmodeled horizontal refractivity structure: In Section 3.11, a set of ~120 profiles were simulated to estimate the sensitivity to non spherical refractivity structure. The average refractivity error of the 120 profiles is less than 5×10^{-4} and generally less than 2.5×10^{-4} for altitudes above 2 km (figure 3.20a). The average temperature error (figure 3.20c) is less than 0.1 K indicating a direct sensitivity to temperature at the retrieval altitude while the average geopotential height error of 1.5 m or less (figure 3.20b) indicates the accuracy to which the average temperature of the atmosphere below that height is measured. The error of 1 meter at the height of the tropical tropopause (~16 km) is equivalent to knowing the average temperature of the troposphere to ~0.02 K, a very sensitive measure of average temperature.

As discussed in Sections 3.11 and 2.4, a sharp refractivity gradient can exist at the tradewind inversion typically at 1.5 to 2 km altitude over the ocean which is sufficiently large to cause critical refraction (Section 2.4) causing problems for the present numerical implementation of the Abel transform and results at and below 2 km are therefore not shown in figure 3.20. As discussed, a more general numerical implementation of the Abel integral will be developed which will significantly reduce this problem.

Persistent dynamical structure will distort surfaces of constant density (and therefore refractivity) relative to the geoid. For example, the strong winter anticyclone over Siberia which disappears by summer in favor of a low-pressure system north of the Indian sub-continent [Peixoto and Oort, 1992] will produce systematic seasonal biases in retrieved refractivity, pressure and temperature if not accounted for. Therefore, minimizing biases in derived refractivity caused by such structures requires that the curvature of refractivity

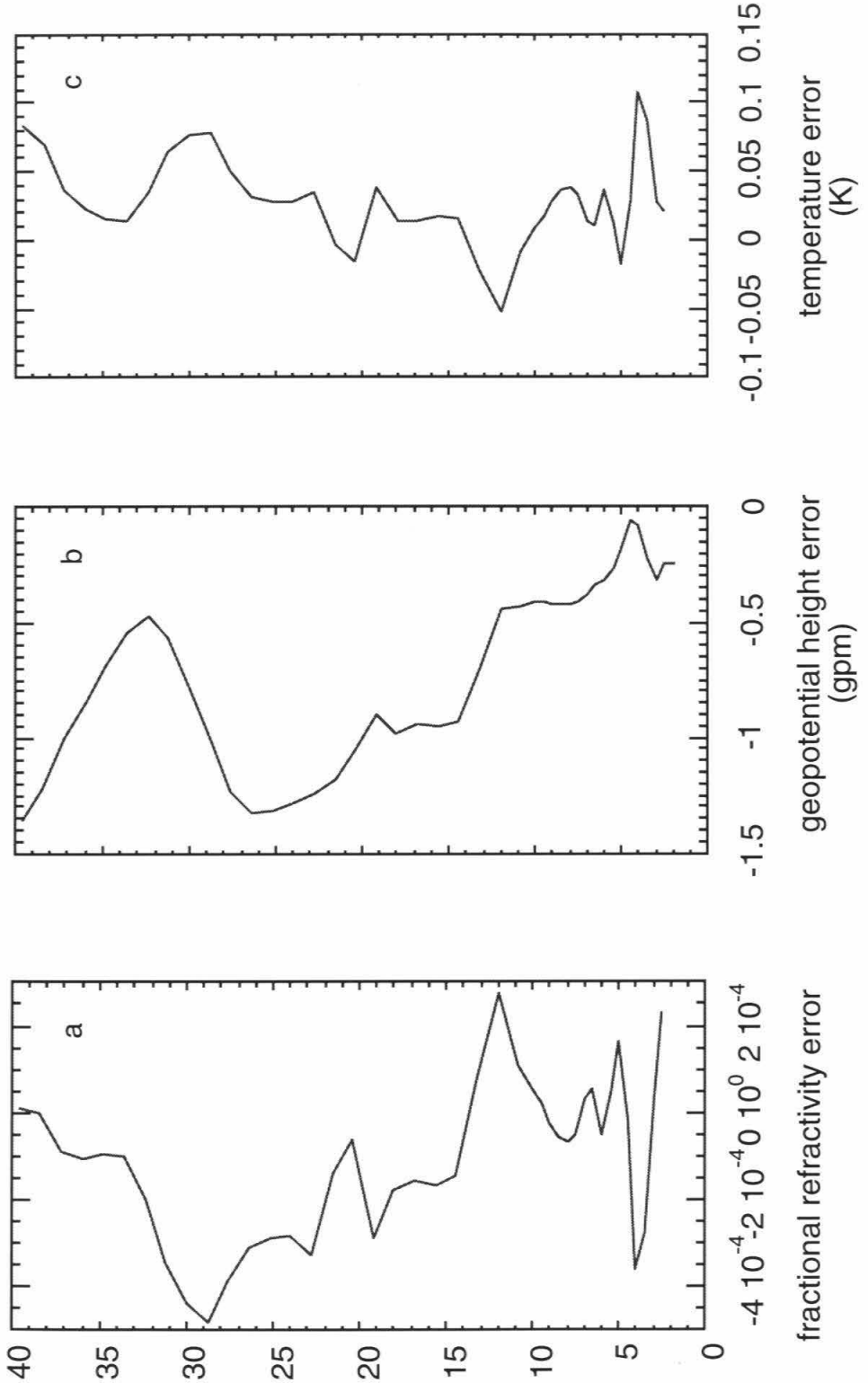

Figure 3.20 Averages of the refractivity, geopotential height and temperature errors in 120 simulated retrievals caused by horizontal, along-track refractivity structure discussed in Section 3.11.

surfaces used in the retrieval process include both the shape of the geoid and the shape of the climatological density contours (Section 3.11). If systematic horizontal structure becomes apparent in the averaged occultation profiles which differs from the climatology, then the climatology should be replaced by the derived systematic horizontal structure and the retrievals should be re-estimated until average retrieved refractivity and curvature of constant refractivity surfaces are self-consistent.

Ionospheric Errors: Incomplete calibration of the daytime ionosphere will probably be the most important source of error to long term climate studies because incomplete calibration will allow a small signature of the solar cycle to leak into the derived occultation profiles creating an apparent physical connection between the solar cycle and atmospheric structure.

The results in section 3.7 indicate the error in derived refractivity when using the ionosphere calibration scheme of Vorob'ev and Krasil'nikova [1993] is approximately independent of altitude, with a magnitude which depends strongly on the diurnal and solar cycles. The greatest systematic variations in error over the solar cycle will occur in daytime observations. Errors during nighttime, solar maximum and daytime, solar minimum conditions are smaller by more than an order of magnitude and insignificant. Daytime ionosphere refractivity error will vary by $\sim 10^{-3}$ and $\sim 10^{-4}$ at 30 km and 10 km altitude respectively over the solar cycle (figure 3.5). Retrieved daytime temperatures at 30 km and 20 km altitude will systematically shift by 0.8 K and 0.2 K over the solar cycle (figure 3.7). Geopotential will likely be the derived variable most affected by daytime ionosphere error because the relatively constant refractivity error with height will integrate to a large fractional error in pressure via the hydrostatic integral. The geopotential height error at 20 km altitude during daytime solar maximum conditions will be of order 10 m which is equivalent to a change in the average atmospheric temperature below 20 km of ~0.1 K,

which as a worst case error still represents a very sensitive measure of average temperature. At 10 km height, the geopotential error will be an order of magnitude less.

Higher order calibration schemes for ground based GPS observations reduce first order calibration errors by an order of magnitude. A higher order occultation calibration scheme which accounts for the vertical structure of the ionosphere (which can be derived from the occultation observations themselves) should reduce these errors by at least a factor of 3. A factor of 3 would reduce the daytime solar cycle signature to $3x10^{-4}$ at 30 km and $3x10^{-5}$ at 10 km representing a significant improvement in sensitivity to changes in climate.

Assumed water vapor: In deriving temperature and pressure estimates in the warmer regions of the troposphere, additional information about water vapor must be used. To the extent that the water vapor climatology is understood, the assumed water vapor should be consistent with the climatology and biases will then depend on how good the climatology is.

In summary, biases and slowly varying errors should be at least an order of magnitude smaller than the rms errors of figures 3.5-3.10 with the one notable exception that the variation in the residual daytime ionosphere over the course of the solar cycle will equal that in figures 3.5-3.7. The variation in daytime temperature error of 0.8K at 30 km over the solar cycle is not large and should be reduced by at least a factor of three with a higher order ionosphere correction. A long-term, database of nighttime observations should provide a still more sensitive set for studying long term variations in climate.

Caveats to estimated accuracy

Several caveats exist regarding the estimated errors and conclusions. First, as a reminder, the measurements are made in a limb viewing geometry so accuracies are relevant

to a sampling volume roughly 200 to 300 kilometers long and ~1 kilometer across and tall (Section 2.5). Caveats regarding removal of the ionosphere have already been discussed in Section 3.7 and will not be repeated here. Horizontal refractivity variations in the lower atmosphere may be larger than those in the 40 km resolution regional model used to derive the horizontal sensitivity in figures 3.5-3.10. Gravity waves with horizontal wavelengths ~100 km may impact accuracy depending on the frequency of occurrence and magnitude of these waves [Belloul and Hauchecorne, 1995].

The accuracy of refractivity retrieved within a few kilometers of the surface is complicated by the effects of diffraction, critical refraction and atmospheric multipath. The error due to horizontal drift of the ray path during the occultation estimated in Section 3.11 did not include horizontal variations of water vapor and therefore underestimates the error in the lower troposphere at low latitudes. Still it remains difficult to imagine that the errors due to horizontal water vapor inhomogeneity could approach a level comparable to the errors due to the water vapor uncertainty.

Future improvements

Improvements can be made such that accuracies will exceed those estimated in figures 3.5-3.10. As mentioned, signal to noise ratio can be increased with higher gain antennas and lower noise amplifiers improving refractivity and temperature accuracy at high altitudes and the ability of the receiver to track the signals at low altitudes. Better antennas will also reduce the magnitude of local multipath and smaller spacecraft will reduce the impact of local multipath by decreasing the phase rate of the multipath error (Appendix A.3). An improved ionosphere calibration scheme would substantially improve stratospheric retrievals and reduce systematic error during daytime conditions of the solar cycle maximum. Ultimately, reducing the impact of the ionosphere would best be accomplished with an additional, higher frequency signal. Sub-Fresnel scale structure may be retrievable by accounting for the effects of diffraction which will improve both

resolution and accuracy. Characterization of vertical structure at the 100 to 200 m level might then be possible within several kilometers of the surface where it would be particularly important.

Improvements beyond the already high accuracy predicted in the upper troposphere through the lower stratosphere can be accomplished by combining multiple occultations with other data sets and models, such as the ECMWF weather analysis model, in a grander scale retrieval or assimilation scheme which would consider and derive both vertical and horizontal structure [Gorbunov and Sokolovskiy, 1993; Eyre, 1994]. The higher density of soundings provided by a constellation of orbiting receivers would be very useful in this regard. Tomographic simulations of Gorbunov and Sokolovskiy [1993] indicate a factor of 10 improvement may be achievable. Based on figures 3.5-3.10, the improvement would primarily be in the troposphere and lower stratosphere.

Use of external information

The retrieval scheme described here uses the occultation data as much as possible, minimizing the use of independent information. Optimization under these conditions is a balance between the relative contributions of two types of error. The first type consists of the Abel and hydrostatic boundary value errors (Section 3.10) where the desire is to raise the upper boundary altitude as high as possible to reduce the impact of the errors at lower altitudes. The other class consists of all other error types where the desire is to lower the upper boundary altitude in order to reduce the accumulation of error, particularly the hydrostatically integrated pressure error, at lower altitudes. The optimum upper boundary altitude which minimizes the error at some lower altitude can be found by approximating the behavior of the dominant error terms with some simple analytical approximations and either directly calculating an optimal altitude or numerically varying the starting altitude to find the minimum error. Taking minimization of error near 100 mbar as the goal, 60 km is not far from optimum. 60 km is a few kilometers higher than the optimum upper boundary

altitude under the daytime, solar maximum and low SNR conditions of figures 3.5-3.7, and it is several kilometers too low under the better conditions of figures 3.8-3.10. This assessment assumes the 5% scale height uncertainty in the upper atmosphere model (Section 3.10) is independent of height.

The analysis here has minimized reliance on external information in the retrieval process in order to characterize the informational content of individual occultation soundings and the accuracy expected using an Abelian-based retrieval capability. External information has only been used at the upper boundary of the retrieval process. Despite the limitations of this approach, the resulting accuracy already exceeds that of present and many planned satellite sensors. For instance, sub-Kelvin temperature accuracy, a goal of NASA's Mission to Planet Earth, appears achievable over a significant fraction of the troposphere and stratosphere. A more optimal use of the occultation data can be achieved when the occultation data is combined with independent information with a weighting scheme consistent with the relative accuracies of the respective data types. In addition, accuracy will be improved when horizontal information available in weather analysis models is incorporated in the retrieval process [Eyre, 1994].

4. Accuracy of Water Vapor Derived from GPS Observations

4.1 Introduction

The need for high vertical resolution satellite observations of water vapor, providing global-scale coverage particularly over oceanic regions, has been discussed by Starr and Melfi [1991]. Water vapor distribution often varies sharply in the vertical dimension as evidenced in the many forms of stratiform cloud systems 10^2-10^3 m thick, and the sudden change in water vapor concentration between the boundary layer and overlying free troposphere. The lack of adequate data limits our ability to analyze or simulate important aspects of the global climate system. Various investigators have identified GPS radio occultation measurements as a possible source of atmospheric water vapor data [Kursinski et al., 1991, 1993; Gorbunov and Sokolovskiy, 1993; Yuan et al., 1993]. The GPS occultation observations offer high accuracy and a viewing geometry complementary to that of nadir viewing weather satellites. The work presented here provides the first global estimate of accuracy of tropospheric water vapor derivations from GPS occultations, calculated as a function of height, latitude and season. To obtain these estimates, it is assumed that independent pressure and temperature data from observations and meteorological analyses can be used to isolate the contribution of water vapor to the atmospheric refractivity profiles retrieved from occultation data. GPS occultation measurements are most sensitive to water vapor in the warmer regions of the troposphere, particularly in tropical regions where abundances are greatest, and their accuracy is limited primarily by uncertainties in the independent pressure and temperature data and errors in retrieved refractivity.

4.2 Water Vapor Accuracy in the Lower Troposphere

It is clear from the refractivity equation (eq. 2.3.1) that an ambiguity exists between the contributions of P, T, P_w and condensed water to refractivity. The contribution of

condensed water to atmospheric refractivity is small because its refractivity is only 25% (liquid) or 10% (ice) of that of an equivalent density of water vapor. Furthermore, condensed water abundances are a small percentage of vapor abundances. For example, very large liquid water abundances of 0.5 g/m^3, found at the top of stratiform cloud deck at the top of marine boundary layers [Paluch et al. 1992], contribute approximately 1.4% to total water refractivity. Therefore, ignorance of condensed water will cause a slight overestimate of water vapor of order 1% or less.

Water vapor profiles can be derived from refractivity profiles given independent estimates of T and hydrostatic equilibrium to constrain P. As discussed in section 2.3, P_w can be expressed in terms of P, T, and N to give:

$$P_w = \left(NT - b_1 P\right)\frac{T}{b_2} \tag{2.3.5}$$

The sensitivity of water vapor partial pressure to errors in pressure, temperature and retrieved refractivity can be assessed by differentiating and manipulating Eq. (2.3.5) to give:

$$\frac{dP_w}{P_w} = \left(B+1\right)\frac{dN}{N} + \left(B+2\right)\frac{dT}{T} - B\frac{dP}{P} \tag{4.2.1}$$

where $B = b_1 TP/b_2 P_w$. Clearly B, and from Eq. (4.2.1), the accuracy of water vapor profiles derived from measurements of microwave refractivity, depends strongly on water vapor mixing ratio. The variation of B with altitude and latitude is shown in figure 4.1. The scale height of B, H_B, is

$$H_B = \frac{H_P H_w}{H_P - H_w} \tag{4.2.2}$$

where H_P and H_w are the pressure and water vapor scale heights respectively. H_w can be estimated from the Clausius-Clapeyron equation to be approximately

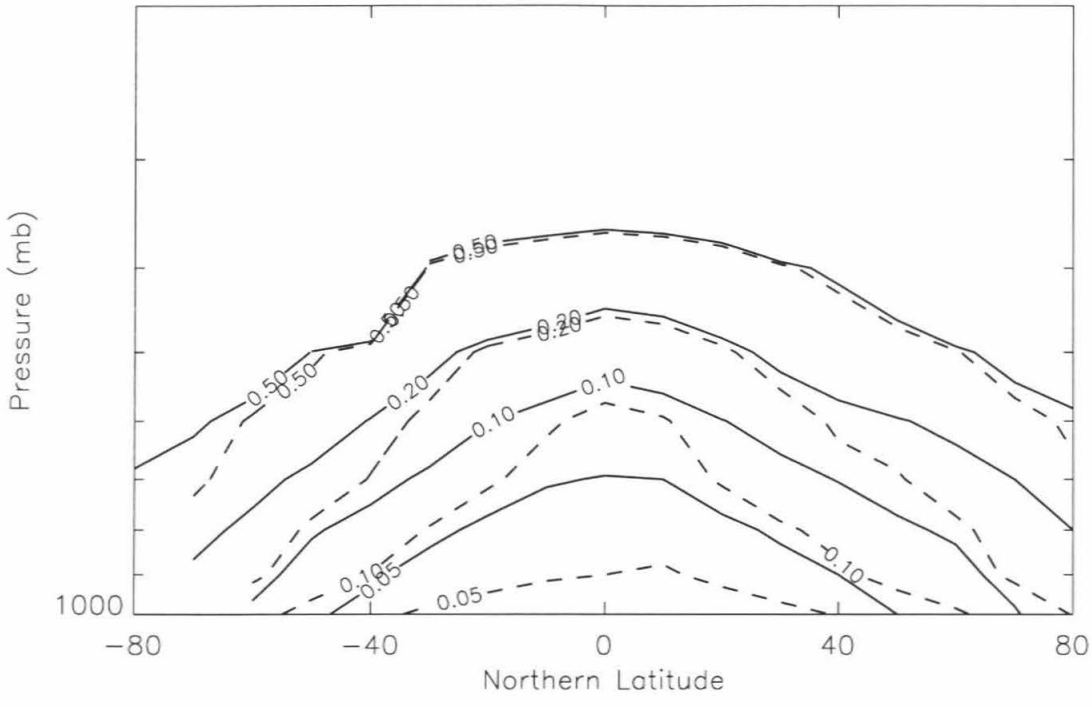

Figure 4.1 Estimated rms fractional error of individual derived water vapor profiles based upon the temperature and water vapor climatologies of Peixoto and Oort (1992). Solid lines contain the temperature and pressure uncertainty contributions only. Dashed lines include the lower troposphere refractivity error in Figure 18 as well.

$$H_w = \frac{T^2 R_d}{0.622 L \, {}^{dT}/_{dz}} \tag{4.2.3}$$

assuming constant relative humidity with height. Since relative humidity decreases with height on average, the true H_w will be somewhat smaller than the constant relative humidity estimate. The decrease in T with height will partially offset the decrease in relative humidity. For T ~ 230 K and H_B is 2.8 and 1.5 km for dT/dz = -4 K/km and -7 K/km respectively corresponding roughly to mid-latitude and low-latitude conditions in the upper troposphere.

Water vapor measurement errors can be derived from eq. (4.2.1) given knowledge of the accuracy of *N*, *P* and *T*. Pressure and temperature error contributions to eq. (4.2.1) are correlated because they are related hydrostatically. Fractional pressure errors at a given altitude consist of a fractional pressure error at some boundary plus an error due to the integrated vertical structure of fractional temperature error. For a surface boundary, the total fractional pressure error at height, *z*, can be written as

$$\frac{\delta P}{P} = \frac{\delta P_s}{P_s} + \int_0^z \frac{\delta T}{T} \frac{dz}{H} \tag{4.2.4}$$

where *H* is the pressure scale height and $\delta P/P$, $\delta P_s/P_s$ and $\delta T/T$ are fractional errors in pressure, surface pressure and temperature. The maximum net error in eq. (4.2.1) occurs when $\delta T/T$ varies randomly with height and the second term on the RHS of q. (4.2.4) contributes little relative to the temperature term in eq. (4.2.1). In this analysis, we therefore ignore the temperature error term in eq. (4.2.4) and assume an rms surface pressure error of 3 mbar.

The determination of the accuracy with which tropospheric temperature can be measured and modeled has received much attention. Passive remote sounding measurements give rms errors typically ranging from 1 to 3 K relative to radiosonde

measurements, depending on the instrument, the retrieval scheme, altitude, cloudiness and vertical temperature structure [Phillips et al. 1988; Kelly et al., 1991]. The rms difference between lower troposphere temperatures in weather model analyses (specifically 6 hour forecasts) and radiosonde measurements ranges from 1 to 2 K [Andersson et al., 1991; Kelly et al., 1991]. As radiosonde data are a primary input to the models, model accuracies in regions devoid of radiosondes will presumably be worse. In addition, radiosonde observations themselves are imperfect so estimates of accuracy made through comparison with radiosondes must be optimistic. The rms temperature error is assumed to be 1.5 K in this analysis.

To complete the water vapor error estimate, a retrieved refractivity error estimate is required. Random and systematic errors in refractivity are introduced by receiver thermal noise, clock instabilities, ionospheric correction residual errors, and the neglect of horizontal structure in the Abel transform retrieval. Based on several studies, the dominant source of error in the troposphere is expected to be the neglect of horizontal structure [Kursinski et al. 1993a,b; Gorbunov and Sokolovskiy, 1993; Vorob'ev and Krasil'nikova, 1993; Hardy et al. 1994]. To characterize refractivity errors, we have simulated retrievals using both temperature and water fields obtained from a 40 km resolution NMC regional forecast ETA model [Janjic, 1990; Mesinger et al. 1988], modified by M. Zupanski for AIRS (Atmospheric Infrared Sounder) data simulation activities. The simulation and results have been described previously in section 3.11 and are shown in figure 3.8 (error labeled "horiz along-track").

If we assume that the refractivity, temperature and pressure error contributions to the water vapor retrieval error discussed above are independent, Eq. (4.2.1) becomes:

$$\frac{\sigma_{P_w}}{P_w} = \left[(B+1)^2 \left(\frac{\sigma_N}{N}\right)^2 + (B+2)^2 \left(\frac{\sigma_T}{T}\right)^2 + B^2 \left(\frac{\sigma_{P_s}}{P_s}\right)^2 \right]^{\frac{1}{2}} \qquad (4.2.5)$$

where σ_a is the standard deviation of quantity, a. Contours of water vapor profile error, derived from eq. (4.2.5) for annual mean water vapor and temperature climatologies, are shown in figure 4.1.

One weakness in the assumed refractivity accuracy from the refractivity accuracy simulation is that only a single vertical dependence of accuracy is provided. There will be substantial latitudinal dependence as well. The simulation covers roughly from Hawaii to the Great Lakes in July. The dominant error in the lower troposphere is due to horizontal variations in refractivity associated with water vapor. At high winter latitudes, the dominant source of horizontal refractivity variations is synoptic scale weather disturbances which are much larger in horizontal scale and therefore will affect the retrieval of refractivity less than water vapor does in the summer hemisphere. Further, under near freezing conditions, water vapor and temperature contributions tend to cancel across a cold front. The warm side has more water but a less dense bulk atmosphere while the cold side has little water but higher bulk density and the two sides of the front have remarkably little variation in refractivity. The point is that the horizontal along-track refractivity error in figure 3.8 is an overestimate of the error in the winter hemisphere and the water vapor accuracies in the winter hemisphere will therefore be closer to the error due to temperature and pressure alone.

The water vapor error will be reduced when the error in assumed temperature has significant vertical correlation such that the temperature and pressure contributions of eq. (4.2.4) cancel to some degree. In the extreme case of a fractional temperature error which does not vary with height, eq. (4.2.1) becomes

$$\frac{dP_w}{P_w} = (B+1)\frac{dN}{N} + \left(2 + B\left(1 - \frac{(z_u - z)}{H}\right)\right)\frac{dT}{T} - B\frac{dP_u}{P_u} \qquad (4.2.6)$$

When z_u-z = H, the scale factor between temperature and water vapor error is reduced from B+2 to 2.

As an example, consider the following low latitude situation. Doubling the pressure errors in figure 3.18 to represent a 40% uncertainty in water vapor in the free troposphere, the low latitude rms pressure error will be 0.1% at ~9 km ($=z_u$). At 5 km altitude, one-half scale height below z_u, the rms refractivity error is 0.4% (figure 3.8). Assuming a temperature error of 1.5 K which does not vary with height and that the refractivity, temperature and upper boundary pressure errors are uncorrelated, eq. (4.2.6) yields a fractional water vapor error at 5 km of 7% roughly 30% smaller than the 5 km, low latitude water vapor error in figure 4.1 derived from eq. (4.2.5). Clearly, the accuracy of derived water vapor increases as the correlation length increases.

A final point to be drawn from figures 3.18 - 3.20 is the new information provided by the refractivity data, in the context of present uncertainties of temperature and water vapor, changes rapidly from dry density to moisture over a 1.5 to 2 km vertical interval. This interval is relatively short because the dependence of the vapor pressure of water on temperature as defined by the Clausius-Clapeyron relation is large. In the future, resolving the wet-dry ambiguity will no doubt make use of the constraint that the high spatial frequency variations of refractivity in the lower to middle troposphere are generally caused by variations in water vapor and retrieval accuracy will therefore be better than estimates made here.

4.3 Low Latitude Water Vapor Retrievals

In order to examine the potential of GPS occultations for low latitude water vapor retrievals given realistic temperature and pressure variations, 63 radiosonde soundings from Hilo, Hawaii, spanning the month of July 1991, have been used to construct a short, single location, sub-tropical climatology. To simulate errors in the derivation of humidity

from retrieved refractivity due to temperature and pressure fluctuations, given a good mean pressure/temperature profile climatology, profiles of water vapor were derived from refractivity profiles calculated for each radiosonde sounding using the mean of all 63 soundings for the temperature and pressure estimates required by Eq. (2.3.5). The mean relative humidity profile for the month and the standard deviation from the mean, measured by the radiosondes, are shown in figure 4.2a. The measured vertical structures of the standard deviation of pressure and temperature over the month are shown in figure 4.2b. Overall the temperature fluctuations are similar to the 1.5 K rms error assumed above while the surface pressure variation is much smaller than the 3 mbar used above. The rms fractional error in water vapor derived in this simulation is shown in figure 4.2c, both with and without the retrieved refractivity error of figure 3.8. The results of figure 4.2c are generally consistent with the low latitude estimates of figure 4.1 with sensitivities of 20% or better achieved up to altitudes of 6 to 7 km. Sharp error increases near 2 and 4 km altitude are caused by day-to-day variations in the altitude of the inversion at the top of the boundary layer. Sensitivities below the tradewind inversion suggest that water vapor abundances within the convective boundary layer could be derived to the 1 to 5% level provided accurate retrieved refractivities are available in this region. This simulation of water vapor accuracy is successful because of the low variability of tropical atmospheric structure, but is conservative in that model atmospheric analyses at low latitudes should represent pressure and temperature structure more accurately than simple climatological averages.

4.4 Climatological Applications

Despite the fundamental role played by water in weather and climate, an adequate climatology of atmospheric water vapor does not exist. Radiosondes presently provide most of the high vertical resolution profiles of humidity with a highly inhomogeneous, land-biased spatial distribution. In the lower troposphere, GPS observations deliver

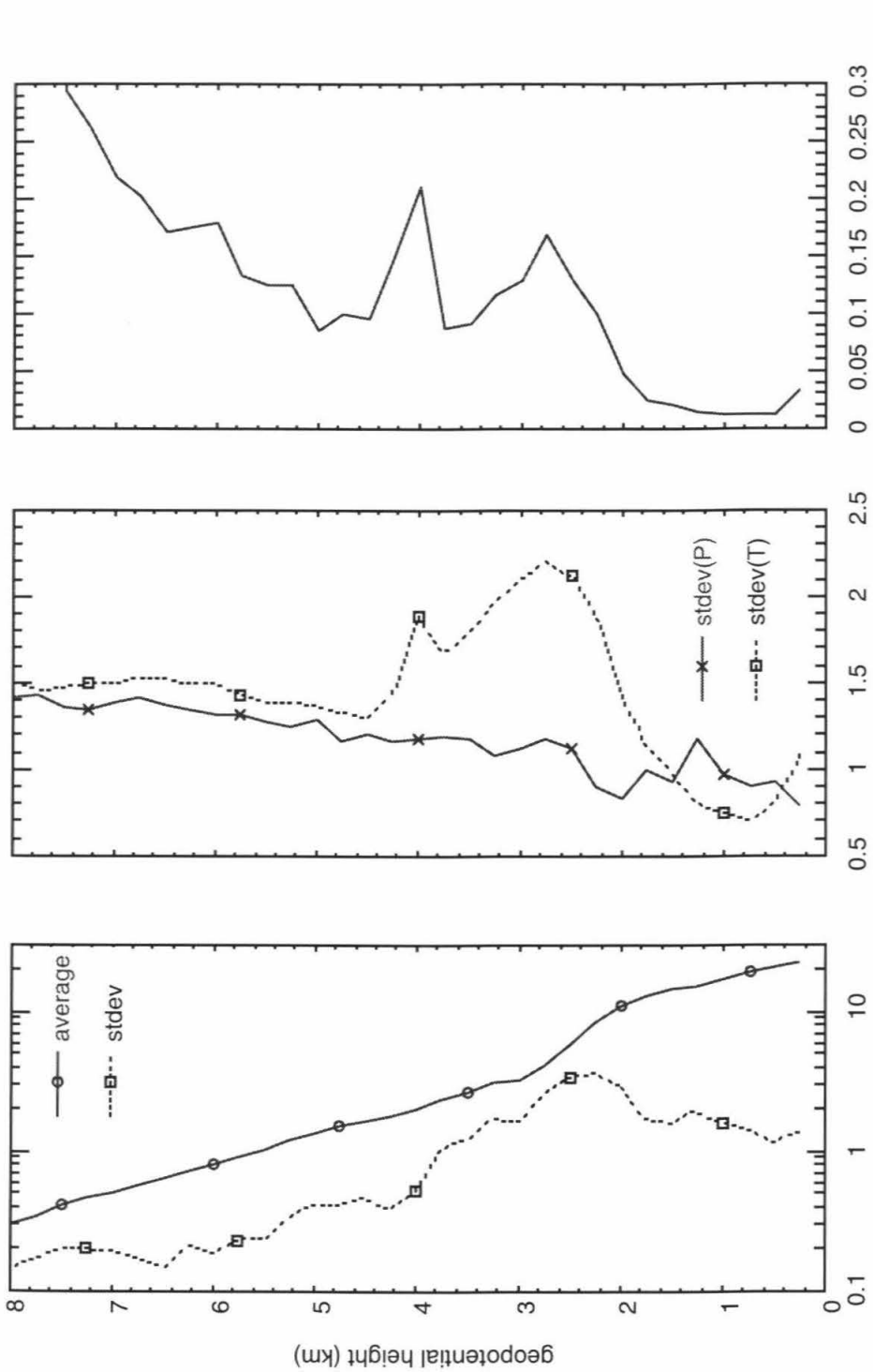

Figure 4.2 Estimated retrieval accuracy based on 63 radiosonde profiles from Hilo, Hawaii for July 1991. a: Average and standard deviation of radiosonde water vapor partial pressure data, b: Standard deviation of radiosonde temperature and pressure variations. c: Simulated accuracy of retrieved water vapor partial pressure. The solid curve shows errors due to temperature and pressure errors alone, whereas the dotted line also includes the refractivity error of Figure 18.

hundred meter to one km vertical resolution with global coverage and 100-200 km horizontal averaging. This vertical resolution lies between that of radiosondes and current satellite remote sounding instruments and should yield a significant improvement in the vertical scales observable globally. In addition, horizontal averaging produces profiles that are more representative climatologically than point measurements. Finally, insensitivity to particulates allows GPS occultations to measure humidity structure representative of the entire range of climatological variation by retrieving water vapor profiles below clouds. Changes in water vapor abundance at low latitudes in the lower and mid-troposphere have been identified as reliable indicators of modeled climate change [Santer et al., 1990; Schlesinger et al., 1990]. These are precisely the regions of greatest accuracy for GPS occultations.

Averaging GPS-derived humidities regionally and temporally to examine climatological behavior will reduce random errors and improve upon the accuracies summarized in figures 4.1 and 4.2c, although improvements will ultimately be limited by systematic errors in the temperature, pressure and refractivity data used in deriving water vapor via Eq. (2.3.5). For the tropospheric refractivity retrieval simulation presented in figure 3.8, the profile of mean error is much smaller than the standard deviation of individual profile errors at all levels, suggesting that regionally averaged refractivity structure may be more accurate than individual refractivity profiles. Refractivity error should average down to the extent that curvature errors are zero-mean. In the simulation discussed in section 3.11, the along-track error did in fact average down significantly roughly as the square root of profiles (figure 3.20). Therefore, refractivity error is not expected to be a limiting error source in general for climate applications. The caveat on this is the ability to retrieve refractivity to the surface.

Biases in the assumed temperature will probably limit the accuracy of climatological water vapor. These biases are difficult to quantify. Biases between NMC and ECMWF

model analyses are of order 0.5 K or less in the lower troposphere [Kasahara and Mizzi, 1992]. A comparison between radiosonde temperatures and those derived from the HIRS (infrared) and MSU (microwave) instruments on the NOAA-9 & 10 satellites indicate that biases at mid and low latitudes in the lower troposphere are generally 1 K or less [Flobert et al., 1991]. A factor of 3 improvement in accuracy gained from averaging, equivalent to a temperature bias of 0.5 K, would give specific humidity accuracies of better than 1.5% within the low latitude boundary layer and raise the 10% accuracy criterion for climatological utility by 3 km. These results would compare very favorably with the goals of 5 and 10% established by Starr and Melfi for the boundary layer and overlying troposphere respectively, and are generally conservative, provided the problem of retrieval through critical layers is solved.

4.5 Summary and Conclusion

We have presented a first order overview of the resolution and accuracy with which water vapor structure can be derived from GPS occultation observations. These observations are extremely sensitive to water vapor, but their accuracy is limited by errors in pressure, temperature and retrieved refractivity. Accuracies are a strong function of water vapor abundance and the best measurements will be obtained in the warmer regions of the troposphere. GPS observations will yield little information on humidity in the upper troposphere. Instead, they will be used to determine density, pressure and temperature structure. The best humidity data will be obtained at low latitudes in the middle and lower troposphere where individual profile accuracies of better than 5% in the boundary layer and 20% up to 6 or 7 km altitude may be achieved. Mid-latitude accuracies will depend upon season. At 45^{o} latitude, the altitude of 20% accuracy varies from approximately 5 km in the summer to 2-3 km in the winter hemisphere. For climatological investigations, averaging has the potential to improve upon these individual profile accuracies. Given its unique combination of good vertical resolution and insensitivity to clouds, this class of

observations will improve our knowledge of the vertical humidity structure of the warmer troposphere significantly, particularly over regions such as the oceans where radiosonde coverage is sparse or nonexistent.

5. Applications

In this section several areas of application for the atmospheric profiles derived from GPS occultation measurements are discussed and summarized.

5.1. Weather

Assimilating observations of important atmospheric variables into weather models and assessing their impact are major activities within a number of research agencies. Existing data assimilation techniques are easily adapted to exploit the full potential of occultation data to be provided by a suite of spacecraft carrying GPS receivers. Such techniques are ideally suited to ingesting the globally dense but irregularly spaced and timed occultation measurements. Comparisons of data acquired from the prototype GPS-MET mission with atmospheric analyses available every 6 hours from ECMWF have already demonstrated desirable properties for their use in numerical weather prediction (NWP), namely generally good agreement with a high-quality NWP analysis plus the ability to identify a minority of cases where there is room for significant improvement in the analysis [Kursinski *et al.*, 1996]. In effect, the GPS occultation system extends radiosonde-like measurements to the full globe. Even the relatively sparse sampling of 500 daily occultations from a single orbiting receiver will more than double the number of high vertical resolution profiles presently available in the southern hemisphere while providing a far more even distribution. Furthermore, occultation measurement horizontal and vertical resolutions are generally similar to the horizontal and vertical resolution used in global models.

Data assimilation systems can use the multiple dependencies of the occultation measurement (eq. 2.3.1) on temperature, pressure and water vapor directly and determine how to interpret tropospheric refractivity in terms of the moist and dry contributions. Studies evaluating the impact of GPS occultation observations on weather models have

either assimilated retrieved refractivity [Zou *et al.*, 1993; Kuo *et al.*, 1996] or bending angle [Eyre, 1994]. Bending angle assimilation uses a forward approach where model variables are adjusted until simulated observations generated via raytracing through the model match the observed bending angles. This approach, while more costly computationally, makes better use of the combined model and GPS data information and constraints.

Kuo et al. [1996] have examined the impact of assimilating refractivity on model extratropical cyclone prediction. Cyclogenesis is driven by vorticity anomalies at the upper and lower boundaries of the troposphere such that the high accuracy of temperatures and pressures derived near the tropopause is well-matched to the dynamical importance of the topography of the tropopause [Hoskins *et al.* 1985]. Kuo et al. examined the impact of refractivity below 3 km by removing it from the data assimilation. The result was a reduction in the rate of deepening of the cyclone in the forecast by more than a factor of two clearly demonstrating the constraints of refractivity on low-level temperature and latent heat to cyclogenesis. They also point out that constraints placed by refractivity are particularly important near the baroclinic wave because of the rapid growth of changes in initial conditions as the baroclinic wave develops. Assimilation of refractivity considerably improved the upper level potential vorticity anomaly, the primary upper level forcing mechanism for low-level cyclogenesis, due to the dynamic coupling between temperature, pressure and winds. The coupling is demonstrated by simple scaling analysis which indicates ~8-10 m geopotential accuracy (figure 3.9) is equivalent to 1% accuracy of the mid-latitude jet velocities determined over scales of the Rossby radius of deformation.

Since refractivity in the warmer regions of the lower troposphere tightly constrains water vapor [Kursinski et al., 1995], precipitation and moisture forecasts will undoubtedly improve as indicated in the simulation experiments by Zou et al. [1995] and Kuo et al. [1996]. Since condensed water content is usually a small fraction of total water content, GPS observations sense most of the total water and do so under any and all weather

conditions unlike IR measurements. Interestingly Kuo et al. found that the improvement in the moisture analysis and precipitation is largely lost 12 hours into the forecast, a duration which is short compared to water vapor's climatological mean time aloft of 9 days (≡ mean water content divided by mean precipitation). Since atmospheric water vapor should in some sense be correlated over the 9 days, the decorrelation over 12 hours implies the processes controlling atmospheric water within the model are less than optimum and will limit the impact of the GPS data on the forecast. Given the current state of models, assimilation of occultation data will at least improve the very short term precipitation forecasts.

Another point is that global weather analyses provide the best global state of the atmosphere and are therefore used in a myriad of other applications such as field campaigns, process studies and development of climatologies. By improving model analysis accuracy particularly in remote regions, GPS observations will indirectly influence many areas of research.

5.2. Climate

Accurate, long term data sets with good vertical resolution provide information necessary to characterize and understand climate. Perhaps the most obvious application of GPS occultation observations is monitoring regional and global atmospheric behavior on seasonal, interannual and decadal time scales, thereby providing an important segment of the data needed to evaluate climate models. GPS radio occultation observations were in fact originally conceived for the purpose of observing predicted climatic changes in temperature structure from the mid-troposphere through the stratosphere [Yunck *et al.*, 1988]. Like the solar occultation observations of SAGE, radio occultations have inherent features of self calibration, high accuracy and vertical resolution with the additional feature of insensitivity to atmospheric particulates. The expense of a radio occultation system is in building and launching the multiple transmitting satellites needed to provide global

coverage. The GPS has been designed and implemented as a long term capability implemented by the U.S. military and funded by U.S. taxpayers and can safely be expected to be available for producing a long term data set needed to observe and understand the behavior of our climate system over the coming decades.

Data used in the study of climate will be temporally and spatially averaged. The horizontal averaging inherent to the occultation limb viewing geometry is therefore a desirable feature in the context of climate research. The high accuracy of individual occultation profiles used in constructing regional averages reduces the number of soundings needed to achieve a specified level of measurement accuracy, thus allowing regional trends to be observed more quickly and longer term trends to be observed more precisely to the extent that measurement accuracy rather than natural variability dominates the uncertainty in identifying longer term trends and behavior.

Changes in propagation delay

To this point, discussions have centered around derivation and interpretation of refractivity, pressure, temperature and water vapor from radio occultations. Yuan *et al.* [1993] have suggested another approach utilizing GPS occultation propagation delay as a potential measure of climate change. Their simulations of changes in atmospheric propagation delay (estimated with a simple ray tracing scheme through the 9 layer NCAR community climate model) between model atmospheres with and without a doubling of CO_2 concentration were as large as 300 m near the surface. In warm regions, increases in delay were caused by increased water vapor abundances associated with warming. In cold, high latitude regions, decreases in delay were observed because near-surface warming caused bulk gas densities and therefore refractivities to decrease. These near-surface signatures are very large relative to the precision to which delay can be measured.

Identifying the signature of an anthropogenically-induced change in climate requires that the signature stand out against natural climatic variations. Unfortunately, the same nonlinear sensitivity which produces large changes in delay in response to modeled anthropogenic climate changes can also produce large natural variability. For example, figure 5.1 shows the variations in propagation delay versus ray path tangent height calculated by ray tracing through a set of 61 atmospheres whose vertical structure is defined by the July 1991 radiosonde data set from Hilo, Hawaii, discussed in Kursinski *et al.* [1995]. The rms variations in delay over the month are large, in fact larger than the predicted climatic change signature, a result which is somewhat surprising because tropospheric pressure and temperature variations over the month are small at 1 mb and 0.5 to 2 K.

Propagation delay of signals passing through the lower troposphere depends on vertical gradients of refractivity whose variability is fairly large, particularly that associated with the strength and height of the tradewind inversion typically found near 2 km altitude during the July 1991 period. For example, the standard deviation of water vapor just above the average height of the tradewind inversion is ~50%. The variability in figure 5.1 demonstrates the sensitivity of the occultation delay to large vertical refractivity gradients implying that climatic changes in the height and magnitude of systematically large vertical refractivity gradients may contribute a significant portion of long term trends and variability in the occultation delay observable. Therefore, to interpret observed changes in delay, climate models must have both sufficient vertical resolution to represent the large vertical gradients as well as adequate physical coupling between the gradients and the changes in radiative forcing which drive the purported changes in climate, posing an interesting challenge for the models.

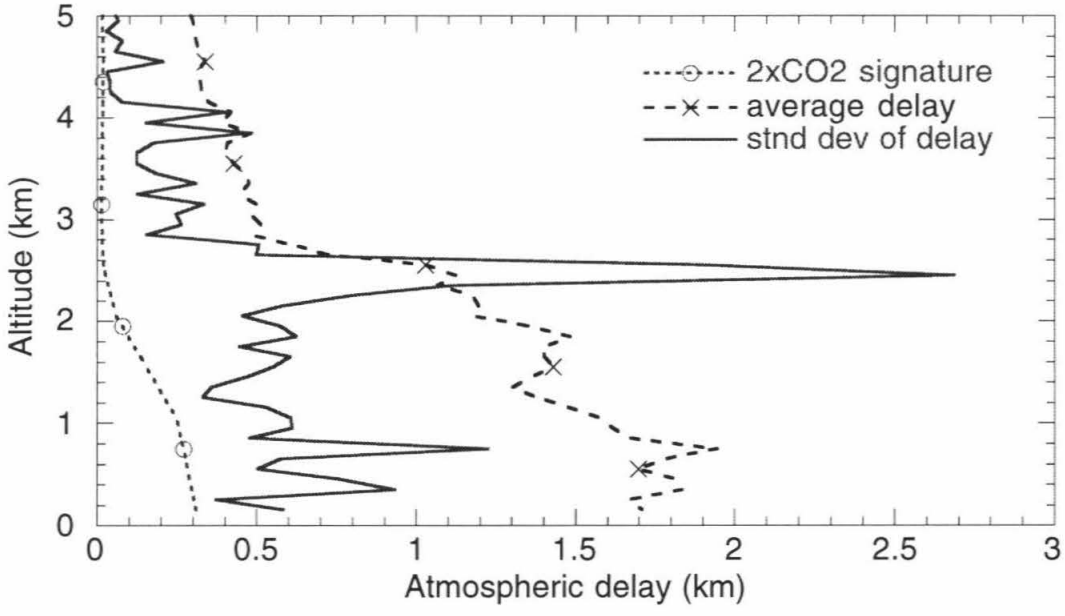

Figure 5.1 Atmospheric propagation delay at low latitudes: Comparison of change in delay at low latitudes due to a doubling of atmospheric CO_2 (short dashed line) as estimated by the NCAR community climate model (Yuan et al., 1993) with the simulated average delay (long dashed line) and standard deviation of delay (solid line) derived from the 63 radiosondes released from Hilo, Hawaii during July 1991.

Changes in Temperature near the Tropopause

Accurate, high vertical resolution temperature profiles of occultation measurements in the mid-to-upper troposphere and stratosphere provide an obvious tool for monitoring climatic variations complementing the very accurate but poor vertical resolution temperatures in this region and below produced by nadir viewing microwave radiometers [Spencer and Christy, 1990]. Temperature monitoring in the stratosphere is expected to provide a relatively straightforward test of cooling predicted in response to observed increases in greenhouse gas concentrations. At somewhat lower altitudes, profiles in the upper troposphere through the lower stratosphere can be compared with the contrast expected between warming in the troposphere and cooling in the stratosphere, a signature which has been proposed as a useful detection fingerprint of climate change [Epstein, 1982; Parker, 1985; Karoly, 1987, 1989] but may be limited by inadequate knowledge of the sources and magnitude of natural variability producing similar signatures [Liu and Schuurmans, 1990]. The altitude at which warming reverses to cooling differs significantly between models, and comparisons between recent observations and model simulations indicate the model altitude is generally placed too high [Ch. 8, Houghton *et al.*, 1990]. Trends reported by Angell [1988] indicate the upper troposphere may in fact be cooling. Results such as these are indicative of the need for accurate, long term and high vertical resolution observations in the tropopause regime.

Observations in the tropopause regime are well suited to observe the atmospheric response to volcanic eruptions. Major volcanic events inject aerosol material into the upper troposphere and lower stratosphere, producing a significant impulsive change in radiative forcing. The global coverage and accuracy of individual GPS soundings in the stratosphere and upper troposphere allow one to peer inside and observe the detailed spatial and temporal evolution of the thermal structure of the atmosphere as it adjusts to the changes in forcing, even in the presence of significant aerosol loading and clouds.

High latitude, near surface warming

Climate models generally predict enhanced near-surface warming at high latitude regions during winter months in response to a doubling of atmospheric CO_2 concentrations [Ch. 5, Houghton *et al.*, 1990]. The models also exhibit relatively large natural variability at high latitudes, implying early detection and interpretation of gradual climatic shifts in temperature may be difficult. Nonetheless, both the natural variations and the gradual climatic shift are of interest to understanding the behavior of the system, and profiles of atmospheric density pressure and temperature derived from the radio occultation observations at high latitudes will be very accurate and therefore useful under winter conditions.

Expansion of the troposphere due to global warming

Because the altitude of an atmospheric pressure level depends on the average temperature at lower altitudes, geopotential height offers a sensitive monitor of tropospheric warming [Gary, 1992], a task for which the accuracy and coverage of GPS occultations are well suited [Kursinski *et al.*, 1994]. Occultation-derived geopotential height is particularly well suited for characterizing climate because of its inherent spatial averaging in both the vertical (through hydrostatic equilibrium) and horizontal (through the limb-viewing geometry) dimensions.

A change in temperature, δT, over a pressure interval, P_{surf} to P, will alter the height of P in accordance with hydrostatic balance such that

$$\delta z \equiv z\, T(z) \left\langle \frac{\delta T}{T^2} \right\rangle \tag{4.1}$$

where z is the original height of P, δz is the change in z due to δT, and $\langle \delta T/T^2 \rangle$ represents the average value of $\delta T/T^2$ over the height interval, 0 to z. Since climate models generally

predict tropospheric warming and stratospheric cooling [Houghton *et al.*, 1992], the largest pressure height increase is predicted near the tropopause. In the tropics, where the height and pressure of the tropopause are ~16 km and 100 mbar, eq. (4.1) indicates a sensitivity of ~55 meters change per Kelvin change in average tropospheric temperature. At 300 mb, the pressure height sensitivity is roughly 30 m/K.

Daytime GPS observations of geopotential height in the upper troposphere will be limited by a systematic 8 to 10 m variation in geopotential height over the course of the solar cycle because of incomplete calibration of the ionosphere, an error source which can be reduced with a higher order ionosphere correction scheme (Sections 3.7 and 3.14). A long term record of geopotential height derived from *nighttime* occultations may be accurate to the 1 m level or better (Section 3.14), which, at the tropical tropopause, represents a sensitivity to fluctuations in average tropospheric temperature of order 0.02 K providing an extremely sensitive tropospheric thermometer!

Identifying anthropogenic changes in climate requires that the changes are distinct from natural climatological variations in geopotential height. Observed variability depends strongly on latitude and season with a distinct minimum at low latitudes [Peixoto and Oort, 1992]. Day to day and seasonal variations of tropical, tropospheric geopotential height are of order 25 m equivalent to fluctuations in average tropospheric temperature of order 1 K or less.

Two representative examples of equilibrium climate models, Hansen *et al.* [1984] and Manabe and Wetherald [1987], which both predict global average *surface* warming of 4 K, predict average *tropospheric* temperature increases of 4.6 K and 3.4 K respectively at low latitudes. The corresponding increases in the height of the 100 mb level of 260 m and 180 m respectively are large in comparison with measurement accuracy and low latitude natural variability. The IPCC average predicted warming rate of 0.3 K per decade

[Houghton *et al.*, 1990] predicts the tropical 100 mbar level will rise ~16 m per decade, a rate at which tropospheric warming could become apparent in 1 to 3 decades. The small observational errors, small natural variability and relatively large global warming signature predictions in the low latitude, upper troposphere are features characteristic of a "fingerprint" variable for diagnosing climate change and features useful in evaluating climate model realism and accuracy.

Changes in humidity

In the lower to middle troposphere, particularly in tropical regions, changes in water vapor abundances provide one of the most sensitive indicators of modeled climatic change [Barnett and Schlesinger, 1987; Santer *et al.*, 1990; Schlesinger *et al.*, 1990]. Water vapor profiles derived from GPS measurements will be most accurate in the lower troposphere at the tropical latitudes particularly in the marine boundary layer where expected accuracies below 7 km altitude are similar to goals established by the Global Energy and Water Cycle Experiment (GEWEX) [Starr and Melfi, 1991] and may therefore yield robust signatures of changing climate.

An understanding of the control of water vapor in the free troposphere is required to make accurate predictions of global warming because of the water vapor feedback effect. Present knowledge of water vapor in the free troposphere is poor. Sun and Lindsen [1993] have pointed out that tropical climatological humidities derived from radiosonde and SAGE observations differ by a factor of two [Oort, 1983; Rind *et al.*, 1993]. Tropical water vapor profiles derived from GPS observations cover the altitude regime where SAGE and radiosonde observations overlap offering an independent data set with sufficient accuracy and insensitivity to clouds (which limits the SAGE data set in particular) to improve the climatology significantly beyond that of present knowledge.

5.3 Stratosphere-Troposphere Exchange

Concern over decreasing ozone abundances in the stratosphere and their possible anthropogenic causes have focused research on understanding exchange between the stratosphere and troposphere. The tropopause is important in this context as the region across which exchange takes place. For instance, the extremely low water vapor mixing ratios observed in the lower stratosphere at tropical latitudes have been used to infer that tropospheric air is migrating upward into the stratosphere at these latitudes where the very cold temperatures of the tropical tropopause and strong temperature dependence of the saturation vapor pressure of water combine to "freeze-dry" the air during its upward passage [Newell and Gould-Stewart, 1981].

There has been much debate over whether the low water vapor mixing ratios require extremely cold temperatures associated with penetrative convection events or whether slow upward migration of air across the tropopause is adequate to explain the mixing ratios. Existing data sets are not adequate to resolve the issue. Radiosonde data has the necessary vertical resolution but coverage in the tropics is spatially very uneven. Further, the low densities of the tropopause affect radiosonde accuracy because of corrections required for effects such as radiation which depends on cloud cover, cloud top temperatures, solar zenith angle, surface temperature, vertical temperature structure and vertical distribution of aerosols, ozone, water vapor and carbon dioxide [Finger and Schmidlin 1991; McMillin *et al.* 1988]. Present estimates of the maximum magnitude of these errors are of order 1 to 3 K [Ahnert 1991; McMillin *et al.* 1988]. Research is under way to improve radiosonde data for climate applications [Luers and Eskridge, 1995]. In-situ airborne observations are extremely important but very limited in spatial and temporal coverage. Tropopause temperatures derived with nadir viewing passive sounders tend to be biased high because

of coarse weighting functions and the combination of a relative temperature minimum and approximately isothermal structure.

GPS occultations will provide globally-distributed, high vertical resolution temperature structure in the vicinity of the tropopause for characterizing the relationship between temperature and lower stratospheric water vapor densities and, in the process, constraining the transport mechanism. Since each 3 K decrease in temperature reduces the saturation vapor pressure of water by a factor of 2 at tropical tropopause temperatures, temperature accuracy of 1 K or better is needed to be consistent with knowledge of water vapor mixing ratios. As shown in figures 3.7 and 3.10, GPS derived temperatures near the tropopause should provide sub-Kelvin accuracy. Further, since the knee of the tropopause can be quite sharp, vertical resolution of ~1 km or better is needed to resolve the minimum temperature achieved by the upward migrating air mass.

Finally, assimilation of occultation profiles into general circulation models should improve the understanding of the basic behavior in the tropopause regime as well as the dynamics of troposphere-stratosphere exchange particularly using profiles retrieved in the vicinity of the ITCZ, tropopause folds and high latitude tropopause region where air is believed to return from the stratosphere to the troposphere.

5.4. Polar Regions

The importance of high altitude clouds in global change lies both in the radiative warming influence they exert [Ch. 3, Houghton *et al.*, 1990; Cess *et al.*, 1990] and the heterogeneous chemistry pathways they enable such as the role of polar stratospheric clouds in the depletion of ozone [Solomon, 1988, 1990]. While the radio occultation observations are quite insensitive to these clouds, the very accurate and high vertical

resolution profiles of density, pressure and temperature which they provide will yield quantitative information on conditions in which these clouds form and decay.

Because of the concern over reduced stratospheric ozone concentrations, much interest exists in understanding the thermodynamic structure and dynamical evolution of the winter polar vortices where ozone depletions are greatest and continue to increase. From a dynamical standpoint, the highly variable behavior of the northern hemisphere vortex is particularly interesting. The very cold lower stratosphere represents near-ideal conditions for GPS observations of density, pressure and temperature and the high density of high latitude occultation profiles from high inclination orbits provide a quantitative, high vertical resolution data set of density, pressure and temperature with which to observe and model polar vortices. Again, these measurements are insensitive to the presence of PSC's.

5.5. Boundary Layer and Surface-Atmosphere Interactions

Energy, momentum and mass are exchanged between the Earth's surface and free atmosphere above across boundary layers. The importance of properly representing the depth and thermodynamic structure of boundary layers in global forecast and climate models has been pointed out by Heckley [1985], Albrecht *et al.* [1986] and Betts and Ridgeway [1989]. Their short vertical extent makes space-borne remote sensing of boundary layers difficult. Boundary layers, in regions of relatively smooth topography can be characterized by radio occultation, particularly layers with distinctive vertical refractivity structure distinguishing them from the free atmosphere. For example, the transition at the tradewind inversion between the convective marine boundary layer and the free troposphere will produce a dramatic signature in occultation measurements as indicated in figures 2.7a and b. Diurnal variations in the structure of continental boundary layers and near-surface thermal inversions characteristic of nocturnal conditions, particularly strong in high latitude

winter climates, will also produce distinctive signatures. The sensitivity of radio propagation measurements to small scale refractivity structure in the atmosphere may therefore prove useful, in the context of boundary layers, in efforts to understand energy, momentum and constituent transfer between surface and atmosphere and their parameterization within GCM's [e.g., Betts *et al.*, 1992].

Evaporation of water provides an upward flux of latent heat which is the largest energy flux on a global scale from the surface to the atmosphere. The upward flux of water vapor has associated with it a large vertical humidity gradient near the surface. It may be possible to retrieve these gradients and therefore infer vertical water vapor fluxes in marine environments if sub-Fresnel scale vertical resolution can be achieved (Section 2.5). The sensitivity of occultation scintillation measurements for inferring structure and intensity of boundary layer turbulence and therefore the turbulent transport of energy, momentum and chemical species should be explored. This may also provide useful constraints on the exchange of energy and mass between the surface and atmosphere.

5.6. Small Scale Waves and Turbulence

Background: Small scale waves transport momentum and energy within the atmosphere. They are important contributors to middle atmosphere circulation and structure whose relative importance increases with height. For instance, gravity waves have been proposed as the cause of the extremely cold summer-time mesopause [Fritts *et al.*, 1984]. Many waves are generated in the troposphere by convective activity, wind shear and flow over topography. Radio occultation observations provide a sensitive, high vertical resolution measure of density and temperature perturbations associated with waves which has been used to characterize waves in the atmospheres of Venus, Jupiter, Titan, Uranus and Neptune [Allison, 1990; Hinson and Tyler, 1983; Hinson and Magalhaes,

1991, 1993; Hinson and Jenkins, 1995]. These features combined with global coverage suggest GPS occultations will provide a quantum step in observational capability for wave related research in Earth's atmosphere.

Resolution: The sensitivity to waves of any remote sensing technique depends on resolution. For occultation observations, waves divide into two classes, those whose vertical wavelengths are either larger or smaller than a Fresnel diameter. Observations of waves in the first class can be interpreted in terms of geometric optics whereas those in the second class produce phase and amplitude scintillations which require interpretation through diffraction theory. As per the diffraction resolution limit described in Section 2.5, GPS observations smoothed to the Fresnel scale are sensitive to waves of vertical wavelengths greater than ~2 km and horizontal wavelengths larger than of order 200 km. Sensitivity to higher horizontal wavenumbers depends on the orientation of the wavefronts relative to the GPS signal ray path. The signature of waves with horizontal wavelengths < 100 km oriented such that the wavefronts lie orthogonal to the GPS ray path will largely average out in the occultation measurements. Optimum sensitivity occurs when the wavefronts lie parallel to the GPS signal ray path such that the density perturbation along the horizontal averaging interval of the radio ray path is roughly constant and therefore adds constructively. Because of this azimuthal dependence, multi-azimuthal slices through a region will be useful.

Sensitivity: Since waves are perturbations of background atmospheric structure, the full precision of the occultation measurements can be utilized when extracting wave profiles, limited only by thermal noise and perhaps residual sub-Fresnel scale ionosphere structure (Section 3.7). When the fractional density perturbation of a small scale wave exceeds the occultation measurement noise, the presence of the waves will be apparent. Based on the fractional thermal noise contribution to refractivity error in figure 3.8, ~1-2 km vertical wavelength wave amplitudes of order 1-2% at about 60 km altitude and

~0.001% at 10 km altitude should be apparent. This high sensitivity of radio occultation to waves has been demonstrated by Hinson and Magalhaes [1991, 1993] where waves with amplitudes of order 0.1 K in the stratospheres of Uranus and Neptune are clearly discernible in Voyager radio occultation data.

Side-looking Occultations: Because of their relatively large vertical component of group velocity, short horizontal wavelength gravity waves may transfer significant momentum and energy from the lower atmosphere to the middle and upper atmosphere. Observing this class of waves has proven difficult [Fritts *et al.*, 1989] and side-looking occultations may offer a new approach. In side-looking occultations, the ray path is approximately orthogonal to the orbiting receiver velocity vector such that the motion of the ray path tangent point through the atmosphere is primarily horizontal. The pencil-beam sampling volume has a cross-track horizontal resolution of ~1.4 km defined by the Fresnel zone diameter, 1 to 2 km vertical resolution roughly defined by the 50% bending interval (Figure 2.8), and a corresponding along-path horizontal averaging interval of 200 to 300 km. Sensitivity to short horizontal wavelengths as small as 1 km or less will occur when the wavefronts are aligned parallel to the GPS ray path as discussed above. Again, multiple azimuthal cuts will be very desirable for constraining wave properties.

Scintillations: Sub-Fresnel scale wave and turbulent structure will create scintillations in the measured phase and amplitude. Diffraction maps the spatial spectrum of the wave and turbulence structure in the dimension orthogonal to the motion of the GPS signal path tangent point to a temporal spectrum of phase and amplitude variations [Hinson and Magalhaes, 1991; Yeh and Liu, 1982]. Scintillations observed during vertical occultations will therefore be due primarily to sub-Fresnel scale structure in the vertical dimension whereas those observed during side-looking occultations will be caused by horizontal structure allowing 1-D spatial spectra of small scale atmospheric structure to be derived from the observations at least when the scintillations are weak. Based upon the

Nyquist criteria, the minimum resolvable spatial wavelength is $2V_T/F_{SR}$ where V_T is the velocity of the ray path tangent point in the atmosphere and F_{SR} is the sample rate. With 50 per second sampling, and V_T equal to 3 and 7 km/sec for vertical and side-looking occultations respectively, a 1-D spatial spectrum can be retrieved down to wavelengths $\geq$ 100 meters in the vertical and $\geq$ 250 meter in the horizontal.

Sub-Fresnel scale waves may be distinguished from turbulence by their azimuthal asymmetry, again indicating the importance of multi-azimuthal scans of a region. Although similar in their effect on occultation observables, sub-Fresnel zone scale waves and turbulence differ fundamentally in that waves are resonant whereas turbulence is dissipative. A non-divergent wave transports momentum and energy through a region without permanent transfer between the wave and the region whereas turbulence is a dissipative mechanism converting bulk energy into heat and transferring momentum between waves and background circulation. The ability to remotely sense turbulence may also be of use to commercial airline traffic.

Sources and sinks: Determination of the spatial and temporal distribution of waves, their sources and sinks and the source and sink mechanisms has been limited by the distribution of the present, land-biased observational database of balloons, radars and lidars. Use of satellite observations required for a global perspective is underway and yielding promising results [Wu and Waters, 1996]. The resolution, sensitivity and global coverage of radio occultation data offers a global database with higher vertical resolution than other satellite techniques spanning an altitude range from roughly the mid-troposphere (limited by the water vapor ambiguity) into the mesosphere (limited by thermal noise and possibly ionospheric scintillations). In addition, electron densities in the ionosphere above 90 km can be derived to examine coupling between the ionosphere and neutral atmosphere below. All of these occultation attributes can be applied to study the sources and sinks of atmospheric waves. For instance, the diffraction-limited vertical resolution of 1.4 km in

the middle atmosphere is in fact adequate for resolving the dominant, vertical, gravity wave scales of 2 to >15 km observed in the middle atmosphere unlike other present satellite based observations [Muruyama *et al.*, 1992].

Since fractional thermal error, which limits occultation sensitivity scales with height approximately as exp(z/H) where H is the atmospheric scale height, whereas wave amplitude grows as exp(z/2H) by conservation of energy, the maximum sensitivity to waves, as defined by the ratio of wave amplitude to observational noise, is largest at lower altitudes. This sensitivity is important for tracing waves back to their source where they may be quite small. In fact, if a wave is discernible in occultation data at high altitudes, the lower altitude source and the wave's evolution during propagation should also be discernible from the occultation observations.

Characterization of changes in a wave field along its direction of propagation is crucial because divergence indicates sources and sinks of momentum and energy. Observed changes therefore can be used to estimate the transfer of momentum and energy and place constraints on the transfer mechanism. The occultation observations of the background atmosphere will characterize background stability and at least the large scale geostrophic wind field at mid to high latitudes which, in combination with the wide range of spatial scales discernible in the occultation data, can be used to characterize and interpret filtering of the wave field during propagation and determine how far waves of different scales are able to propagate from the lower atmosphere into the middle atmosphere.

Present and future observations: At equatorial latitudes, a number of wave modes exist including planetary waves such as Kelvin waves, Rossby-gravity waves and equatorial gravity waves whose long horizontal and short vertical dimensions are well suited for characterization by radio occultation. Mechanisms have been proposed in which these waves are coupled with and produce the Quasi-Biennial Oscillation (QBO). However, the source of these waves, particularly the Rossby-gravity waves, is open to

question. These waves have been observed with radiosondes since the 1960's but not from space due to their short vertical wavelengths. The initial GPS/Met observations have revealed the presence of waves at equatorial latitudes in the lower stratosphere. Waves with a ~3 km vertical wavelength have been identified as either a Rossby gravity or inertio-gravity wave because of their large horizontal wavelength and extent, clear asymmetry across the equator of the vertical phase structure, and a zero amplitude at the equator [Kursinski *et al.*, 1996]. These results demonstrate the resolution and sensitivity of the data as well as the importance of multiple samples of the same wave field. While individual occultation profiles provide a sensitive measure of wave structure in the vertical dimension, multiple observations sufficiently close in space and time will also constrain the wave's horizontal and/or temporal dimensions such that the wave type and phase and group velocities can be determined through its dispersion relation. The sampling density of a constellation of orbiting receivers would prove very useful in both this regard and the constraints of multi-azimuthal limb scans mentioned previously.

6. Initial Temperature Results

6.1 Introduction

In Section 3, sources of error for occultation-derived refractivity, pressure and temperature were characterized theoretically culminating in an error budget for these quantities. Here we examine initial results from GPS-MET and characterize their accuracy and resolution via comparisons with "ground truth" in the form of radiosondes and primarily the European Center for Medium-range Weather Forecasting (ECMWF) global analyses. We begin with a brief overview of the instrument and processing system developed at JPL for GPS-MET (Section 6.2). In Section 6.3 temperature results are compared against ECMWF global analyses. A sampling of climatological behavior is presented in Section 6.4. In Section 6.5, temperature data is used to make a preliminary characterization of waves in the equatorial lower stratosphere. These results also lead to some ideas concerning exchange between the stratosphere and troposphere.

6.2 Data Flow and Processing System Overview

We will discuss the occultation processing path following the data flow summarized in figure 6.1.

LEO receiver

The LEO GPS receiver is a TurboRogue receiver developed at JPL and Allen Osborne Associates modified for launch into and operation in LEO. It tracks 3 signals (CA, L1 and L2 defined in Section 2.1) from each of 8 GPS satellites simultaneously. The LEO GPS receiver continually acquires low rate data (0.1 to 0.03 samples per second) for orbit determination. During each occultation, the LEO receiver acquires 50/second phase and amplitude data while the ray path tangent point is below ~100 km altitude. The 50/second phase data is acquired from both the occulted GPS satellite and a non-occulting

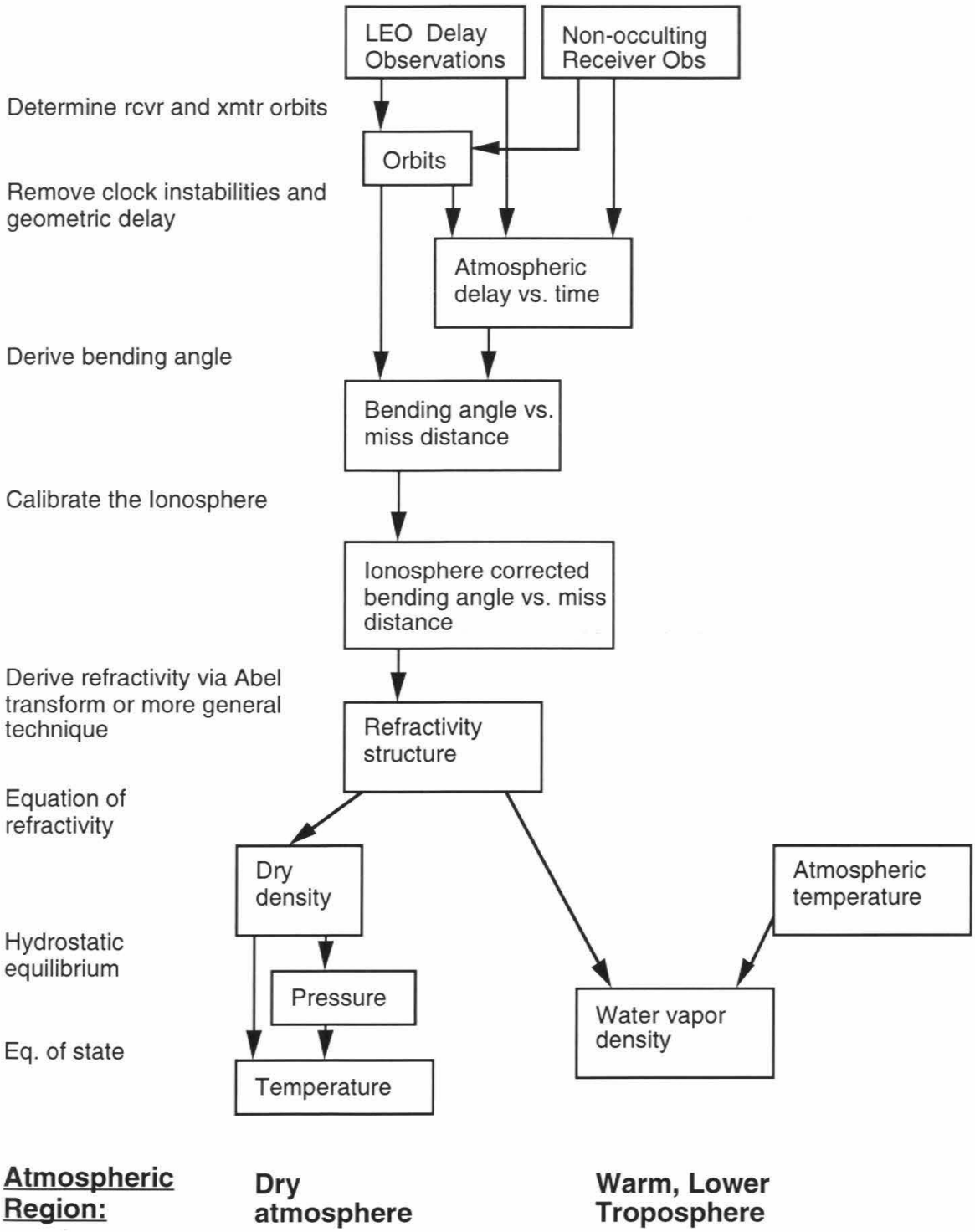

Figure 6.1 Generic procedure for deriving atmospheric profiles from occultation data.

satellite for the purpose of differencing out the unstable reference clock in the LEO receiver (Section 3.8). In the present configuration, both the occulted and non-occulted 50/sec data are downlinked to the ground where the differencing is performed. In future receivers, differencing may be done on-board to halve the downlink data rate. 10 MB of on-board storage downlinked twice a day limits the number of occultation acquired by GPS-MET to ~100 to 150 per day.

There have been four periods since the launch of GPS-MET when the GPS Anti-Spoofing (AS) encryption was turned off, three week periods in April-May, June-July and October 1995 and February 1997 during which the L2 signal is recovered to full precision and the ionosphere correction is therefore far more precise than periods when AS is on. In the examination here, we will discuss results from the April-May and June-July 1995 periods. In the next generation of receiver, beginning with the Oersted and SUNSAT launches at the end of this year, the quality of the ionosphere corrected data when AS is on will be as good if not better than GPS-MET data with AS off.

To acquire the phase, amplitude and group delay of GPS signals, the receiver continually updates a model of these variables in a process referred to as a phase-locked loop. Because of the 10 MHz pseudo range modulation, the received signal is initially downconverted and sampled at 20 MHz and these samples are then multiplied by the model signal phase and summed over 20 msec intervals. The model is continually propagated one 20 msec sample interval into the future such that when the next 20 msec sum is generated, the difference between the actual and predicted phase and group delay are used to adjust the signal model. The purpose of the phase-locked loop is to narrow band filter and reduce noise of a dynamically varying signal. The reported phase and group delay are the predicted signal plus the residual difference between the actual and predicted which provides the actual received values (plus thermal noise) when the receiver is properly in lock. Because of problems in initially acquiring a dynamical signal with a phase locked

loop, only setting occultations have been acquired thus far. A more general acquisition scheme is under development at JPL.

The 50/sec sample rate was originally chosen for the TurboRogue design because of a low rate telemetry modulation at that rate. 50 Hz is also appropriate for capturing the range of Doppler frequencies produced by atmospheric multipath in the lower troposphere (Section 2.4). The maximum altitude separation of multipath signals is ~4 km set by the height of the tradewind inversion. When the same signal arrives at the receiver via two separate paths, with tangent heights respectively at the tradewind inversion and the surface, the angle separating the two signal paths at the LEO receiver is $\sim \Delta a/D \sim \Delta z/D \sim 4/3000 =$ 1.3 mrad, which converts to a Doppler frequency separation between the two signals of ~20 Hz via eq. (3.10.1). Therefore, 50 Hz is in theory wide enough to capture multipath signals if the receiver phase locked loop can in fact track them.

Ground receiver network

A globally distributed network of high precision receivers acquires phase and group-delay data at intervals of 30 seconds up to 5 minutes for the purpose of precision positioning including the orbit of the LEO receiver. A subset of six of these receivers have been modified to acquire 1/second phase data for differencing out the GPS SA encryption on the occulted GPS signals during each occultation. The six receivers are located in Hawaii, Alaska, eastern Australia, Germany, McMurdo station in Antarctica and JPL. A table of the locations of this subset of ground receivers is uplinked to the LEO receiver which acquires only those occultations where the occulted transmitter is visible simultaneously from one of the 1/second ground receivers. Data is automatically returned to JPL once per day from each receiver in the ground network.

Precise orbit determination

Precise orbits are determined by the JPL GYPSY-OASIS system [Bertiger et al., 1994]. The geometric strength when tracking four or more GPS satellites allows instantaneous point positioning. However, because of noise, the optimum orbit solution is weighted combination of dynamic (celestial mechanics using gravity and drag models) and kinematic or point positioning information. The GYPSY-OASIS system was used to determine TOPEX orbital position to the 2-3 cm level. For occultations, where the more critical variable is orbital velocity rather than position (Section 3.6), weighting was shifted towards a dynamic solution relative to the weighting used for TOPEX. GPS-MET velocity determination accuracy has been estimated to be ~0.05 mm/sec [Bertiger and Wu, 1996] a level which contributes negligibly to overall error (Section 3.6).

Removal of clock instabilities and geometric delay

Two important calibrations are performed following orbit determination. First, using the geometry derived in the orbit determination process, the 50/sec LEO and 1/sec ground receiver data acquired during each occultation are essentially differenced to remove clock error, particularly the unstable LEO receiver clock and the SA encryption on the occulted GPS satellite signals (Section 3.8). The effects of the ionosphere on the three non-occulting differencing signal links are removed using the linear combination of two GPS signal phase delays (Section 3.7). The 1/sec ground receiver data are interpolated to the 50/sec rate of the flight receiver data before differencing. Differencing incurs a thermal noise penalty discussed in Section 3.8.

In the second step, the geometric delays associated with the orbital geometry and general relativity are removed which yields the best estimate of the phase delay due to the atmosphere.

Determination of bending angle

Determination of bending angle involves two steps. First the time derivative of the the atmospheric phase delay (= atmospheric Doppler shift) is estimated. This is accomplished presently via a quadratic fit to segments of phase data. The segment length is chosen to be the length of time over which the ray path tangent height descends the diameter of the first Fresnel zone, Z_F (Section 2.5). Thus the resolution of derived atmospheric structure is set at this step. Z_F is estimated using eq. (2.5.3) where the Fresnel scale factor, M, is estimated via eq. (2.4.1) using the received signal amplitude. The output Doppler points are oversampled by a factor of ~3 relative to Z_F to better constrain the numerical interpolation used to calculate the Abel transform.

In the second step, the bending angle and asymptotic miss distance are derived from the Doppler shift and orbital geometry using equations (2.2.9), (2.2.10) and (2.2.11) as discussed in Section 2.2. An important point discussed in Section 3.11 is the geometry is defined relative to a center of curvature which is not the center of the Earth but rather that defined by the local curvature of the geoid at the ray path tangent altitude in the plane defined by the transmitter, receiver and local orthogonal to the geoid (Section 3.11). In practice the center of curvature is determined by fitting a circle to match the curvature of the ellipsoid at the ray path tangent point.

Calibration of the ionosphere

Calibration of the ionosphere along the occulted ray path is accomplished using a linear combination of the bending angles estimated at the two GPS signal wavelengths as described in section 3.7.

Derivation of refractivity

Refractivity is derived via the abel transform of bending angle versus asymptotic miss distance (eq. 2.2.7). Presently, the abel transform is initialized at 50 km. In the future this altitude will be raised depending on the quality of the ionosphere calibration.

Since the upper integration limit of the abel transform is infinity whereas the data's finite signal to noise ratio limits the practical upper integration limit, the portion of the abel transform above the upper altitude limit is extrapolated as a simple exponential whose value at the altitude limit and the scale height are determined from the bending angle data. In the future, the extrapolation model will be generalized to represent the change of temperature with altitude and make use of mesospheric climatological constraints.

Construction of density, pressure and temperature

As discussed in Sections 2.2 and 2.3, measured refractivity profiles are converted directly to density profiles which are then integrated hydrostatically to determine pressure. Given density and pressure, temperature is obtained from the ideal gas law. An assumed profile of water vapor is used derived from a model or climatology. A minimum water vapor volume mixing ratio of $3x10^{-6}$ is assumed in the stratosphere.

Based on noise considerations, the hydrostatic integral is initialized presently at 50 km requiring a temperature estimate at 50 km as a boundary condition. This temperature is derived from a model like the NMC stratospheric analysis or climatology. A 10 K error in the 50 km temperature estimate produces a temperature error of approximately 0.1 K at the 100 mbar level (~16 km).

Water vapor retrieval

Derivation of water vapor from refractivity requires knowledge of temperature. In the present scheme, temperature is taken from either climatology or a weather model. The details of the present scheme are discussed in Section 7.

Interpolation scheme

An interpolation method is needed in many stages of the processing which must keep information as local as possible while achieving high accuracy. For this purpose a

cubic polynomial suggested by Stephen Leroy is generally used. Given a set of points and a desired location of interpolation, the nearest four points, two to either side of the desired location, are referred to as points 1, 2, 3 and 4. The desired location falls between points 2 and 3. The cubic polynomial is constrained to pass through points 2 and 3 and the slopes at points 2 and 3 are defined as the slopes between points 1 and 3 and between points 2 and 4 respectively. This form gives a fit which is continuous through the first derivative and quite local as demonstrated in the frequency and phase error response discussed in Section 3.2. Cubic splines should be avoided because they are remarkably non-local.

6.3 Comparisons with ECMWF

Here we evaluate retrieved occultation temperature profiles by comparing specific examples with radiosonde and ECMWF model profiles (figure 6.2). At high latitudes, the cold and dry conditions allow accurate temperatures to be derived almost to the surface (figure 6.2a). The retrieved profile is similar to model and nearby radiosonde profiles, with differences smaller than 1 K through most of the troposphere. In the vicinity of the tropopause and above, temperature differences are comparable to those between the radiosonde and model analysis. Agreement with the radiosonde in resolving the sharply defined tropopause and the lapse rate change below 3 km is illustrative of the sensitivity and vertical resolution of the occultation technique. At low latitudes, the ability of the technique to measure the high and cold tropopause structure characteristic of the tropics is illustrated (figure 6.2b).

To evaluate the accuracy of the occultation temperature profiles, given the rarity of close coincidences with radiosondes, we have compared them with the ECMWF global analyses (figure 6.3). The analyses used here have 30 vertical levels from the surface to

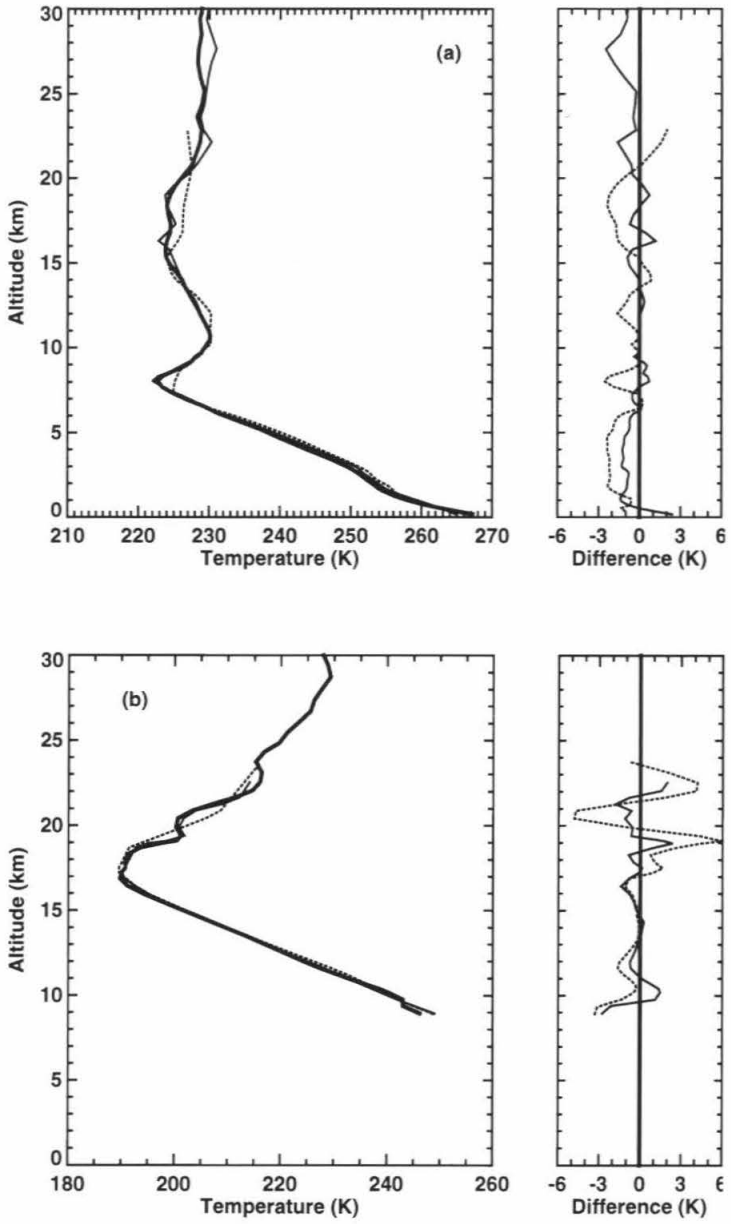

Figure 6.2 Comparisons between occultation, radiosonde, and ECMWF temperature profiles. The left panel compares the profiles and the right panel displays temperature differences (Occultation - Radiosonde/Model) as a function of altitude. (a). Occultation obtained at 01:33 UT on 5th May, 1995 over Hall Beach, Northwest Territories, Canada (69.2°N, 82.6°W). The radiosonde (00 hr UT, 68.8°N,81.3°W) is 65 km from the occultation location, and the model analysis from 00 UT is spatially interpolated to the occultation location. (b). Occultation obtained at 12:40 UT on 4th May, 1995 in the south Pacific (7.9°S, 167.5°E). The radiosonde profile (12 hr UT, 6.0°S, 170.4°E), obtained from a ship, is 350 km from the occultation location and the model analysis from 12 hr UT is spatially interpolated to the occultation location. Thick solid line is the retrieval, the thin solid line is the radiosonde and the dotted line is the ECMWF analysis.

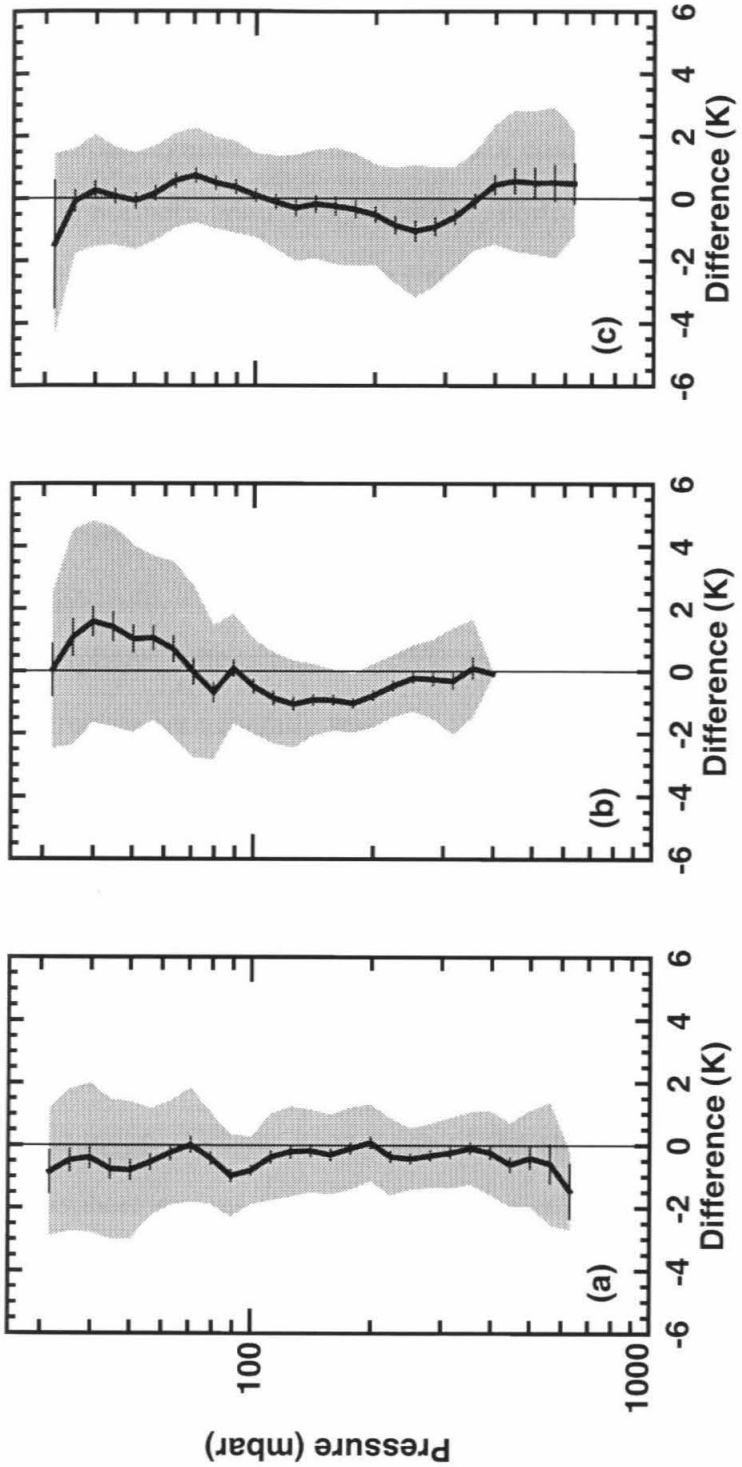

Figure 6.3 Statistical comparisons between the 6 hour ECMWF analyses and temperature profiles retrieved from radio occultations on 4th and 5th May, 1995. The panels plot mean temperature differences (retrieved-ECMWF) for (a) 34 profiles in the northern hemisphere (> 30°N), (b) 32 profiles in the tropics (30°S to 30°N), and (c) 33 profiles in the southern hemisphere (> 30°S). The vertical curve represents mean temperature differences and the horizontal error bars depict the standard error in the mean. The shaded area is defined by the mean temperature difference plus or minus the standard deviation of the temperature difference about the mean.

10 mb with the greatest vertical resolution near the surface and 100 km horizontal resolution. The ECMWF analyses have assimilated radiosonde and satellite data in a physically consistent manner such that they provide the best available interpolation globally with a resolution similar to the occultation data itself and therefore provides the best data set to compare against. For the purposes of comparison, the globe has been divided into three zones, northern hemisphere (latitudes north of 30N), tropical (latitudes between 30S and 30N) and southern hemisphere (latitudes south of 30S). The three zones were selected on the basis of the distinct difference between the height of the tropopause at low versus mid-to-high latitudes and differences between the northern and southern hemispheres in terms of surface conditions, seasonal contrast and radiosonde density and coverage. Tropospheric temperatures exceeding 250 K were excluded from the temperature comparisons to minimize issues related to the water vapor ambiguity. In the northern hemisphere, where the ECMWF analyses are the most accurate, mean temperature differences are generally less than 0.5 K and standard deviations of the differences range from 1 to 2 K. These differences include vertical structure that is not resolved by the ECMWF analysis, especially above the 100 mbar level.

Although radiosonde and TIROS Operational Vertical Sounder (TOVS) [Smith et al., 1979] data are assimilated into the ECMWF model, the analyses are less accurate in some regions of the southern hemisphere apparently due to the sparse distribution of radiosondes. Greater knowledge of the structure of the atmosphere over the southern hemisphere oceans is essential for studies of climate and the global energy and water cycles. As the occultation retrieval process has little dependence on latitude, these temperature measurements can be used to characterize atmospheric structure in the southern hemisphere in more detail. In the southern hemisphere, mean temperature differences and standard deviations increase at lower altitudes (figure 6.3c). This feature is produced by eight occultation profiles which are concentrated far from radiosonde ascents, primarily in the southeastern Pacific in the southern hemisphere storm track and close to the ice edge,

where problems in the assimilation of TOVs data are known to arise [Eyre et al., 1993] The statistics for these 8 profiles when compared with those of the remaining 25 southern hemisphere profiles, reveal a well-defined bimodal signature in model accuracy (figure 6.4). Variations in the temperature differences of the 25 profiles (figure 6.4a) are somewhat larger than, but generally comparable to, those in the northern hemisphere (figure 6.3a). In contrast, the 8 profiles (figure 6.4b) show that the model tropopause altitude is 1-2 km too low, and its temperature is ~3 K too low in the troposphere and ~2 K too high in the lower stratosphere. These differences are larger than the predicted decadal climate variations and imply that caution is appropriate when using weather model data to establish climatological behavior and study climatic changes in regions devoid of high vertical resolution observations. The temperature biases and errors in tropopause height, which must be significant given the importance of the height and topography of the tropopause to tropospheric dynamics [Hoskins et al., 1985], suggest that GPS occultation measurements will improve medium range weather forecasts in regions where weather systems move from remote oceans onto continents.

Temperature differences at tropical latitudes also display distinctive structure (figure 6.3b). On average, retrieved profiles are colder than the ECMWF analyses between 300 and 70 mbar, with a maximum difference of about 1 K near 150 mbar whereas above the 70 mbar level, they become warmer by a similar amount. Retrieved temperature gradients between 80 and 60 mbar are therefore systematically larger than model gradients. Although a little warmer than the retrievals, radiosonde data exhibit similar temperature structure in this altitude range suggesting that the model does not have sufficient resolution to represent these gradients. As equatorial waves in the lower stratosphere are not resolved by the model, they are probably responsible for the increase in standard deviation above the 100 mbar level in figure 6.3b.

The source of the temperature biases in the upper tropical troposphere (figure 6.3b) is not understood. These biases could bias estimates of convective available potential energy in the troposphere, and therefore energy transfer within the atmosphere and the severity of convective storms. They could also affect radiative emission by cirrus clouds, an important component of the greenhouse effect, and troposphere-stratosphere exchange through the thermal control of water vapor transfer. Given the preliminary nature of these results, the biases seen in figure 6.3b could be due to occultation measurement error. However, the good agreement between the retrievals and model in the northern hemisphere (figure 6.3a) argues against this. Errors in the model data are due primarily to incomplete model physics and imperfect radiosonde observations. Model physics becomes important when the model extrapolates to locations and times far from the ground truth provided by radiosondes. However, radiosonde temperatures are themselves imperfect and require corrections for absorption of solar and IR radiation, thermal emission, and conduction and convection of heat [Luers and Eskridge, 1995]. Inadequate calibration could contribute to the temperature biases seen in figure 6.3b, to the extent that the model is constrained by radiosondes in this region.

Figures 6.2, 6.3 and 6.4 indicate that future measurements, if available in near real-time, could play a significant role in numerical weather prediction (NWP). The density of 500 globally-distributed measurements per day provided by a single orbiting GPS receiver would exceed that of the radiosonde network by a factor of two in the southern hemisphere, making a significant contribution to the global observing system. A constellation of orbiting receivers could make a major contribution to fulfilling the stated temperature observation requirements for global NWP. The results presented here have demonstrated desirable properties for use in NWP namely generally good agreement with a high-quality NWP analysis, plus the ability to identify a minority of cases where there is room for significant improvement in the analysis.

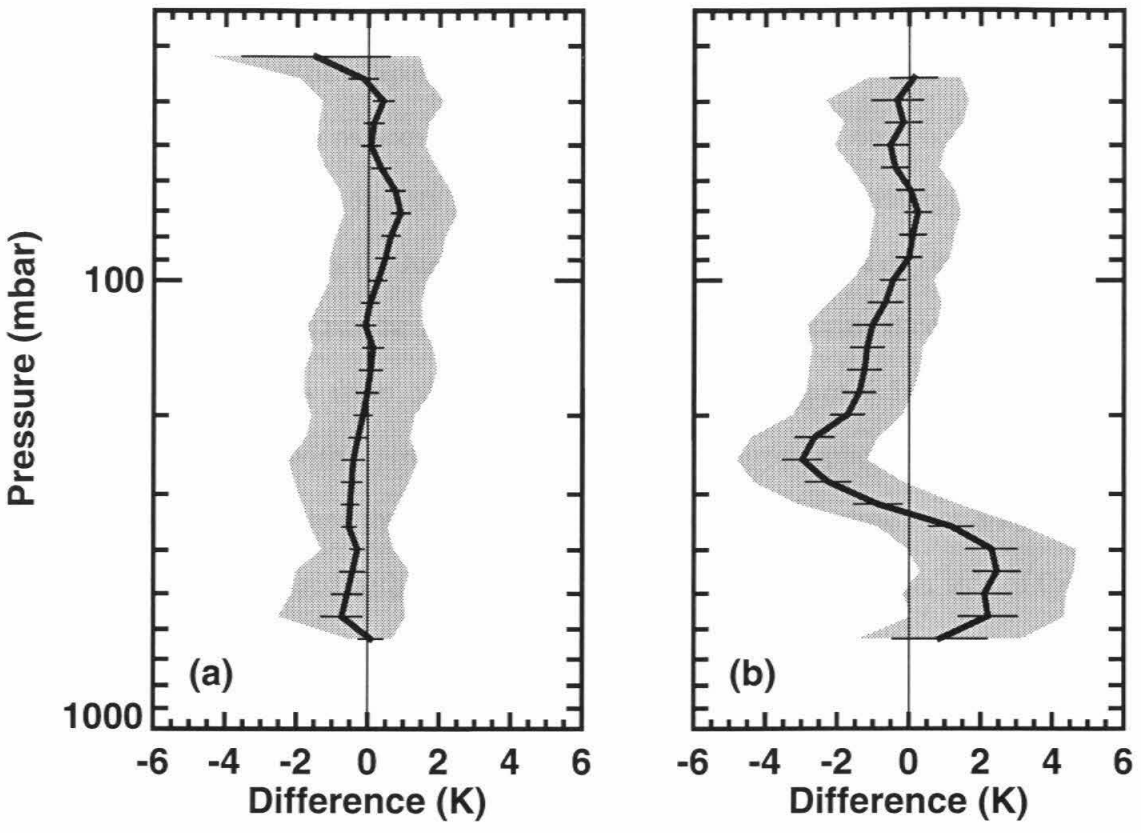

Figure 6.4 Statistical comparison of the 33 southern hemisphere profiles from April-May 1995 period divided into two groups: (a) 25 profiles which differ little relative to the model and (b) 8 profiles with the largest deviations relative to the model.

6.4 A Sampling of the Climatological Range of Temperature Structure

We will now discuss several example profiles chosen to demonstrate the range of climatological behavior of temperature in the lowest 30 km of Earth's atmosphere. Figure 6.5 shows eight occultations acquired by GPS-MET over an hour on June 22, 1995. The solid lines are the occultation data and the dotted lines are profiles interpolated from the NMC stratospheric analysis for June 22. Since June 22 is the longest day of the year in the northern hemisphere, the two hemispheres exhibit significant seasonal contrast. The first four panels are profiles from 50N to 65N. Panel a of figure 6.5 shows a rounded stratospheric structure with a possible hint of waves. The troposphere in panel a of figure 6.5 is somewhat colder than those in panels b-d with a low tropopause near 8 km and 235 K which may be associated with cyclonic storm activity or air on the polar side of the polar jet stream. The profiles of figure 6.5 b, c and d are similar to one another with a more distinct transition between the troposphere and stratosphere than figure 6.5a. The troposphere is evident by the large negative temperature gradient capped off by a tropopause between 11 and 12 km altitude roughly between 210 and 220 K. Structure near the tropopause has a distinct vertical notch-like signature with a relatively large positive vertical temperature gradient on the topside of the minimum temperature extending over a 1 to 2 km interval. The sharp variations in tropopause structure demonstrate the vertical resolution of the occultation technique. Deriving such structure from passive nadir sounders is difficult if not impossible because of coarse vertical resolution, combined with a local temperature minimum and somewhat isothermal structure of the tropopause region. The notch structure may result from vertical motion in the upper troposphere where rising air overshoots the height of radiative equilibrium creating a tropopause which is higher and colder than the radiative equilibrium structure. The high vertical gradient over the 1 to 2 km interval on the topside of the tropopause reflects the interval over which the atmosphere returns to radiative equilibrium which largely controls the thermal structure in the

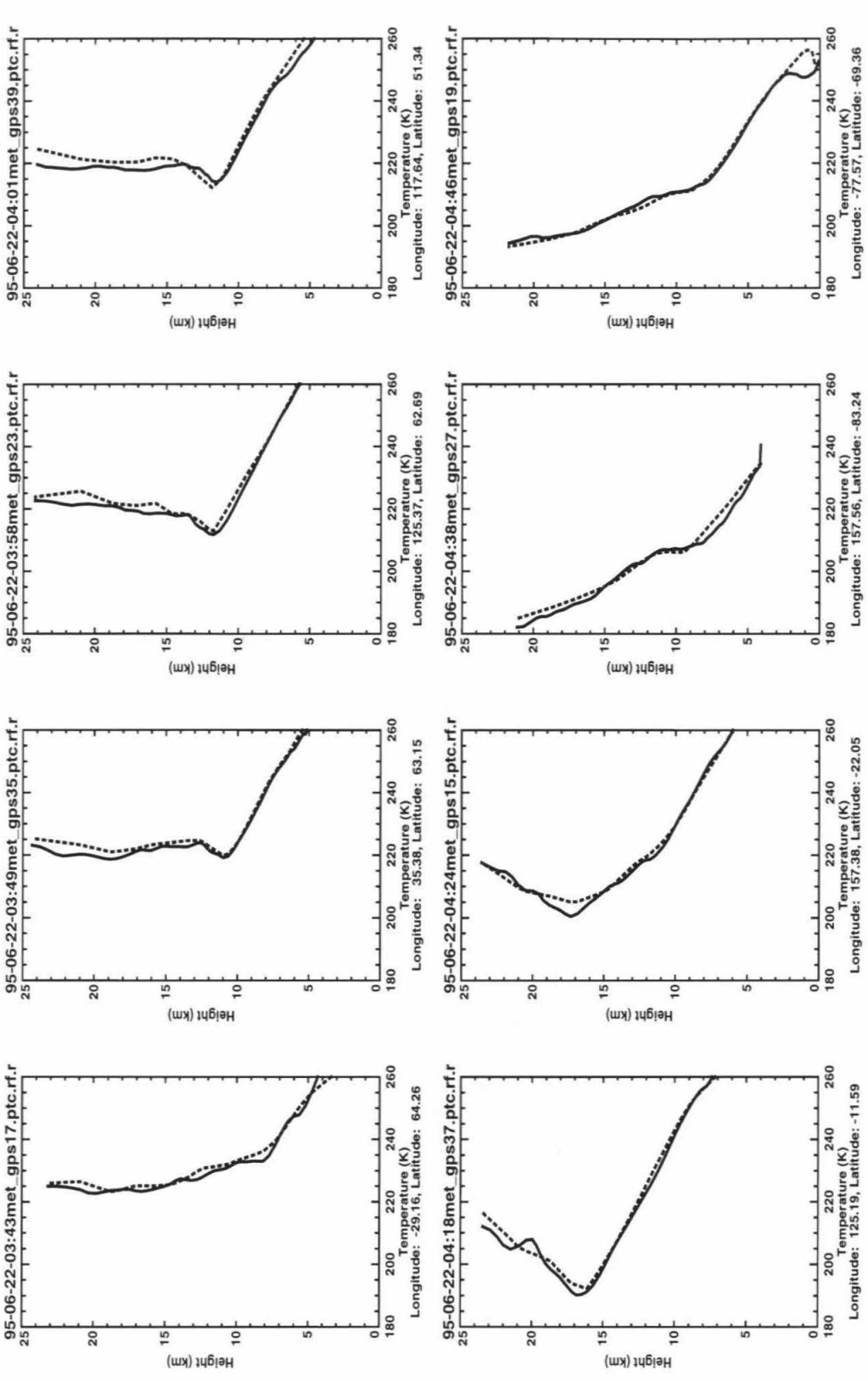

Figure 6.5 8 profiles from June 22, 1995 demonstrating the range of temperature structure observed by GPS-MET. Solid lines are occultation profiles. Dashed lines are interpolated from ECMWF analyses

stratosphere. In the northern (summertime) hemisphere, stratospheric temperatures generally increase with altitude toward the relative maximum temperature at the stratopause.

Figure 6.5e (12.8°S) shows a temperature profile representative of tropical structure with a high and cold tropical tropopause of about 190 K near 17 km. The vertical influence of convection in the tropics causes the troposphere to extend much higher at tropical latitudes. Stratospheric temperatures increase more rapidly with height than at higher latitudes because of the high altitude and lower temperatures of the tropopause. Another feature of tropical profiles is the waviness in the lower stratosphere which we will discuss in Section 6.5. Figure 6.5f shows a double tropopause characteristic of the 20 to 30° latitude band. The lower altitude tropopause defines the top of the troposphere near 11 km and 220 K. The minimum temperature near 17 km is related to the tropical tropopause at lower latitudes. The region between the two tropopauses has become known as the middle world between the troposphere and the rest of the stratosphere where material apparently flows meridionally between the upper tropical troposphere and the middle world.

Figures 6.5 panels g and h demonstrate high southern latitude winter conditions. Note the contrast between these two profiles and those in figure 6.5 panels a-d. Tropospheric temperatures are much colder in the winter hemisphere such that at 6 km altitude about a 40 K difference exists between figure 6.5 panels b-d (260 K) and g and h (220 K). The tropopause is near 8 km with temperatures near 210 K. There is a transition region again on the topside of the tropopause about 4 km in extent across which there is a small negative temperature gradient between the colder tropospheric temperatures and the radiatively controlled stratospheric temperatures above. Stratospheric temperatures are noticeably colder than their summertime counterparts decreasing to minimum values near 25 km altitude reaching minimum temperature less than 185 K in figure 6.5g. Note also how the agreement between the NMC model and the occultation profiles is much better in

the northern than in the southern hemisphere reflecting reduced information used by the analysis in the southern hemisphere.

The notch may be an indicator of vertical motion in the upper troposphere and the lack of the notch structure at the southern hemisphere tropopause may indicate less vertical motion in the upper troposphere in the southern hemisphere winter than the northern hemisphere summer. However, because little wave activity at 1 to 3 km vertical scales is observed in the high latitude lower stratosphere profiles whereas most low latitude lower stratosphere profiles do exhibit wave activity at these scales makes this conclusion questionable particularly given that vertical convective motion is believed responsible for the low latitude wave activity. Additional comments on the implications of the notch are made at the conclusion of Section 6.5.

6.5 Tropical Waves Near the Tropopause

In this section we discuss some initial occultation observations of small vertical scale waves in the equatorial lower stratosphere and reflect on some of their possible implications. The tropopause regime is well suited to probing by GPS occultations because of the resolution, accuracy and the very dry conditions which eliminate the wet-dry ambiguity. A comparison of low latitude ($< 25^{o}$) and higher latitude profiles reveals that equatorial temperature profiles in the lower stratosphere exhibit far more wavy perturbations (Figure 6.6). Several planetary scale wave types are believed to exist in the equatorial regime whose dimensions are well matched to the occultation resolution. We will explore the ability of these modes to explain the observed wavy behavior. The occultation data represent the first observations of these waves from space. The general inability to connect the phases of the waves from profile to profile indicates the waves are much smaller than the dominant observed planetary-scale waves in the equatorial lower

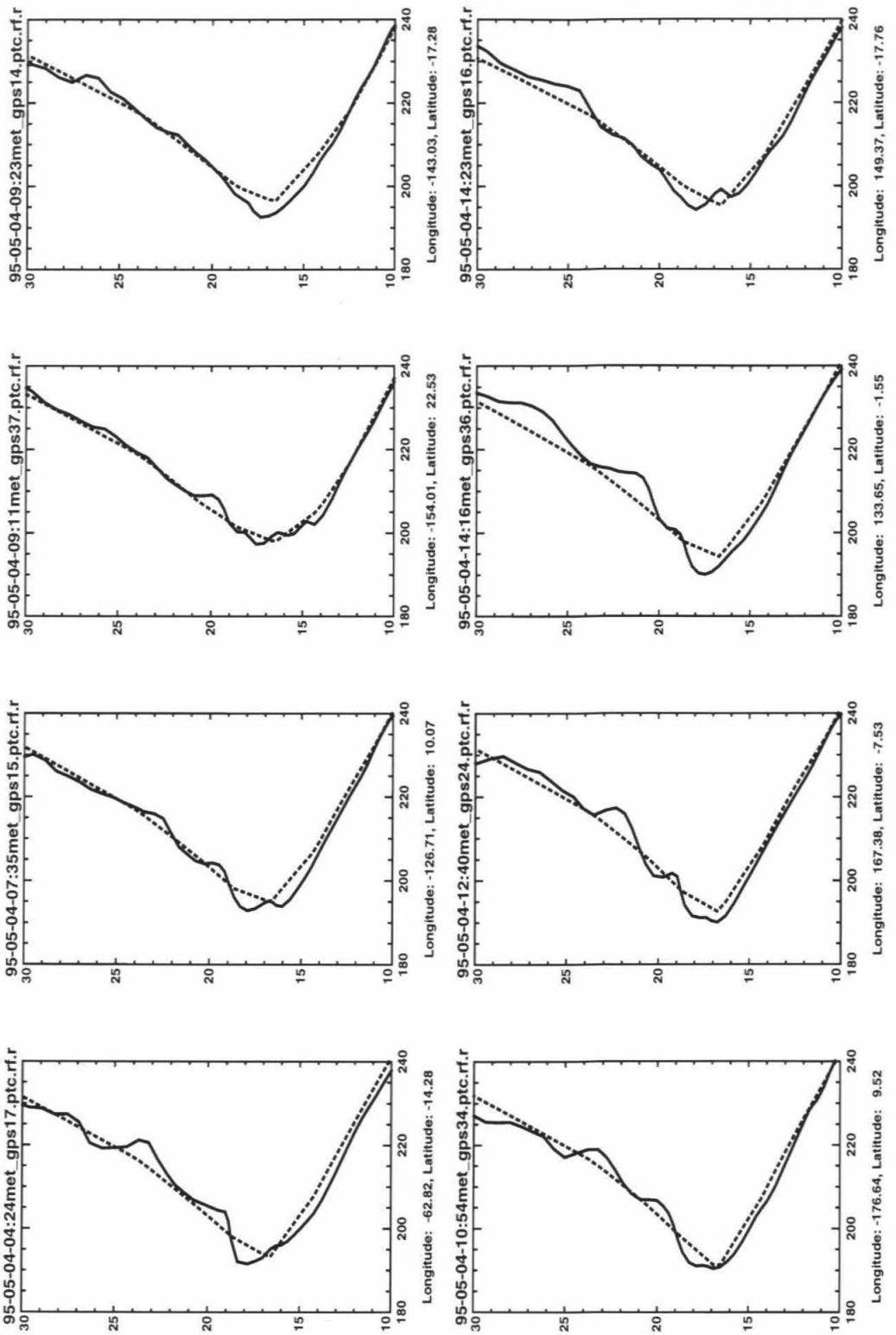

Figure 6.6 Examples of low latitude temperature profiles. Solid lines of occultation profiles. Dashed lines are interpolated from daily NMC stratosphere analyses

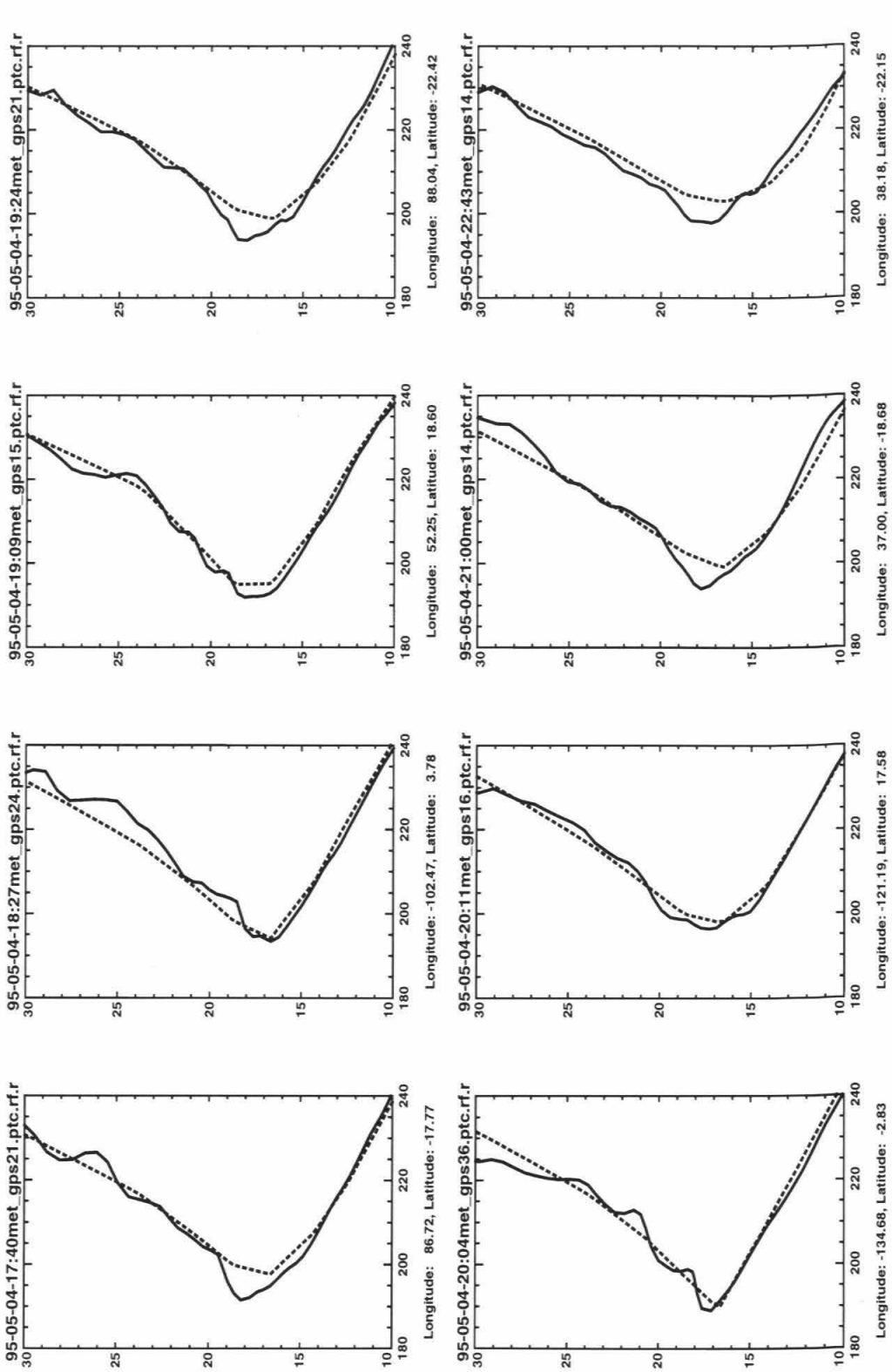

Figure 6.6 (cont'd.) Examples of low latitude temperature profiles. Solid lines of occultation profiles. Dashed lines are interpolated from daily NMC stratosphere analyses

stratosphere stated by AHL to be wavenumber 1 to 4. One pair of profiles separated by 12 hours and more than 1000 km does demonstrate coherence over ~1000 km scales. A fairly close coincidence between an occultation and radiosonde profile reveal close agreement and rapid appearance and disappearance of the wave providing further indication of wavelengths of order 1000 km or less and may hint at nonlinear behavior near the wave packet edges. The presence of equatorial waves near the tropopause suggests two roles which they may play in exchange between the troposphere and stratosphere. First, they modulate the temperatures and therefore water vapor saturation mixing ratios of rising air. Second, the waves may affect vertical advection across the tropopause. Still another point is the altitude range over which these waves exist coincides approximately with that of the Quasi-Biennial-Oscillation (QBO) suggesting these waves may play a role in driving the QBO. A more in-depth examination of the occultation data and the increase in temporal and spatial sampling density of future occultation observations will produce a better climatology of equatorial wave types and characteristics and better answers to some of the issues raised here.

Overview of equatorial wave types

At equatorial latitudes, Kelvin waves, Rossby-gravity waves (RGW), inertio-gravity waves (IGW) and Rossby waves form a set of low order planetary waves whose long horizontal and short vertical dimensions are well suited for characterization by radio occultation. The four wave types are characterized by dispersion relations which couple each wave's zonal and vertical wavelengths and period (Appendix A.5). Some key features of the waves are summarized in Table 6.1. The source of these waves is believed to be tropospheric convective activity. Uncertainty exists as to how waves of low horizontal wavenumber can be excited. Mechanisms have been proposed in which these waves are coupled with and produce the QBO. Equatorial waves were first observed in radiosonde data in the 1960's. However, radiosonde sampling is limited by extremely

inhomogeneous spatial distribution with large regions of the tropics left routinely unsampled. The occultation observations represent the first satellite data set with sufficient vertical resolution to observe these equatorial waves globally and routinely. Waves in the lower stratosphere with short vertical wavelengths appear in most occultation profiles within 20° of the equator. The preliminary results demonstrate the resolution and sensitivity of the occultation data as well as the need for multiple samples of the same wave field.

Table 6.1

Summary of Important Equatorial Wave Characteristics

Wave	Equatorial Symmetry	Meridional Flow	Frequency	Speed	c_x	c_{gx}
Kelvin	symmetric	No	low	slow	eastward	eastward
RGW	antisymmetric	Yes	low-med	med	either	eastward
IGW	either	Yes	high	fast	either	same as c_x
Rossby	either	Yes	low	slow	westward	west ($\|\tilde{\omega}\| < 0.5$) east ($\|\tilde{\omega}\| > 0.5$)

Note: dispersion relations and other equations are given in Appendix A.5.

Individual occultation profiles

Identifying an equatorial planetary wave as one of the four requires that two out of the wave's three scales, zonal wavenumber (k), vertical wavenumber (m) and frequency (ω), be determined. Equatorially trapped waves are simpler than the general form of waves because their horizontal direction of propagation is limited to the zonal direction. Since an

individual vertical occultation profile provides only m, resolving k or ω requires two or more profiles in the same vicinity within the same wave cycle. The three sampling requirements to acquire a snapshot of a wave and identify it are:

1. Several samples are needed on both sides of the equator to determine the wave's meridional structure.

2. Longitudinal separation between samples must be a significant fraction of the horizontal wavelength to determine k.

3. The time span of the samples should be less than a wave period to eliminate cycle ambiguities.

The zonal wind which is available from meteorological stratospheric analyses is required to determine the wave's phase speed in the medium (although one must wonder how analyses are able to properly interpret the undersampled equatorial radiosonde network data in the presence of the waves discussed here). A more general characterization of a wave packet can be achieved using profiles acquired over a wider range of space and time such that the phase and group velocities and amplitude evolution of the wave can be determined.

Occultation sampling

During each overpass, a LEO GPS receiver will acquire occultations within roughly $\pm$2000 km of the suborbital track of the receiver assuming an azimuthal cutoff of $\pm 45^{o}$ with respect to the LEO receiver's velocity vector. This cutoff allows capture of the occultations where the ray path motion is still primarily vertical (section 3.11). Since Earth rotates by 25^{o} during the ~100 minute orbital period of an LEO, the tracks of consecutive orbits are separated longitudinally by ~2800 km at the equator such that consecutive swaths of occultations will overlap. According to AHL, the dominant modes of the Kelvin and RGW are wavenumber 1-2 and 4 respectively with periods of 15 and 4-5 days respectively.

Sampling over a half cycle of wavenumber 4 (5000 km) requires three consecutive orbits and 200 minutes, a small fraction of the wave's period. Analogous sampling of the Kelvin wave would require 2 to 4 times as much time which is still a small fraction of its 15 day period. Occultations from a single LEO receiver will provide sampling density adequate to characterize the dominant Kelvin and RGW modes. The same conclusion may be true for equatorial Rossby waves which have longer periods than Kelvin waves at comparable vertical wavelengths.

Achieving sampling adequate to resolve the smaller horizontal and temporal scales of IGWs presents a significantly greater challenge. IGWs with ~3 km vertical wavelength and 3 to 6 hour periods, which have been proposed to explain radiosonde observations (Hamilton and Vincent, 1995), have zonal wavelengths of only 150 to 300 km. In order for an occultation to recover the density perturbation of such high zonal wavenumber waves accurately, the phase of the perturbation must change little (< cycle/4) over the ~300 km occultation averaging interval. Therefore, the azimuth of the occultation ray path must be less than $\sim\lambda_x/(4*300)$ radians relative to north where the units of λ_x are km. For λ_x = 150 km, the resulting maximum azimuth is ~8°, which is much tighter than the $\pm$45° azimuthal cutoff mentioned previously which will significantly reduce the number of available occultations. For azimuths somewhat larger than 8° the wave may be apparent but its amplitude will be attenuated. High wavenumber waves will probably have to be identified from the set of occultations from a single orbital track because of the many wavelengths separating between consecutive orbital tracks. Multiple receivers in the same orbital plane will improve the sampling density problem. The suborbital tracks of 3 coplanar LEO receivers would be separated by 33 minutes in time and 900 km in longitude which is still large relative to the zonal wavelengths of 3 to 6 hour IGWs.

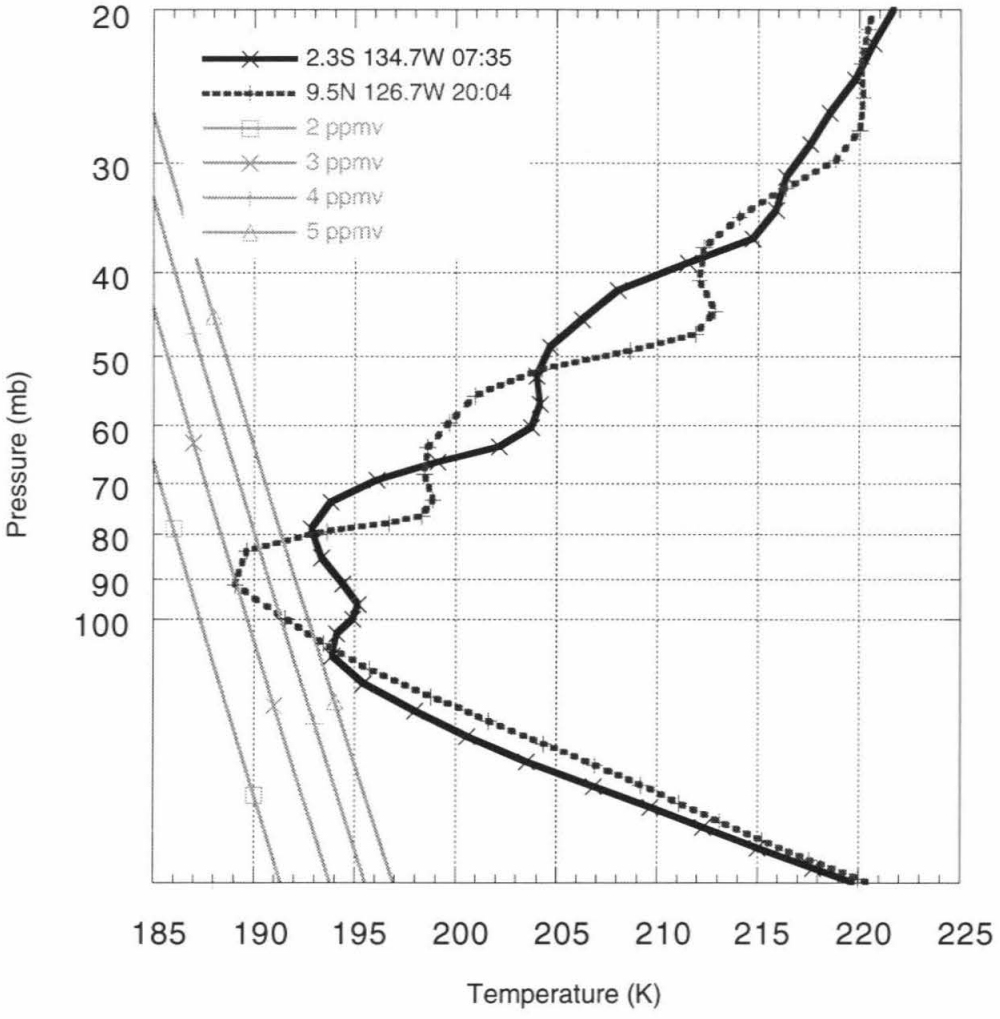

Figure 6.7 Comparison of two low latitude occultation temperature profiles 12.5 hours and 1500 km apart exhibiting apparently the same atmospheric wave approximately 180° out of phase. Also shown are 4 saturation vapor mixing ratios over ice versus pressure to demonstrate how the waves may freeze dry air near the tropopause (ppmv = parts per million by volume).

Preliminary attempts to interpret wave profiles

We will now focus on two sets of profiles from the May 4-5, 1995, period. The first set includes two profiles which show clearly anticorrelated wave perturbations plus a third profile which appears to see the same wave and places much tighter constraints on the wave type than the first two profiles alone. The second set consists of an occultation and a 300 km distant radiosonde observation which contain very similar wave structure. Interpretation of the waves is based primarily on the phase relation between wave profiles although amplitude is considered as well. This examination demonstrates the complexity of the observed wave structure, uncertainties in interpretation and the need for additional sampling in the future.

We have eliminated Rossby waves (as opposed to RGW) from contention because, at the observed short vertical wavelengths, Rossby wave periods are 20 days or more and therefore too slow to be consistent with the rapid appearance and disappearance established by the radiosonde constraint.

Two anticorrelated occultation profiles

The strongest argument for planetary scale equatorial waves in the occultation observations examined thus far from May 4 and 5, 1995, comes from the two temperature profiles at 05/04-07:35UT (9.5N, 126.7W) and 05/04-20:04UT (2.3S, 134.7W) (figure 6.7). The ray paths for each occultation are oriented nearly north-south. Comparison of the two profiles reveals large perturbations approximately half a cycle out of phase with respect to one another. The vertical wavelength of the perturbation is roughly 3.2 km and its amplitude is 2 - 3 K. The two profiles are separated by about 8^{o} in longitude, 11.8^{o} in latitude and 12.5 hours in time; the perturbations are apparently associated with a large wave. To further constrain the possible wave types, a third profile at 11.1S, 116.7W taken at 05/05-06:55 is used which exhibits a weaker perturbation than the other two profiles

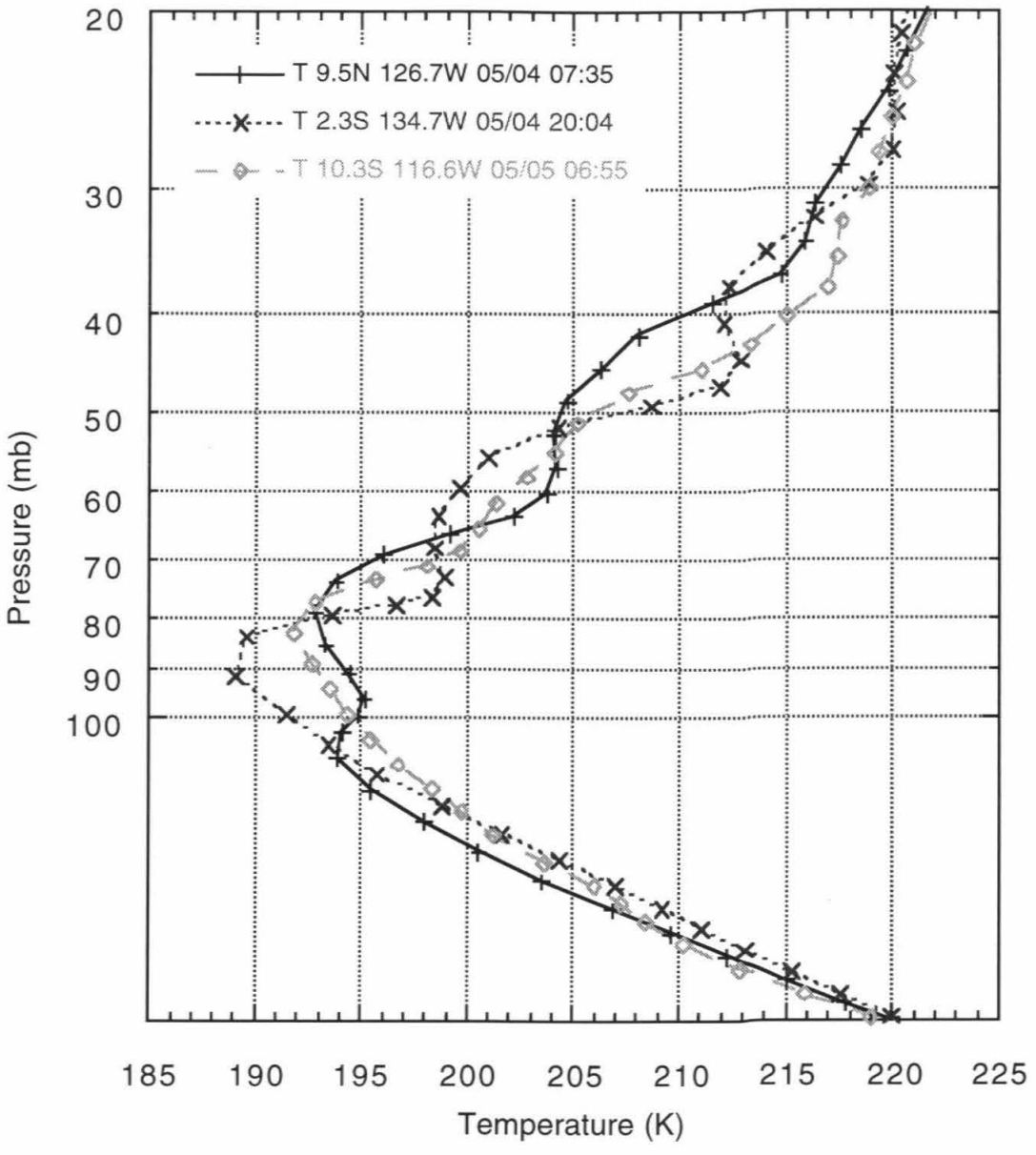

Figure 6.8 Comparison of the occultation profile taken on May 5, 1995 at 06:55 with the two low latitude profiles in figure 6.7. The three profiles span 23.5 hours and 2000 km apparently sampling the same atmospheric wave.

(figure 6.8). Under the assumption that vertical group velocity is directed upwards and vertical phase velocity is downward (consistent with a tropospheric convection source), the phase of the 20:04 profile is roughly 0.2 cycles ahead of that of the 06:55 profile which is roughly 0.2 cycles ahead of the 07:35 profile.

We will now attempt to determine the wave type by assuming that the three wave profiles are sampling a common wave which is either a Kelvin wave, RGW or IGW. The procedure is to define a set of waves consistent with the phase of the first two profiles (07:35 and 20:04) and then compare the observed phase of the third profile (06:55) with that predicted by the set of waves until matches are found. A match is defined as the phase of the 06:55 profile leading that of the 20:04 profile by approximately 0.7 to 0.9 cycles. The NMC stratospheric analysis estimate of the zonal wind velocity of -1 to -2.5 m/sec in this region for May 4 and 5 is small compared to the wave's phase velocity and will only slightly affect phase propagation. The results indicate the zonal wavelength is ~2000 km or less.

Kelvin wave: Kelvin waves are the simplest equatorial wave type. The anti-correlated phase across the equator immediately eliminates from possibility Kelvin waves of wavenumber 1 or 2 which have been stated to be one of the two dominant equatorial wave modes (AHL, p. 211). However, shorter wavelength Kelvin waves are possible solutions. To be a Kelvin wave, the change in the kx-ωt phase between the two profiles must be one-half cycle under the constraints that $k > 0$ (AHL convention), $\omega = -Nk/m$ (Kelvin dispersion relation [Appendix A.5] where N is the buoyancy frequency), phase propagation is eastward ($\omega/k > 0$) and $m < 0$. The observational constraint is

$$k\Delta x - \omega\Delta t \simeq (j + 0.5 - 0.1)\, 2\pi = (2j+1-0.2)\, \pi$$

where j is some integer and $\Delta x = -900$ km and $\Delta t = 45000$ seconds. The factor of -0.1 accounts for the fact that the phase of the wave in the 07:35 profile is ~0.4 rather than 0.5

cycles behind that of the 20:04 profile. Applying the Kelvin wave dispersion relation yields

$$k\,(\Delta x + N\Delta t/m) = (2j+1-0.2)\,\pi$$

such that

$$k = \frac{(2j + 1 - 0.2)\,\pi}{\Delta x + N\Delta t/m} \qquad (6.5.1)$$

Since Δx and m are negative, the denominator is negative, and since k must be positive, j is negative. The lowest order solution (j = -1) yields a zonal wavelength of 2650 km (~24° of longitude), a period of 2 days and an eastward phase (and group) velocity of 15.6 m/sec. However, the phase of the j=-1 Kelvin wave in the 05/05-06:55 profile should be 0.4 cycles ahead of that in the 05/04-20:04 (2.3S, 134.7W) profile, and 0.1 cycle behind the 05/04-07:35 (9.5N, 126.7W) profile, whereas in the observations, the 06:55 profile phase is 0.2 cycles *behind* that of the 20:04 profile. Further, assuming the wave packet extends at least a few zonal wavelengths in longitude (as required for the dispersion relation to hold), this wave should be visible across at least half of the Pacific basin which it is not. Therefore, the observed wave is not a j=-1 Kelvin wave. Table 6.2 shows shorter wavelength Kelvin solutions (up to six wavelengths between the 07:35 and 20:04 profiles) indicating that the only Kelvin wave solution which is approximately consistent with the phase constraints of these three profiles, has a zonal wavelength of 280 km, much smaller than the dominant Kelvin mode of wavenumber 1 - 2 quoted by AHL.

RGW: RGWs are asymmetric about the equator. The phase, kx-ωt, at the locations of the two anticorrelated profiles must therefore be roughly the same (to an integer cycle ambiguity) to account for the anticorrelated structure, a relation which can be expressed as

Table 6.2

Predicted phase relationships for Kelvin waves

j	k	wavelength (km)	period (days)	x-phase (cycles)	t-phase (cycles)	06:55 - 20:04 net phase (cycles)
-1	2.37E-06	2645.78	1.97	0.76	0.23	0.53
-2	6.33E-06	992.17	0.74	2.01	0.61	0.40
-3	1.03E-05	610.57	0.45	3.27	1.00	0.28
-4	1.42E-05	440.96	0.33	4.53	1.38	0.15
-5	1.82E-05	345.10	0.26	5.79	1.76	0.03
-6	2.22E-05	283.48	0.21	7.05	2.14	0.90

Note: phase here is the phase of the 06:55 profile relative to that of the 20:04 profile. Net phase equals the spatial phase (x-phase) minus the temporal phase (t-phase) modulo 1 because phase equals ($kx-\omega t$)

$$k\Delta x - \omega\Delta t \simeq 2\pi (j - 0.1) \qquad (6.5.2)$$

As for the Kelvin wave, the factor of -0.1 accounts for the phase of the wave in the 07:35 profile being ~0.4 rather than 0.5 cycles behind that of the 20:04 profile. This observational constraint can be combined with the RGW dispersion relation,

$$k = -\frac{m|\omega|}{N} - \frac{\beta}{\omega}$$

where β is $2\Omega \cos\phi / R_{Earth}$ and Ω is the angular rotation rate of Earth, ϕ is the latitude, R_{Earth} is Earth's radius. Ignoring the zonal wind, which is apparently small, yields a quadratic expression for ω. For RGWs of westward phase propagation, $\omega < 0$ and $m > 0$ and

$$k = \frac{m\,\omega}{N} - \frac{\beta}{\omega} = \frac{\omega\,\Delta t + 2\pi(j - 0.1)}{\Delta x}$$

such that

$$\omega^2\left(\frac{\Delta t}{\Delta x} - \frac{m}{N}\right) + \omega\,\frac{2\pi(j - 0.1)}{\Delta x} + \beta = 0 \qquad (6.5.3a)$$

For eastward propagation, $\omega > 0$ and $m < 0$ and the quadratic expression for ω is

$$\omega^2\left(\frac{\Delta t}{\Delta x} + \frac{m}{N}\right) + \omega\,\frac{2\pi(j - 0.1)}{\Delta x} + \beta = 0 \qquad (6.5.3b)$$

Table 6.3 shows the results for ω and k for different values of j. We consider first the lowest possible wavenumber solution, j=0. Although the observed wave has a shorter vertical wavelength, the anticorrelated vertical wave structure of the 07:35 and 20:04 profiles combined with their 8 degree longitudinal and 12.5 hour separation are consistent with a westward propagating ~wavenumber 4 RGW (j=0 solution in Table 6.3), one of the two dominant equatorial waves quoted by AHL. However, as with the Kelvin wave solutions, the lowest wavenumber solution is inconsistent with the observed phase of the 06:55 wave profile. This conclusion is further borne out by the fact that, at such a large wavelength, even a very short wave packet a few wavelengths in length would be visible around most of the equatorial Earth which they are not. The j = -1 and -2 eastward solutions satisfy the observational phase constraints with wavelengths of 1650 km and 800 km respectively. Note these are *eastward* propagating whereas the dominant RGW wavenumber 4 mode quoted by AHL travels *westward*. The westward j=-2 solution has approximately the correct phase relation for the three profiles, but it is an extremely slow wave which turns out to be inconsistent with the near coincident radiosonde observations discussed shortly.

Table 6.3

Predicted phase relationships for Rossby-Gravity waves

j	ω	period (days)	k	wavelength (km)	c_x (m/s)	x-phase (cycles)	t-phase (cycles)	net phase (cycles)
0	-1.13E-05	-6.41	1.28E-06	4902.55	-8.85	0.41	-0.07	0.48
-1	-2.81E-06	-25.85	7.93E-06	792.77	-0.35	2.52	-0.02	0.54
-2	-1.52E-06	-47.94	1.49E-05	420.68	-0.10	4.75	-0.01	0.76
-1	7.06E-05	1.03	4.21E-06	1492	16.76	1.34	0.44	0.90
-2	1.31E-04	0.56	8.24E-06	763	15.89	2.62	0.81	0.80

Note: phase here is the phase of the 06:55 profile relative to that of the 20:04 profile.

Net phase equals x-phase - t-phase + zero crossing phase modulo 1

c_x is phase velocity, ω/k

IGW: IGW solutions are symmetric (asymmetric) about the equator for odd (even) n (Appendix A.5). The number of sign changes or zero crossings in a meridional cross section of an IGW equals n+1 and the approximate latitudes at which the zero crossings occur are given in Table 6.4. In the lowest symmetric mode (n=1), sign changes occur near $\pm 7^{o}$ latitude such that profiles at 2.3°S and 9.5°N latitude will be a half cycle out of phase. For the lowest asymmetric mode (n=2), the sign changes at the equator and near $\pm 11^{o}$ latitude again causing profiles at 2.3°S and 9.5°N to be a half cycle out of phase. Therefore, both the n=1 and n=2 modes can be consistent with the 07:35 and 20:02 profile phases with relatively large zonal wavelengths. Additional higher n IGW solutions exist indicating the need for denser sampling for ambiguity resolution. The ~300 km along track averaging of the occultations in the meridional direction limits n to be ~6 or less.

Solutions consistent with the observational constraints (eq. 6.5.2) and the IGW dispersion relation (Appendix A.5) were found numerically. All the westward IGW solutions which satisfy the $\phi_{06:55} - \phi_{20:04} \sim 0.8$ cycles requirement have small zonal wavelengths less than 200 km and periods of 2 to 3 hours (Table 6.5). The best agreement for an eastward IGW mode is n = 2 with zonal wavelength of 1600 km and period of ~1 day. An eastward IGW with n = 4 (and possibly 6) is roughly consistent with the observed phase requirement as well. An n = 5 solution is unlikely because a zero crossing occurs very close to the latitude of the 06:55 profile such that this profile would have ~zero amplitude. Before moving on, we also note an unusual IGW solution of wavenumber 4, n=6, with a phase velocity of -110 m/s and a very slow group velocity of -1.6 m/s. Such a wave would be virtually a standing wave. It is not clear such a wave could ever be excited because the time required for wave energy to propagate and fill out a wave packet a few zonal wavelengths in length (as required for the dispersion relation to hold) would take of order $3 \times 10^7 / 1.6$ seconds or ~200 days!

Table 6.4

Approximate Latitudes of Meridional Zero Crossings for IGWs

n	Latitude (degrees)				07:35-20:04 shift	20:04 - 06:55 shift
1	-	7			π	π
2	0	11			π	0
3	-	4	14		π	π
4	0	7.5	16.7		0	π
5	-	3.3	10.2	19	π	0
6	0	6	12.5	21.2	0	π

Table 6.5

Predicted phase relationships for Inertio-Gravity Waves

Westward phase propagation

n	$\phi_{20:04}$-$\phi_{07:35}$ (cycles)	wavelength (km)	period (days)	c_x (m/sec)	c_{gx} (m/sec)	x-phase (cycles)	t-phase (cycles)	zero crossing phase (cycles)	$\phi_{06:55}$-$\phi_{20:04}$ (cycles)
1	-0.1	1382	-0.96	-16.66	-14.48	1.45	-0.47	0.50	0.42
1	-1.1	171	-0.13	-15.58	-15.55	11.68	-3.56	0.50	0.74
2	-0.1	1185	-0.80	-17.14	-14.07	1.69	-0.57	0.00	0.25
2	-1.1	170	-0.13	-15.62	-15.51	11.75	-3.59	0.00	0.34
3	-0.1	1053	-0.70	-17.41	-13.86	1.90	-0.65	0.50	0.04
3	-1.1	169	-0.13	-15.59	-15.54	11.82	-3.60	0.50	0.92
4	-0.6	299	-0.22	-15.77	-15.36	6.69	-2.06	0.50	0.25
4	-1.6	117	-0.09	-15.60	-15.53	17.08	-5.21	0.50	0.79
5	-0.1	898	-0.59	-17.77	-13.59	2.22	-0.77	0.00	1.00
5	-1.1	168	-0.12	-15.64	-15.49	11.92	-3.65	0.00	0.57
6	0.4	10078	-1.07	-109.46	-1.63	0.20	-0.42	0.50	0.12
6	-0.6	292	-0.21	-15.86	-15.28	6.84	-2.12	0.50	0.46
6	-1.6	117	-0.09	-15.61	-15.52	17.14	-5.23	0.50	0.88

Eastward phase propagation

n	$\phi_{20:04}$-$\phi_{07:35}$ (cycles)	wavelength (km)	period (days)	c_x (m/sec)	c_{gx} (m/sec)	x-phase (cycles)	t-phase (cycles)	zero crossing phase (cycles)	$\phi_{06:55}$-$\phi_{20:04}$ (cycles)
1	-1.1	1542	0.99	18.02	13.57	1.30	0.46	0.5	0.34
2	-1.1	1599	0.96	19.38	12.66	1.25	0.47	0	0.78
3	-1.1	1654	0.92	20.80	11.84	1.21	0.49	0.5	0.22
4	-1.6	1074	0.67	18.55	13.13	1.86	0.67	0.5	0.69
5	-1.1	1791	0.86	24.04	10.30	1.12	0.52	0	0.59
6	-1.6	1114	0.65	19.90	12.26	1.79	0.70	0.5	0.60

Notes: The second column is the number of cycles between the 07:35 and 20:04 profiles

The phase, $\phi_{06:55}$-$\phi_{20:04}$, equals x-phase - t-phase + zero crossing phase

c_{gx} is group velocity. Equations for group velocity are given in Appendix A.5

Table 6.6

Summary of Possible Wave Solutions to Figure 6.8

Wave type	n		period (days)	λ_x (km)	$\frac{\Phi_{2.3^o}}{\Phi_{9.5^o}}$	$\frac{\Phi_{2.3^o}}{\Phi_{10.3^o}}$	c_{gx} (m/sec)
Kelvin		eastward	0.21	280	2.2	2.3	15.6
RGW		eastward	0.6 1.0	760 1490	0.6	0.6	15.3 14.5
IGW	1 3 4 6	westward	~-0.1	< 200	2.4 0.7 1.2 1.0	2.1 0.8 1.0 1.1	-15.5
IGW	2 4	eastward	0.96 0.67	1600 1070	2.3 1.2	5.7 1.0	12.7 13.1

Table 6.6 shows expected amplitude relationships between the profiles which may further constrain the solutions. The observed ratio of the perturbation amplitudes of the 20:04 and 07:35 profiles, $|\Phi_{-2.3^o}/\Phi_{9.5^o}|$, is ~1 with the 2.3S latitude profile amplitude possibly being slightly larger. The observed ratio of the perturbation amplitudes of the 20:04 and 06:55 profiles, $|\Phi_{-2.3^o}/\Phi_{-10.3^o}|$, is ~2. Of the possibilities in Table 6.6, the IGWs of n = 3, 4 and 6 may provide the best amplitude match although none of the predicted amplitude ratios match those of the observations terribly well. In this regard, it is important to realize that when several wavelengths separate the occultation profiles (see Table 6.5, columns 7 and 8), the growth and decay of the wave amplitude across the wave packet in the zonal direction becomes important, an effect which is not reflected in the meridional envelope relation shown in Table 6.6. Therefore, for waves of zonal wavelengths < 400 km, the amplitudes of three profiles do not provide much constraint. A related issue is that the phase of solutions with very short zonal wavelengths is suspect because several cycles exist between profiles that small changes in the zonal wind field can substantially alter the phase between profiles.

All of the possible solutions in Table 6.6 have wavelengths of ~2000 km or less and periods of order 1 day or less, much smaller and faster than the dominant Kelvin and RGW modes quoted by AHL. As discussed next, a comparison of radiosonde profiles with an occultation also indicates the observed waves are of short period and zonal wavelength.

Implications of the occultation near the WTEA radiosonde

The second set of profiles we will examine are the occultation and nearby shipborne radiosondes mentioned in Section 6.3 located roughly 1000 km west of the International Date Line (figure 6.2b). The occultation and radiosonde profile show a very similar vertical temperature structure including the modulation of the temperature associated with a wave. The agreement demonstrates the accuracy and resolution of the occultation profile as well as the wave's large zonal wavelength implied by the 300 km of longitude separating the soundings. In combination with two occultations east of this occultation, I initially believed the wave might be a RGW. However, the lack of the wave's signature in radiosonde profiles taken 12 hours earlier and 36 hours later indicates the wave's zonal extent is probably too small to be a RGW. While a smaller scale IGW is the next natural choice, there are issues even in this case which may suggest the wave has a solitary character.

Combining occultation profiles with radiosondes can significantly constrain the properties of a wave. Unfortunately, given the sparsity of radiosondes in the largely oceanic tropics, spatial and temporal coincidence between occultation and radiosonde profiles is difficult to find. A fairly good spatial-temporal match exists between an occultation on May 4, 1995, at 12:40UT at 7.9S 167.5E and a shipborne radiosonde (SHIP WTEA) launched from 6.0S 170.4E at 12:00UT. Agreement between the two temperature profiles including the presence of a wave in the lower stratosphere is striking and was used as an example profile in reporting the initial GPS-MET results in *SCIENCE*

(Kursinski et al., 1996). Some care in interpretation is required because the radiosondes capture a wider spectrum of scales in comparison to the low-pass filtered vertical and horizontal scales of the processed occultation results. The simplest interpretation of the agreement between the vertical wavelength, phase and amplitude in combination with the ~320 km longitudinal separation and relatively short time separation of 40 minutes is the two profiles have sampled a common wave whose horizontal wavelength is large (>>300 km).

Based on its implied large horizontal scale, the wave should be apparent in other nearby occultations. The two closest occultations at 05/04-10:54UT (9.00N, 176.59W) and 05/04-23:24UT (0.29N, 175.41E) together with the 05/04 12:40UT occultation form a southwest to northeast diagonal across the equator near the international dateline spanning 16 degrees of longitude, 17 degrees of latitude across the equator and 12.5 hours in time. The vertical structures of the first and third profiles are somewhat out of phase with respect to one another between 60 and 25 mb but almost identical between 100 and 60 mb. The middle profile, which is very close to the equator, shows essentially no sign of any wave below the 40 mb level. The lack of wave at the equator and asymmetry about the equator between 60 and 25 mb would be inconsistent with a Kelvin wave and consistent with a RGW or an asymmetric IGW. However, the virtually identical structure just above the tropopause would require that the 100 mb to 60 mb interval be below the wave packet spanning the 60 to 25 mb interval.

This dubious large-scale RGW or IGW interpretation is further undermined by the additional constraints of the preceding and following SHIP WTEA radiosonde profiles. These two radiosondes were released 12 hours before and 36 hours after the May 4 at 12:00UT sonde. The surprise in light of the wave's implied large zonal scale, is that the signature of the wave is not obvious in either of the two profiles (figure 6.9). Instead, at

00hr on May 4, a shorter vertical wavelength structure (~1 km wavelength) is present bearing little relation to the 12 UT wave.

The 12 hour onset combined with equatorial wave group velocities of order 12-16 m/sec (Table 6.6) imply the amplitude envelope of the wave packet envelope rises from zero to a large value over less than 500-700 km of longitude. This leaves us with apparently conflicting constraints of a large (>> 300 km) zonal wavelength as implied by the close agreement between the radiosonde and occultation profiles near 05/04 12:00UT and a sharp wave amplitude envelope which rises over less than 700 km implied by the appearance of the wave within 12 hours.

Given the wave's rapid appearance, a RGW solution with a 2.7 km vertical wavelength and a period ~0.25 days would have a zonal wavelength of several hundred kilometers (Table 6.3). The shortness of period suggests an IGW solution, although the < 12 hour onset is rapid even for an IGW. To satisfy the >> 300 km requirement suggests wavelengths $\geq$ 600 km. However, since 600 and 1200 km wavelength IGW have periods of 12 and 24 hours (Table 6.5), the appearance of the wave in 12 hours and disappearance within the next 36 hours indicates the envelope of the wave is comparable to its wavelength which is inconsistent with the assumptions used in deriving the IGW and RGW equations. There are alternative scenarios:

The wave is a short zonal scale IGW propagating with a large eastward phase velocity and each profile has sampled essentially the same portion of the wave's short zonal extent.

The wave's zonal wavelength is smaller than 300 km. Since the occultation ray path is aligned close to north-south, the ~300 km averaging interval applies to the meridional direction which allows an integer number of zonal wavelengths between the occultation and radiosonde profiles, with $n\lambda = 300$ km.

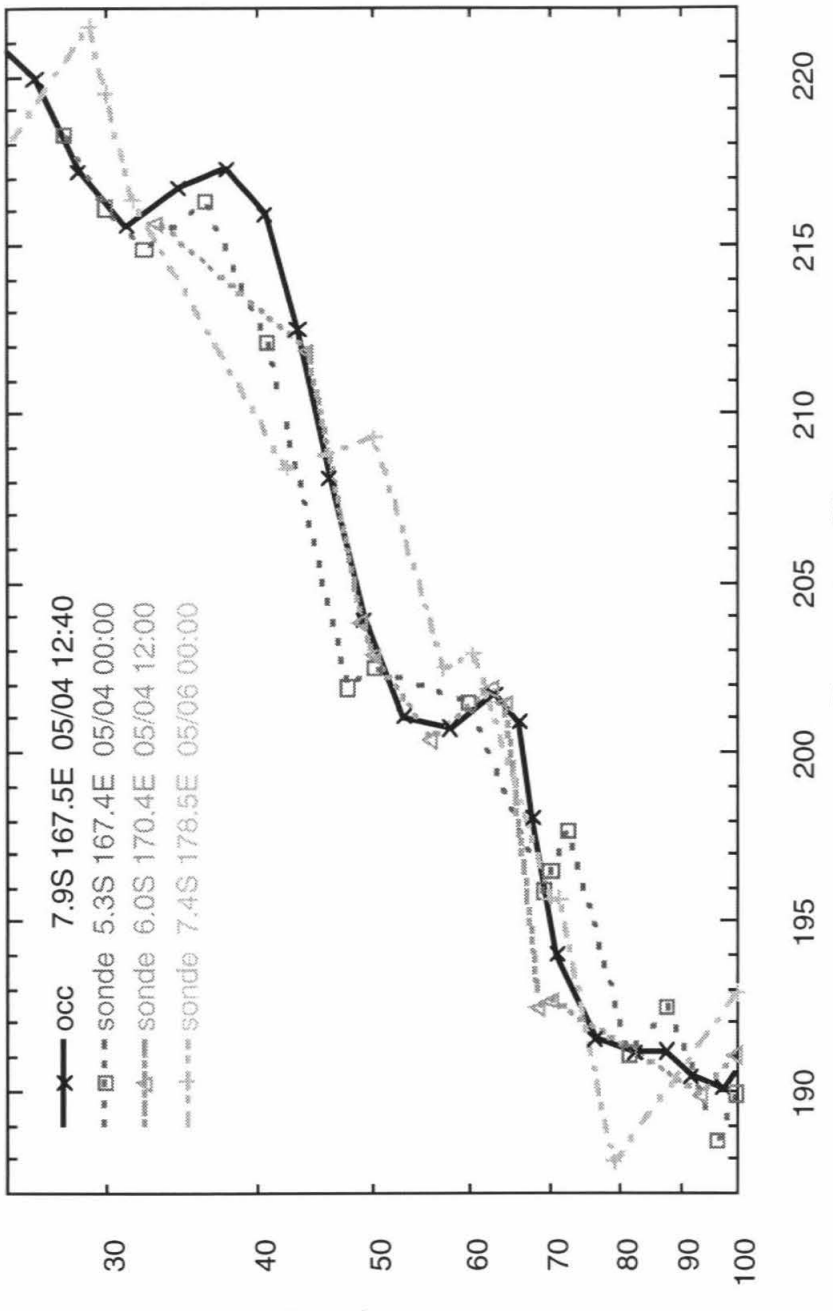

Figure 6.9 Comparions of occultation and nearby radiosonde profiles in south-central Pacific. Radiosonde profiles from the ship WTEA occur roughly 13 hours prior, 1 hour prior and 35 hours after the occultation on May 4 at 12:40 UT. The nearest radiosonde matches very well with the occultation. It is not clear if there is resemblance between the May 4 00:00 UT and May 6 00:00 UT sonde profiles and the occultation profile.

The wave's appearance in 12 hours could indicate a local source and a short existence.

The large and rapid amplitude modulation implies the wave may be solitary.

Scenario 1 is impossible because traveling 2.5° of longitude in 40 minutes requires a phase velocity greater than 100 m/s whereas realizable phase velocities of IGWs are of order 15 m/sec or less. Scenario 2 says that the phase and amplitude agreement between the radiosonde and occultation profiles is an unlikely coincidence. Scenario 3 would be surprising because of the wave's large latitudinal and longitudinal extent. This scenario would mean more waves exist in the equatorial region than already implied and that the waves must be removed locally via a rapid dissipation mechanism. It further implies large vertical group velocities such that the wave packet can extend vertically two or more wavelengths in the short time before the wave dissipates. It is not clear in Scenario 4 if it is possible for the wave to be solitary in the x-direction and still exhibit multiple wavelengths in the z direction. This idea needs to be pursued further.

In summary, when the two sets of observational constraints, the three occultation profiles and the radiosonde-occultation profile comparisons, are combined, it appears the observed waves must have relatively short zonal wavelengths, ~ 2000 km or less and more probably several hundred km or less. Still smaller scale solutions are possible as long as the wave's horizontal extent along the ray path is larger than the horizontal averaging interval of the occultation geometry. One could in fact argue that these profiles have sampled unrelated waves in which case the zonal scale of the waves will be even smaller.

The rapid appearance and disappearance of a wave and the near exact phase agreement at a 300 km separation in longitude raise the question of whether the observed equatorial waves match any of the four equatorial wave types. Clearly more occultation profiles need to be examined to better constrain the wave features and types. Higher sampling densities in the future might also determine directly the group velocity, both its

direction and speed, by tracking wave packet motion which would further distinguish the wave types as well as constrain their sources.

6.6 Thoughts Concerning Exchange Between the Troposphere and Stratosphere

Exchange between the troposphere and stratosphere has been a topic of intense research for more than a decade, driven largely by the need to understand the chemical cycle of life-enabling ozone and the flow into the stratosphere of cycle-altering anthropogenic constituents. One of the dominant channels of tropospheric to stratospheric flow is believed to occur across the tropical tropopause inferred from extremely low specific humidities in the equatorial lower stratosphere apparently caused by freeze drying of air as it rises across the extremely cold tropical tropopause. The debate over the style of vertical flow, a broad and slow versus local penetrative local convection, continues. While occultation sampling density is insufficient and along-track resolution too coarse to examine individual convection events, occultation temperature data can address the question of whether freeze drying associated with the slow upward migration is consistent with observed mixing ratios above the tropopause.

High wave activity in the occultation observations of the equatorial lower stratosphere raises questions about the role or roles these waves play in the exchange process. Two possible roles are suggested by the occultation observations. First, the waves modulate temperatures at and just above the tropical tropopause to the point that the tropopause is often ill-defined. Given the extreme sensitivity of saturation mixing ratios to temperature, the waves probably affect water mixing ratios in the lower stratosphere. Second, certain equatorial waves create vertical flow across the tropopause through net vertical displacement of air parcels and as such may be directly responsible for some exchange between the two regions.

Modulation of tropopause temperatures

The saturation vapor pressure of water vapor varies nonlinearly with temperature according to the Clausius-Clapeyron relation. To zeroth order, tropospheric water vapor mixing ratios are limited by temperature whereas in the stratosphere they are limited by the water vapor source. A wave's ability to alter moisture content depends on the wave's amplitude and period. Long period waves will reduce mixing ratios whereas short period waves may actually increase mixing ratios.

The amplitude of a wave determines the range of the temperature variations and saturation mixing ratios. To affect freeze drying of air, a wave must modify temperatures where they are otherwise minimum, i.e., the tropopause. It is not obvious that waves can in fact accomplish this because a sharply defined tropopause represents a stability boundary. In deriving the dispersion relations of the four equatorial wave types, N was assumed constant which is really only valid above the tropopause. In the uppermost 2 km of the tropical troposphere, the occultation-derived lapse rate generally decreases, indicative of reduced coupling to convective activity below and possibly increased coupling to waves above. The two profiles in figure 6.7 exhibit dramatically different tropopause structure which appear consistent with the phase and amplitude of the wave at higher altitudes. The 05/04 20:04UT profile at 2.3S has the more sharply defined tropopause near 90 mb (~17 km) and a minimum temperature of 189 K, some 5.4 K colder than the 90 mb temperature in the 05/04 07:35UT profile. This example indicates waves do modify minimum temperatures at the tropopause.

According to the Clausius-Clapeyron equation, the saturation vapor of water, P_w^*, is an exponential function of temperature, T:

$$P_w^* \;\alpha\; \exp\left[-\frac{0.622\,L}{R_d\,T}\right] \qquad (6.6.1)$$

where L is the latent heat of sublimation (~2.8×10^6 J/kg) and R_d is the gas constant for dry air. The 3 K perturbation representative of wave amplitudes observed in the occultation profiles will reduce the minimum saturation vapor pressure by 40%, which is a significant modification. Differences in the two temperature profiles in figure 6.7 cause their minimum saturation mixing ratios to be 3 ppmv near 17 km and 5.5 ppmv, near 15 km. Minimum mixing ratios observed by LIMS in December 1978 were roughly 2.5 ppm near 19 km [Remsberg et al., 1984]. The lower saturation mixing ratio is sufficiently close to the minimum mixing ratio to suggest that wave modulation may explain the observed lower stratospheric mixing ratios without need to invoke penetrative convective injections into the lower stratosphere. A more in-depth examination of the equatorial occultation data is clearly warranted.

Reducing temperature to form ice does not by itself guarantee freeze drying. Time is required for the ice to fall out the air parcel in which it formed and reduce total water content (vapor plus condensed phases) to the saturation mixing ratio. In the simplest case, freeze drying by a wave requires that ice form and fall out over the one-quarter to one-half of a wave cycle when air is colder than average. A problem is the descent rate of the ice crystals is not clear because it depends on the size of the ice crystals. Simulations of cloud ice particle formation at the tropopause by Jensen et al. [1996] indicate that 1-2 hour gravity waves produce little dehydration because the ice crystals are too small to descend much over the wave cycle. Vertical displacement of 500 m by 24 hour waves produce more dehydration ranging from 0.4 ppmv to 0.9 ppmv. The dehydration range depends on the number of particles formed. If many form, each particle tends to be small and descend slowly resulting in relatively little dehydration. If the number of nucleation sites is limited somehow, the particles grow larger and descend faster, significantly increasing dehydration.

Another possibility is ice formed by a cloud may alter the opacity of the air sufficiently to cause local convection as suggested by Danielsen [1993] such that the top layer of ice may descend quite rapidly. Enhanced radiative cooling might also nonlinearly amplify the wave's amplitude during its cold phase. Another question is what happens to the water if the time required for ice particles to fall out of a parcel increases beyond a wave cycle. In the presence of a short period wave, the fraction of total water which is above the saturation vapor pressure of the unperturbed temperature spends less time in the ice phase than in unperturbed conditions creating at least the possibility that average total water content of air parcels increase relative to unperturbed conditions. Whether or not such an increase is possible depends on the vertical location of the water while in the vapor form. Freeze drying will occur even if descending water moves back and forth between the ice and vapor phases as long as there is net downward displacement during the ice phase and the displacement is maintained during the vapor phase. If, on the other hand, the vapor is mixed around during the warm phase, memory of the vertical displacement of the ice will be lost. This is a complicated issue and the microphysics of the ice formation must be understood to resolve it.

Upward migration of air parcels due to equatorial planetary waves

The mechanism or mechanisms causing air to rise across the tropical tropopause remains a fundamental question. Holton has proposed a scenario where mid-latitude waves break and deposit momentum in the mid-stratosphere, creating a suction pump which moves mid-stratospheric air from lower latitudes to higher latitudes and, by continuity, pulls air up across the tropical tropopause. The impact of equatorial waves on tropopause structure suggests they might play a role in forcing vertical advection of air across the tropopause, an idea which I have begun to investigate. Equatorial waves which produce meridional winds such as RGWs and IGWs move air parcels in trajectories which result in net vertical displacement over a wave period. I have verified this with both a Lagrangian

parcel trajectory integration and using the transformed Eulerian mean (TEM) equations which produce the same result for waves of small amplitude at the equator (Appendix A.6). In contrast, parcel motion produced by Kelvin waves, which lack meridional winds, integrates to zero over a complete wave cycle. The sign of the vertical displacement depends on the sign of the vertical component of group velocity. RGWs and asymmetric IGW's create upward displacement at the equator whereas symmetric RGWs produce downward displacement.

Figure 6.10 shows examples of integrated parcel motion in the y-z plane due to a RGW. Parcels within several degrees of the equator are clearly displaced upward. Figure 6.11 shows an x-y cross section of the vertical displacement integrated over a cycle of a RGW with a 1K amplitude, a vertical wavelength of 6 km and a four day period, the dominant RGW mode according to AHL. Figure 6.12 shows zonally averaged vertical displacement velocities for several different RGWs. Zonally averaged vertical displacements shown in figure 6.12 are upward within ~7^{o} of the equator and downward at higher latitudes. The peak zonally averaged, vertical velocity at the equator for a 1K amplitude is ~0.14 mm/sec (~12 m/day) or ~50 m per wave cycle, in approximate agreement with 0.1 mm/sec estimated from eq. (A.6.7). The peak vertical displacement velocity increases to 0.32 mm/sec (28 m/day) for a 2K RGW. Peak downward velocities of -0.06 and -0.22 mm/sec for 1 and 2 K waves respectively occur near $\pm$12^{o} latitude. As wave amplitude increases, the peak average velocity moves off the equator due to increasing meridional excursions of air parcels shown in figure 6.11. The region of upwelling widens nonlinearly as the wave amplitude increases further.

IGWs produce more complicated patterns of net upward and downward displacement. The lowest IGW mode (n=1) which is symmetric about the equator produces systematic downward motion within 5^{o} of the equator peaked at the equator and upward motion at higher latitudes peaked near 8^{o} latitude (figure 6.13). The lowest order

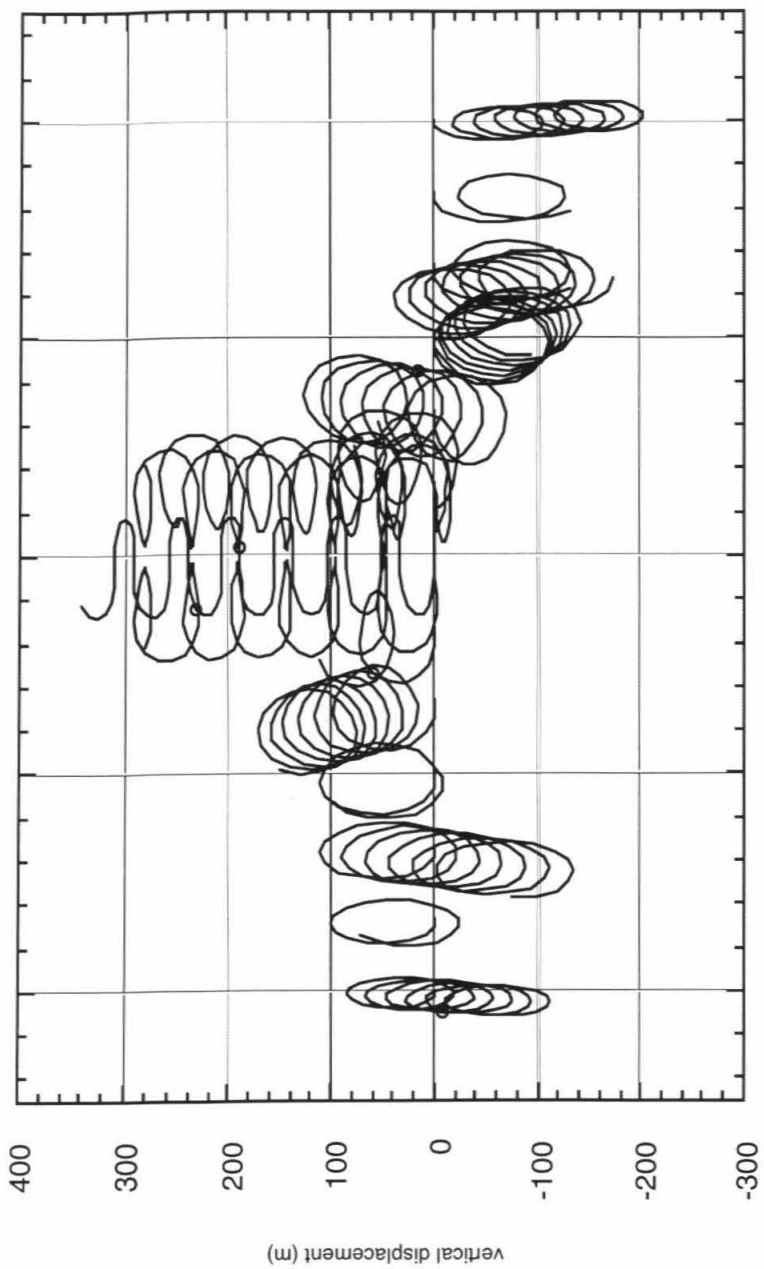

Figure 6.10 Air parcel trajectories in the vertical-meridional plane during the passage of a Rossby-Gravity Wave of 1 K amplitude, wavenumber 4, 8 km vertical wavelength and 3.6 day period

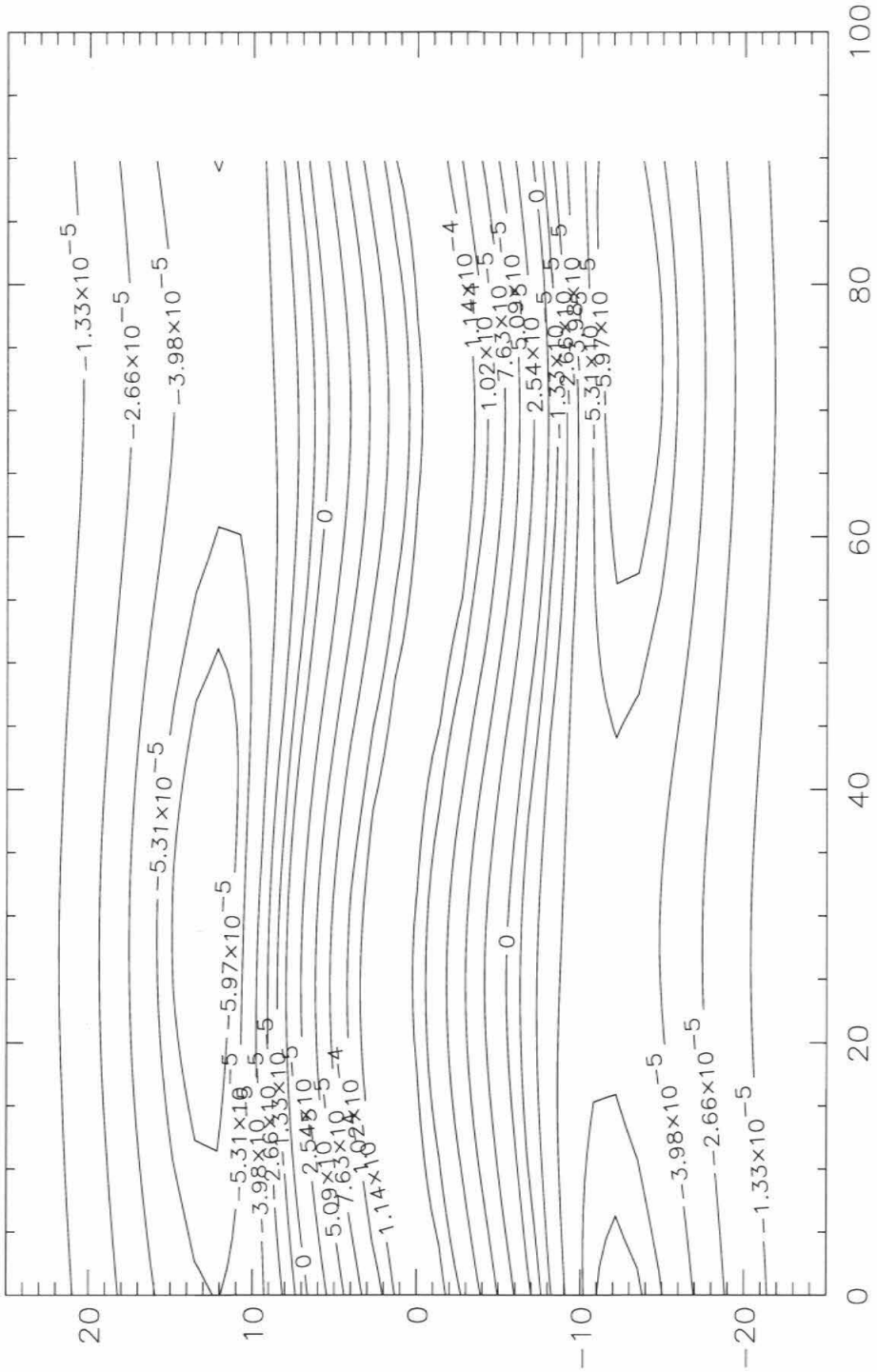

Figure 6.11 Average vertical velocity in m/sec of air parcels displaced by a Rossby-gravity wave of 1 K amplitude, wavenumber -4, vertical wavlength = 6 km and a 4 day period

asymmetric IGW mode (n=2) produces upwelling within a couple degrees of the equator peaked at the equator, downwelling between 2.5 and 10° latitude peaked near 6° latitude and upwelling at higher latitudes peaked around 12.5° latitude (figure 6.13). The meridional width of the upwelling region is narrower than that of the RGW by a factor of ~2.

The TEM energy equation

The impact of waves on net or residual vertical motion must be determined in the framework of the Transformed Eulerian Mean (TEM) equations. The TEM energy equation is

$$\bar{\theta}_t + a^{-1}\bar{v}^*\bar{\theta}_\phi + \bar{w}\bar{\theta}_z - \bar{Q} = -\rho_0^{-1}\left[\rho_0\left(\overline{v'\theta'}\,\bar{\theta}_\phi / a\bar{\theta}_z + \overline{w'\theta'}\right)\right]_z \tag{6.6.2}$$

where $\bar{v}^*$ is the residual mean meridional velocity defined as

$$\bar{v}^* \equiv \bar{v} - \rho_0^{-1}\left(\rho_0\overline{v'\theta'} / \bar{\theta}_z\right)_z \tag{6.6.3}$$

The mean crosscorrelation perturbation terms in eq. (6.6.2) and (6.6.3) can be evaluated using the relationships in Appendix A.5. With an RGW under the assumptions of steady state, constant N (which was used to derive the equation set of equatorial waves) and negligible zonal wind, the resulting energy equation is

$$\frac{\bar{v}\bar{\theta}_\phi}{a} + \bar{w}\bar{\theta}_z + \overline{v'\theta'}\left(-\frac{\tan\phi}{a} + \frac{1}{y} - 2\frac{\beta|m|y}{N}\right) = \bar{Q} \tag{6.6.4}$$

where

$$\overline{v'\theta'} = -\frac{\hat{v}_0^2}{2}\exp\left[-\frac{\beta|m|y^2}{N}\right]\exp\left[\frac{z(\kappa+1)}{H}\right]\omega y m\frac{H}{R} \tag{6.6.5}$$

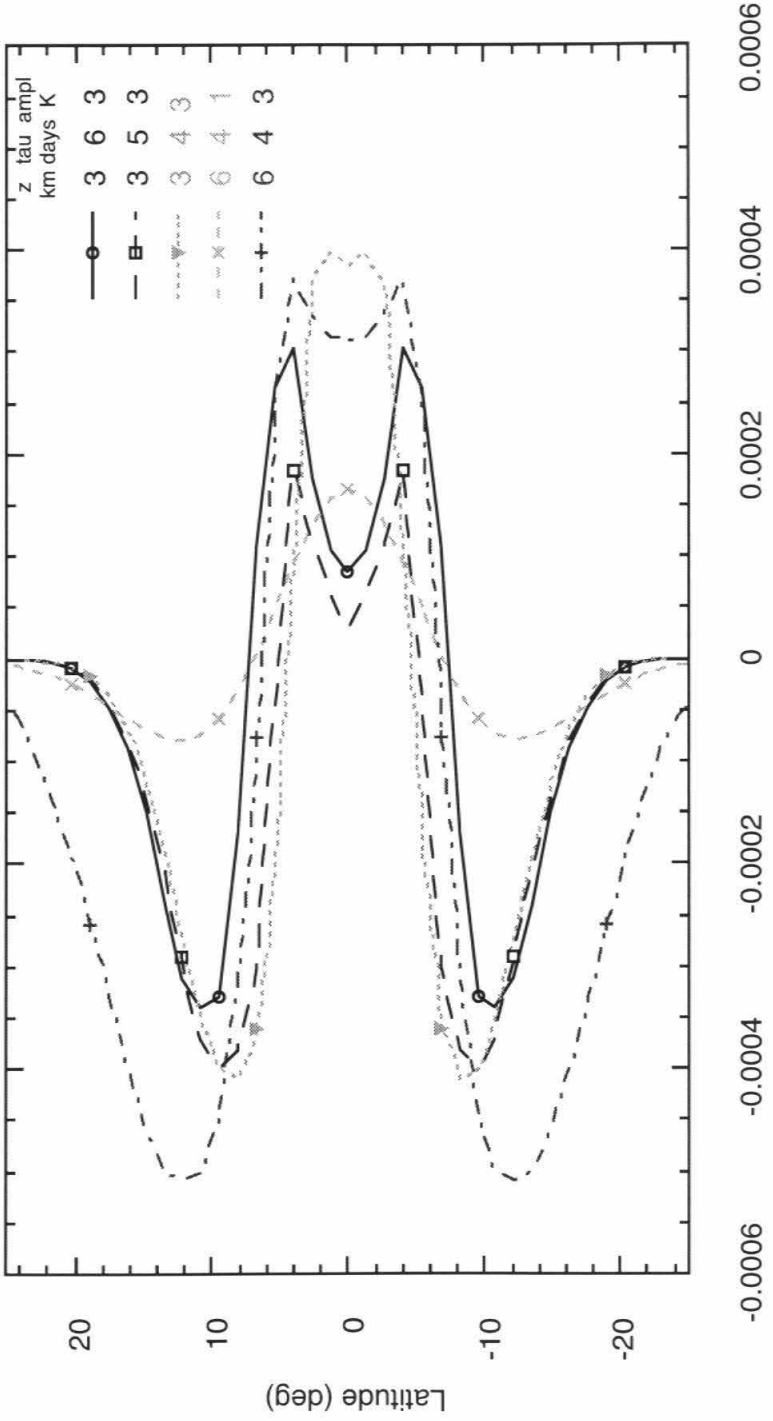

Figure 6.12 Vertical component of integrated Lagrangian velocity of air parcels due to equatorial Rossby-gravity waves with westward phase velocity

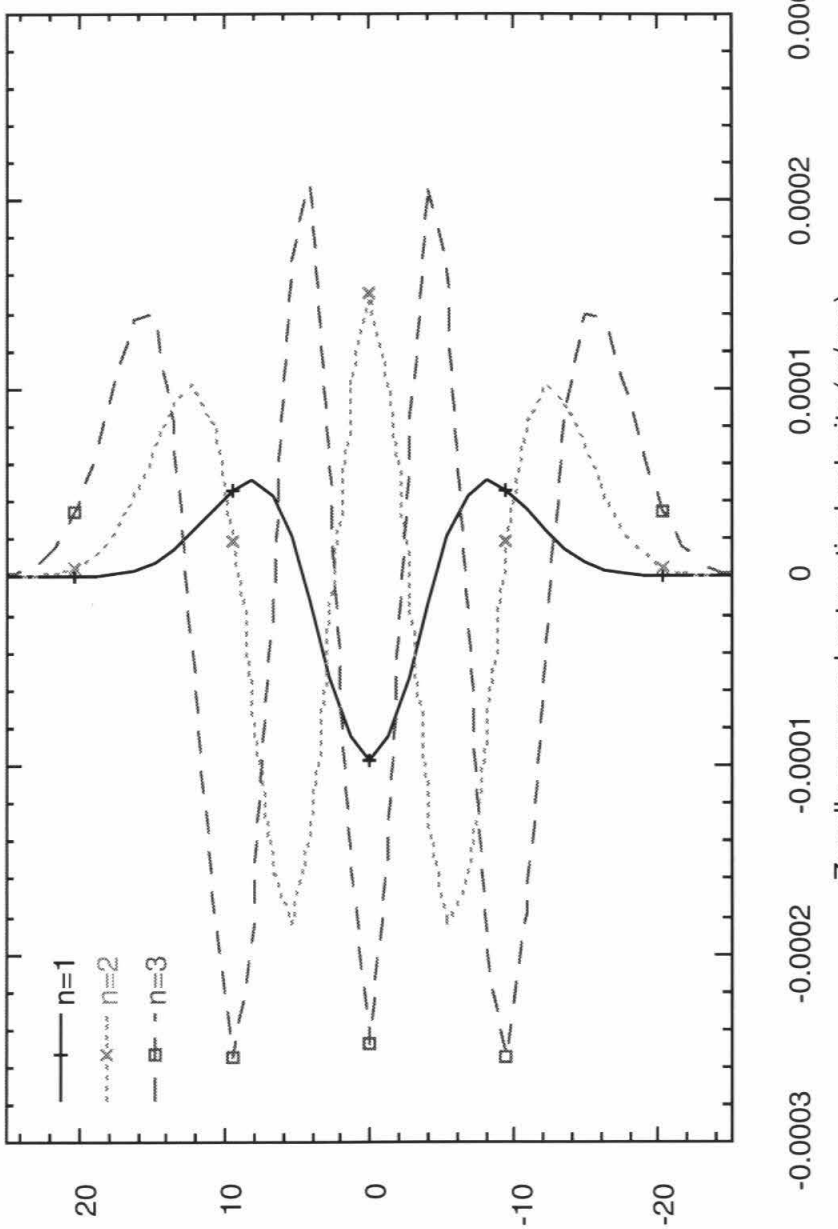

Figure 6.13 Zonally averaged vertical component of integrated Lagrangian velocity of air parcels due to three lowest order inertio-gravity wave modes of 2.7 km vertical wavelength, 12 hour period, 3 K amplitude and westward phase velocity

Note that $\overline{v'\theta'} > 0$ because ω and m have opposite signs for a positive vertical component of group velocity.

Order of magnitude estimates concerning vertical velocity and heating balance can be made using eq. (6.6.4) and (6.6.5). First, we assume the longitudinal term is small because meridional gradients at low latitudes are very small (Section 7.5.3). If net heating ($\bar{Q}$) is 0, the vertical mean velocity required to balance $\overline{v'\theta'}/y$ in the tropical lower stratosphere using the dominant RGW mode according to AHL (1 K amplitude, $\hat{v}_0 \sim 2.5$ m/s, a four day period and 6 km vertical wavelength) and $d\theta/dz \sim 16$ K/km is $\bar{w} = 0.08$ mm/s. Since $\overline{v'\theta'}$ scales as the square of the wave amplitude, a 2K wave would require 0.3 mm/s which is similar to the 0.2 to 0.3 mm/sec rate of vertical motion inferred by Boering et al. [1996] from studies of tracers. If the mean vertical flow were 0 and the wave-induced motion were balanced by heating, the heating required to balance the dominant RGW mode at the equator is ~0.1 K/day and 0.4 K/day if the wave amplitude is doubled which are in approximate agreement with that derived with UARS data [Eluszkiewicz, 1996]. RGW's also play a key role in the theory of the Quasi-Biennial Oscillation (QBO), a 22 to 34 month cyclical pattern of equatorial stratospheric winds between 80 and 10 mb. If they are also important in driving exchange between the equatorial troposphere and stratosphere, a correlation is implied with the QBO during the QBO phase when winds in the lower stratosphere are eastward and RGW's can propagate upward. As mentioned, Kelvin waves do not produce any net displacement of parcels. The implication is that RGW and IGWs are important and should be considered in modeling flow across the tropical tropopause.

Radiation and vertical motion across the tropopause

If we take the balance between the wave-induced and mean vertical motion required when heating is zero, then we discover that the wave-induced vertical flow at the equator

(eq. A.6.12) is equal and opposite to the mean vertical velocity such that there is not net vertical flow. Therefore, according to the TEM energy equation, the dominant factor determining vertical motion of air across the tropopause is net radiation which is not surprising because in the absence of diabatic heating, flow across the tropopause would be adiabatic and the tropopause would be replaced by a region of extended adiabatic lapse rate. Upward flow requires that air be heated because potential temperature is increasing. In downward flow, air must cool because potential temperature is decreasing. This conclusion is a bit disconcerting because it says that if there is no net radiation at the tropopause, all the net displacement due to RGW's and IGWs is somehow magically balanced by the mean vertical flow (whatever that is).

Carrying this further, a conjecture can be made based on the role of radiation and the tropopause observations discussed in Section 6.4. In the discussion on tropopause structure at various latitudes in figure 6.5, it was pointed out that a relatively short transition layer exists on the top side of each and every tropopause between the tropopause temperature minimum and the radiative equilibrium stratospheric temperature profile above. The transition between the tropopause and stratospheric radiative equilibrium curve is distinct, suggesting it is not due to radiative transfer unless there is a sharp change in opacity. The transitional temperature gradient is also more positive than the overlying radiative equilibrium curve indicating that temperatures in the immediate vicinity of the tropopause are colder than temperatures at radiative equilibrium. The coldness of the tropopause and overlying transition structure is consistent with upward motion of air parcels in a low dissipation troposphere which systematically overshoot the height of radiative equilibrium in the upper troposphere. Since air at the tropopause and transition region is colder than radiative equilibrium temperatures, the air is warmed radiatively and must therefore rise unless a compensating eddy cooling term exists. This suggests that, except under perturbed conditions, air is rising on average across the tropopause *at all latitudes*. Stratospheric air must return to the troposphere across discontinuities or folds in

the tropopause. Due to the thermal wind equation, these folds coincide with the jet streams at the transitions between mid-latitude and tropical tropopauses and the polar and mid-latitude tropopauses. These conclusions are general implying that in any low dissipation planetary atmosphere of sufficient mass (surface pressure ~ 1 bar or greater) to produce a well defined tropopause, discontinuities in the tropopause must exist in order to provide a path for returning stratospheric air. Injection of stratospheric air into the troposphere during mid- to high latitude storms is consistent with this conclusion as these storms are associated with tropopause folds. The vertical extent of radiative imbalance does not extend far into the stratosphere implying the upward motion dies off vertically requiring a compensating horizontal motion to satisfy continuity. The conjecture can be carried to the cold Antarctic winter tropopause where a transition zone is seen (figure 6.6 g and h) implying air is rising across that tropopause even though the overall atmosphere is cooling. The tropopause continues to warm because vertical motion is required to remove heat from the tropopoposhere. This conjecture should be testable by combining temperature profiles, radiative calculations and tracer measurements.

6.7 Conclusions

Occultation data is consistent with ECMWF weather analyses at the ~1 K level. Vertical resolution has been demonstrated at the ~1 km level. Accuracy can only be determined to the level of the analyses. The occultation data are probably not yet achieving the accuracy predicted in Section 3.

The overall climatological behavior of the profiles in figure 6.6 agree with the latitude versus height temperature climatology for July derived from Nimbus 5 and 6 satellite data and radiosondes (AHL, figure 5.1).

The equatorial waves examined thus far in the April-May 1995 occultation data exhibit qualities summarized as follows. Most occultation profiles within 20 to 25 degrees of the equator exhibit "waviness" in the lower stratosphere consistent with the predicted meridional envelope of equatorial waves. This equatorial wave activity is distinctly larger than that observed at higher latitudes. Vertical wavelengths range from 1 to several km with amplitudes near ~1 to 3 K. The altitude range of waves extends from between 200 and 100 mb through between 30 and 10 mb consistent with a tropospheric source and radiative damping of slowly upward propagating waves mentioned in AHL. The waves seem to be present at all longitudes sampled by the occultations which is a global (but uneven) distribution implying either a global distribution of sources or long-lived waves. The waves exhibit a vertical and horizontal structure which seems complex in comparison to the simple large scale structure of theoretical RGW and Kelvin waves. The complexity is not surprising given the variable nature of convective activity in the Earth's troposphere.

The waves modify thermodynamic structure near the tropopause and must influence exchange between the troposphere and stratosphere in terms of temperature and therefore saturation mixing ratios. Long period waves must lower water content at the tropopause. Short period waves may raise water content at the tropopause. They may also affect vertical advection around the tropopause. To the extent that they do, there may be a correlation between vertical motion and the phase of the QBO.

Clearly higher spatial and temporal sampling density is needed to identify the observed wave types along with their sources and sinks including coupling to tropospheric convection. The soliton hypothesis may need to be explored if further observations indicate the wave envelope is indeed a problem. The climatology of the waves should be determined including seasonal and longitudinal dependence requiring an automated processing approach to isolate the wave perturbations.

The relationship between equatorial waves and ice needs to be considered. An ambiguity exists in interpreting wave-like structure in radiosonde profiles which may be waves or temperature structure due to ice clouds, and the two may be coupled.

Observations of equatorial waves may yield a monitor of convection frequency and intensity on regional and global scales if and when coupling between convection and the waves is understood. This could prove important as a diagnostic of Earth's heat engine and climate.

Vertical flow may be systematically upward across well defined tropopause in low dissipation atmospheres. If so, tropopause folds or discontinuities are required to provide a path for stratospheric air to return.

7. Initial Water Vapor Results from GPS-MET

7.1 Introduction

As discussed in Section 4, water vapor is extremely important and poorly observed globally in the vertical dimension. GPS occultations can remotely sense tropospheric water vapor with vertical resolution presently unobtainable from space, a vantage point required for global coverage. GPS observations are insensitive to clouds which plague IR observations and should yield a consistent level of accuracy regardless of cloud fraction. As an active system GPS observations are also insensitive to variations in surface emissivity which complicate interpretation of passive observations.

GPS observations yield refractivity. To extract water vapor, temperatures from the ECMWF global analyses available every 6 hours are interpolated to the occultation locations and used to estimate and remove the dry contribution to refractivity. Water vapor results derived in this manner are referred to as GPS-ECMWF Water vapor (GEW). We present and discuss a set of GEW profiles derived from GPS-MET data. The occultation data acquired during the period of June 21 through July 4, 1995, routinely penetrate to within 1 to 3 km of the surface, offering the best qualities yet available from GPS-MET for characterizing the lower troposphere.

The latitude-longitude distribution of the occultations is given in figure 7.1. Although sampling is sparse, the GPS occultation-ECMWF based water vapor (GEW) data offers the first globally-distributed high-vertical resolution water vapor dataset. We compare GEW results with other data sets to assess accuracy. However, there are no absolute standards against which to compare, and the GEW data therefore also serves as an independent data source to assess current state of knowledge and biases related to vertical structure unresolvable using TOVS data which is assimilated in the ECMWF analyses. Weather analyses such as ECMWF assimilate much available in-situ and remote sensing

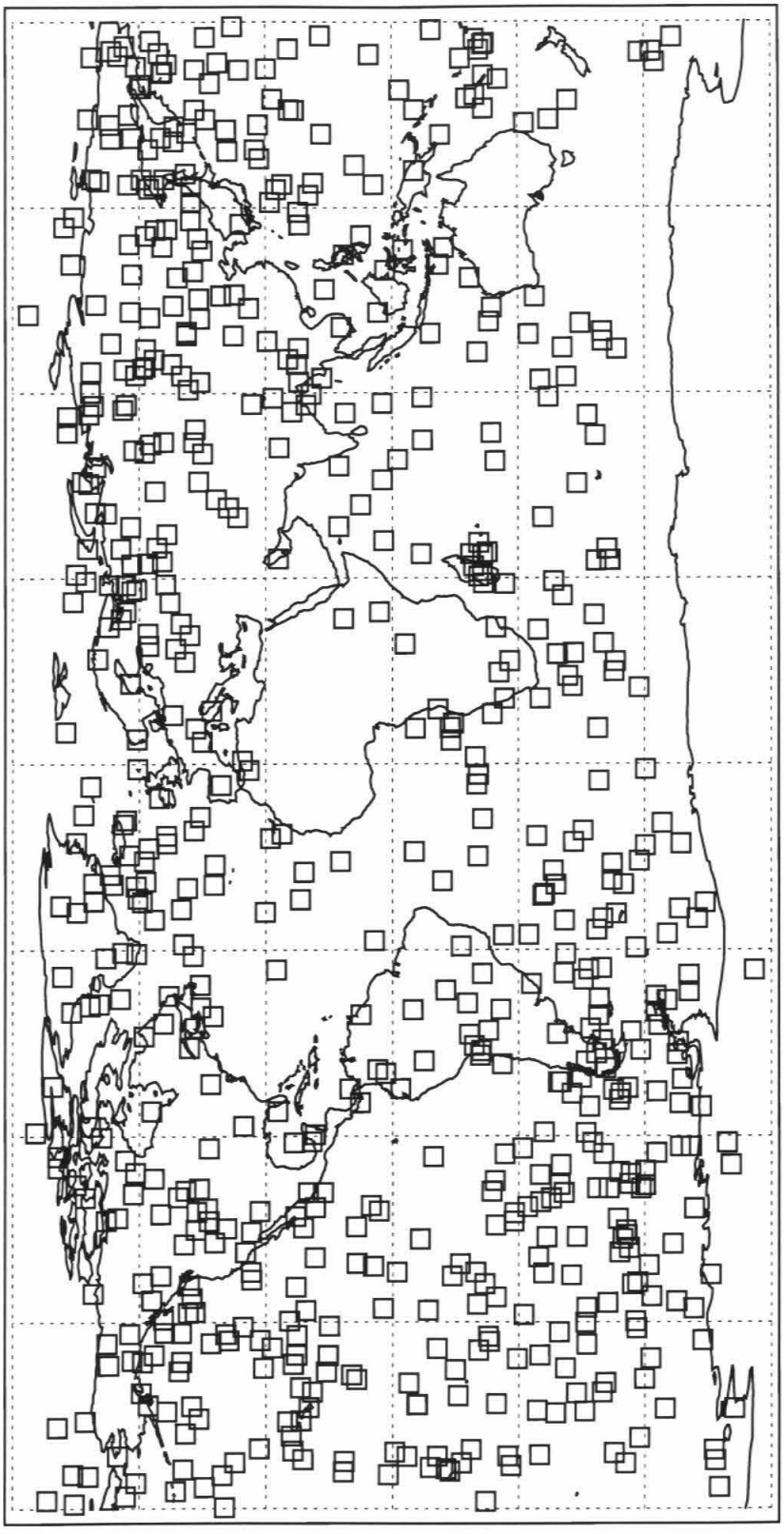

Figure 7.1 Occultation coverage for June 21 to July 4, 1995

data in a physically consistent way and provide the best estimates of the global atmospheric state. As such they are extremely valuable not only for weather prediction but as climatological databases for process studies and climate research, and assessing and hopefully eventually improving their accuracy is an important goal. The averaging inherent in the limb viewing geometry improves the accuracy and spatial extent of utility of the GPS observations beyond that of individual profiles.

The purposes of this section are to

Demonstrate the refractivity retrievals in the mid- to lower troposphere from GPS occultations.

Isolate and derive water vapor structure in the mid- to lower troposphere using GPS refractivity and ECMWF temperature analyses

Derive a brief latitude vs. height climatology using 660 occultation profiles acquired during June 21 to July 4, 1995

Compare and contrast GPS occultation-ECMWF temperature based water vapor (GEW) climatology with that of the ECMWF humidity climatology for the period and the P&O June-July-August (JJA) climatology to evaluate all three data sets

Begin an interpretation of the water vapor distribution

We begin with an overview of the instrument and processing system (Section 7.2). In Section 7.3, the accuracy prediction of Section 4 is extended to include relative humidity and northern summer conditions. Examples of individual profiles are compared with ECMWF and nearby radiosondes when we can find them (Section 7.4). In Section 7.5, we present the results of assembling the water vapor profiles into a latitude vs. height climatological snapshot of atmospheric water vapor for the 11 day period. The 2-D snapshot is compared to the ECMWF climatological snapshot of moisture derived at the same times and places as well as the complete ECMWF climatology for the period and the P&O climatology for JJA. Statistical comparisons are made against the ECMWF analyses

and we use the GEW climatology itself to infer accuracy. We conclude with some embryonic interpretation of observed behavior and control of atmospheric water vapor (Section 7.6).

7.2 Processing overview

In this section we discuss details of the data acquisition and signal processing unique to the derivation of water vapor during the June 21-July 4, 1995, period.

Special features of the June-July 1995 acquisition period

The June 21 -July 4, 1995, period was one of four periods thus far during the GPS-MET mission when the GPS Anti-Spoofing (AS) encryption was turned off (the other two being three week periods in April-May, 1995, October 1995 and February 1997). The signal strength of the L2 signal (required for ionosphere correction) as acquired by the prototype receiver is significantly higher during AS-off allowing deeper penetration into the troposphere during these periods. The June-July period is particularly useful for occultation water vapor studies because as northern summer, maximum specific humidities exist where the ground truth is most accurate.

Further, during the June 21 -July 4, 1995, period the GPS-MET receiver had special software, developed at JPL, uploaded to improve signal tracking in the lower troposphere. Normally, the receiver operates in a phase-lock mode where a model of the signal carrier phase and modulation is continually updated in real-time. If the signal phase or amplitude changes sufficiently rapidly or the signal disappears momentarily, the receiver loses phase lock and the data acquisition for that occultation is concluded. A general problem is that sharp layers in the atmosphere, particularly those associated with water vapor, can cause the occulted signal to drop below the receiver's detection threshold and then reappear several seconds later (Section 2.4). Unfortunately, with a simple phase-locked loop system, the vertical extent of the atmosphere below a sharp layer is lost. A

special "bridge" mode was in operation during the June-July period referring to the ability of a GPS receiver mounted in an automobile to track when the automobile passes momentarily under a bridge. In this mode, when the signal amplitude drops below the detection threshold, the receiver extrapolates the phase for up to 5 seconds according to the most recent model waiting for the signal to reappear. If it does and the model is sufficiently accurate, phase lock will resume immediately and the occultation continues allowing occultations to probe significantly deeper into the troposphere. The final 5 seconds of every occultation acquired in this mode must be removed as the receiver is systematically out of lock during this period.

The altitude coverage of the occultations for this period versus latitude is shown in figure 7.2b. Most of the occultations extend to below 3 km altitude and a large fraction extend to within 1 km of the surface. The results for the April-May 1995 period shortly after launch are shown in figure 7.2a. Clearly the flywheel mode has significantly extended the lowest altitudes toward the surface. Data acquisition with higher SNR and improved signal processing which dramatically reduces sensitivity to AS will further improve results beginning with the launch of SUNSAT in late-1997.

Screening process

In assembling the set of occultations for the 2-D GEW climatology, screening criteria were defined consisting primarily of three checks: 1. the profiles must not contain jagged unphysical structure, 2. water vapor abundances had to be at crudely climatologically reasonable and 3. humidities at low altitudes had to be positive. Negative humidities values at higher altitudes were selected if the structure at lower altitudes was positive and somewhat smooth. A goal was to derive representative mean humidities at high altitudes even when individual profiles might be quite noisy. The approach was to allow selection of occultations with negative and supersaturated values at higher altitudes to

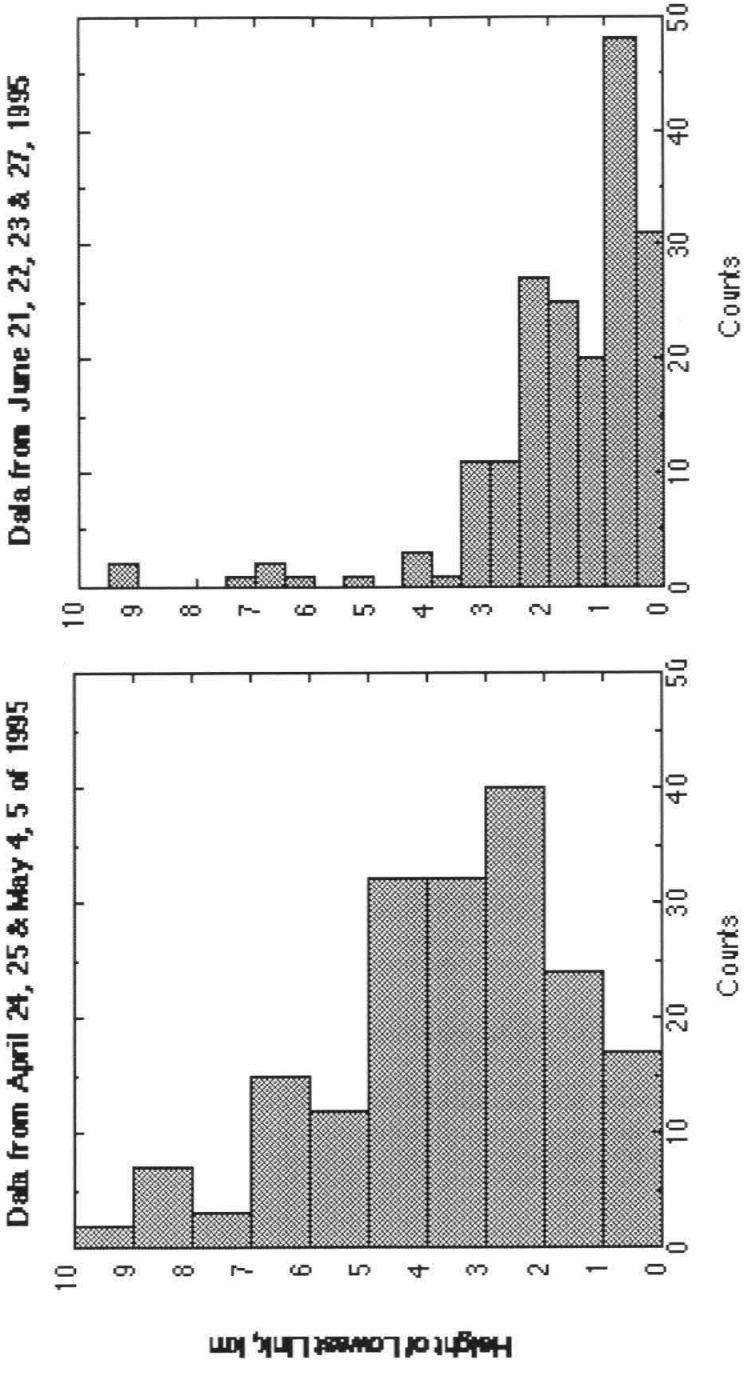

Figure 7.2 Histograms of Lowest Height of GPS/MET Occultations Before (April and May) and After (June) Tracking and Processing Enhancements

leave the mean unbiased by the selection process. Of the roughly 1000 occultations acquired over June 21-July 4, 1995 (minus June 24-26), 660 passed the selection criteria.

Processing approach

The derivation of water vapor from the occultation observations described here uses temperature information from the nearest 6 hour ECMWF global analysis. ECMWF was chosen as the source of temperature information needed to isolate water vapor because of its combination of global coverage, high accuracy as demonstrated in the temperature comparisons at higher altitudes (Section 6.3) and its high resolution temporally (every 6 hours), vertically (31 levels to 10 mb) and horizontally (~100 km). Further ECMWF has converted much of the available data to a resolution similar to the GPS observations. The ECMWF temperature information is spatially interpolated to the location of each point in the occultation profile. The derivation of water vapor from refractivity is initiated in the troposphere at the height where the occultation-derived temperature exceeds 230 K. Below this height the ECMWF temperatures are used. The hydrostatic integral below the 230 K level uses the ECMWF temperatures plus the pressure at the 230 K level which is determined by hydrostatically integrating the density derived from the occultation refractivity profile. The occultation hydrostatic integral is initiated at 50 km. The temperature profile combined with the constraint of hydrostatic equilibrium is used to remove the contribution of the dry density to refractivity and isolate the water vapor contribution.

According to the analysis in Section 4, specific humidities at 230 K are too small to be derived accurately from GPS occultation data. Nonetheless, initiating the water vapor derivation at the 230 K altitude allows the accuracy of the combined error biases of N, P and T to be assessed. Near 230 K, small errors in N, P and T produce fractional humidity errors which can exceed the fundamental humidity range from zero to saturation. As

mentioned, in the selection process, derived water vapor profiles with negative humidities for temperatures near 230 K were allowed to pass to avoid introducing biases in the resulting climatological averages. At low altitudes, negative or supersaturated values were thrown away. One could modify the temperatures to produce positive and subsaturated water vapor values near the 230 K level. However, these modifications might introduce biases which could dominate the climatologically averaged results. Further, altering upper troposphere temperatures requires modification of the entire temperature profile which can alter atmospheric stability in undesirable and dynamically inconsistent ways. Modification of the temperatures can really only be done in a self-consistent physical model such as the ECMWF weather model which is beyond the scope of the present effort. Nonetheless, it is clear, based on the analysis in Section 7.5, that some adjustment of ECMWF temperatures at high southern latitudes is needed to reduce noisy and systematic biases in temperatures there. By allowing impossible values of water vapor to pass the selection criteria, the mean value may lie outside the realm of physical possibility. In such cases, biases in the occultation data and assumed temperatures become evident and can be quantified.

One additional change was made to the occultation processing for the 2-D climatology which significantly improved the quality of the ionosphere correction at low altitudes. The ionosphere correction uses the L1 and L2 GPS signals to isolate and remove the dispersive effect of the ionosphere (eq. 2.3.1). A problem exists because the L2 signal is weaker than the L1 signal and the L2 signal amplitude therefore systematically drops below the signal detection threshold before the L1 signal in the lower troposphere due to defocusing. If early loss of lock of the L2 signal is not properly taken into account in the retrieval process, the ionosphere-corrected occultation data becomes extremely noisy in the lower troposphere. Since the ionosphere contribution to the occulted signals varies slowly when the ray path tangent height is near the surface, loss of the L2 signal can be dealt with via extrapolation. The processing was modified to take the ionosphere correction while L1

and L2 were still both in lock and extrapolate it through the period where L2 is out of lock but L1 remains in lock. This change has produced a dramatic improvement in the quality of lower troposphere retrievals.

The geometry of comparisons with ECMWF

The occultation results are derived using a center of curvature defined by fitting a circle to the plane through the Earth ellipsoid containing the transmitter, receiver and ray path tangent height. The vertical coordinate of the ECMWF model structure is pressure. The lowest level of the model is defined relative to sea-level, a surface of constant geopotential. In order to compare occultation results with ECMWF, the heights of the occultation results are defined relative to the sea-level geoid. The correction between height defined relative to the ellipsoid and the geoid varies by as much as 100 m.

Future improvements in processing

At the moment, the phase data acquired while the signal strength is below the detection threshold is presently used in the frequency estimation process. In the near future, data from these periods will be removed prior to the frequency estimation process.

It is clear that a refractivity retrieval based on both phase and amplitude data will further improve lower troposphere results because the amplitude information places a strong constraint on the vertical second derivative of the refractivity which must be met in the retrieval process.

In the future, when refractivity is assimilated into weather models, the model will separate refractivity into the dry and wet atmospheric densities in a more optimal fashion.

Data types

The following three humidity-related quantities have been derived in the GEW climatology:

Specific humidity, q, which is the water vapor mass mixing ratio typically in units of g/kg.

Relative humidity which is the amount of vapor present relative to the maximum at saturation and provides an indication of the role of temperature in controlling water vapor concentrations.

Equivalent potential temperature, defined in Section 7.5.7, which is an entropic measure of the conditional stability of the atmosphere which includes the latent heat of evaporation of water. It is conditional because air must reach saturation to release latent heat.

7.3 Predicted GEW Accuracy

Before discussing the water vapor results, some additional discussion of predicted accuracy is appropriate. In Section 4, we estimated the accuracy to which water vapor can be derived for annual mean conditions based on the relationship between refractivity, temperature, pressure and water vapor combined with estimated temperature accuracy and simulated refractivity accuracy. Here we will extend that analysis to June-July-August (JJA) conditions. Further we estimate the accuracy of relative humidity estimated from a combination of GPS refractivity and model temperature. Estimates of accuracy of the actual observations will be made in Section 7.5 as the actual results are discussed.

Expected fractional water vapor accuracy for JJA

Establishing the accuracy of humidity derived from actual GPS observations presents a challenge because no absolute standard exists for comparison. To define a baseline of expected accuracy, we use the approach in Section 4 with the P&O JJA moisture and temperature climatology and assumed temperature accuracy of 1.5K to produce the rms accuracy of individual water vapor profiles shown in figure 7.3. In comparison to predicted annual mean behavior in figure 4.1, the accuracy contours have shifted northward toward the wetter summer hemisphere.

The accuracy contours are representative of individual profiles of q. When individual profiles are averaged, the contribution of random errors in occultation refractivity and ECMWF temperature will average down as the square root of the number independent points averaged. As discussed in Section 4, climatological averages may be a factor of 3 better than the results in figure 7.3 because simulated refractivity accuracy averages down to very small mean error (Section 3.14) such that accuracy should be limited by model temperature accuracy which may be of order 0.5 K in the mean. Averaging ~20 GEW profiles which is representative of the averaging in the 2-D snapshot climatology discussed in Section 7.5 will reduce random error by a factor of 4.5.

Relative humidity accuracy

Relative humidity is defined as

$$U = P_w/P_w^* \qquad (7.3.1)$$

where P_w is water vapor partial pressure and P_w^* is the saturation vapor pressure at the given temperature. The numerator, P_w, is derived from the occultation refractivity and the ECMWF temperature as described in Sections 2.3 and 7.2. The saturation vapor pressure in the denominator is derived from the ECMWF temperature. Since a temperature error causes errors in the numerator and denominator of the same sign, the numerator and denominator errors are positively correlated. The correlation has two significant

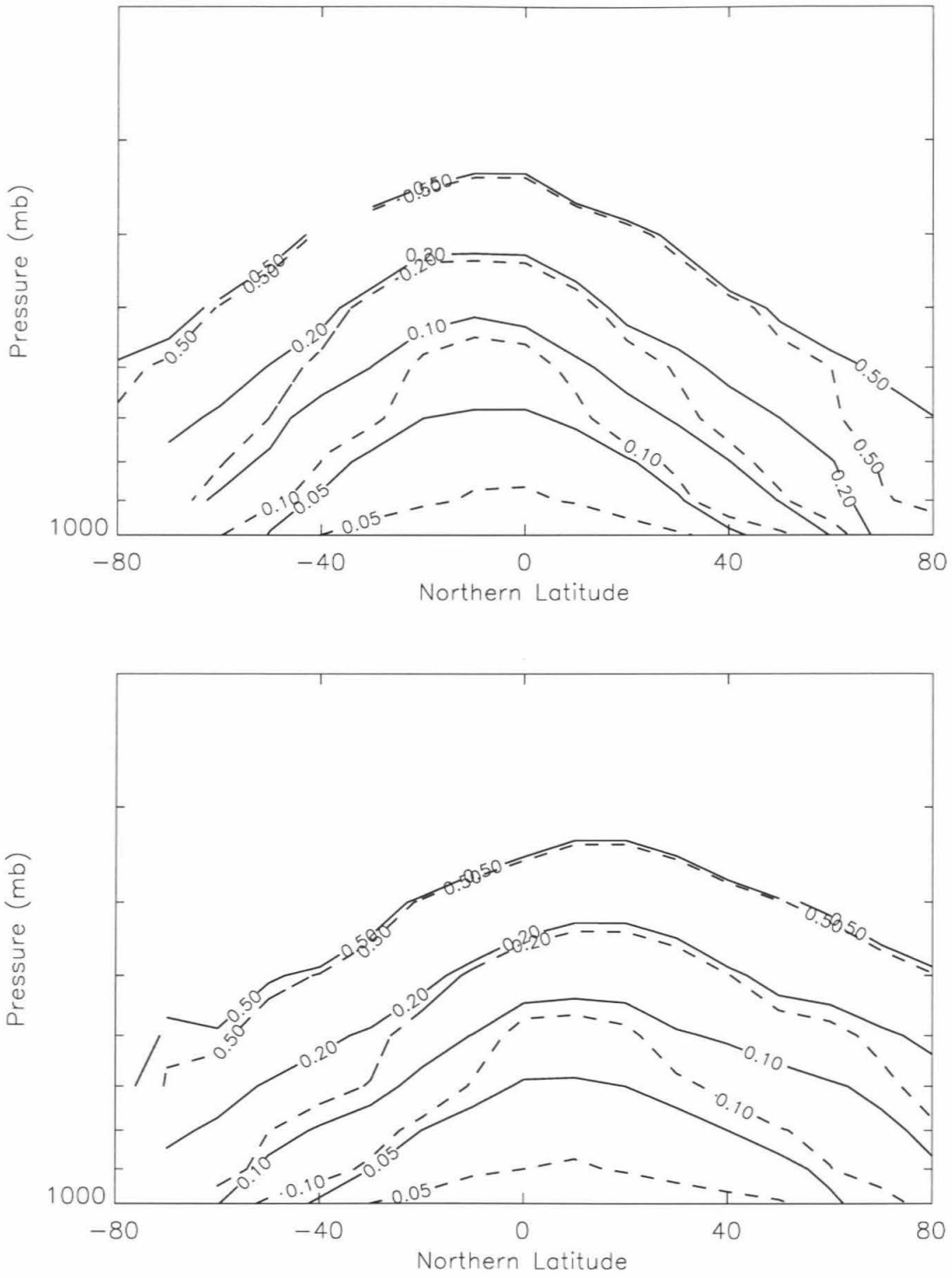

Figure 7.3 Predicted rms fractional error in individual derived water vapor profiles a. DJF and b. JJA.

implications. First, an optimum height exists where the contribution by the temperature error to the numerator and denominator cancels. Second, the derived $\bar{U}$ can be negative in regions where <P_w> and <P_w*> are both positive.

We will now examine the first implication. The second implication is discussed in Section 7.5.8. Consider the terms contributing to $\bar{U}$ estimated via eq. (7.3.1). From eq. (4.2.1), the estimate of P_w ($\equiv \tilde{P}_w$) is the true P_w plus error

$$\tilde{P}_w = P_w\left(1 + \frac{\delta P_w}{P_w}\right) \cong P_w\left(1 + B\left[\frac{\delta N}{N} - \frac{\delta P}{P} + \frac{\delta T}{T}\right]\right) \tag{7.3.2}$$

where we have assumed that B >> 1 in the regions of interest to simplify eq. (7.3.2). δN is the error in the retrieved refractivity, δT is the error in the ECMWF temperature and δP is the error in the hydrostatic pressure estimate. At the altitude where the water vapor retrieval begins, δP is due entirely to refractivity error at higher altitudes. At lower altitudes, δP accumulates error primarily from errors in the ECMWF temperature.

Now to express the saturation vapor pressure, P_w*, we will write the ECMWF temperature as the true temperature, T, plus an error, δT.

$$\tilde{T} = T + \delta T$$

From the Clausius-Clapeyron equation,

$$P_w^* \;\alpha\; \exp\left[-\frac{0.622\,L}{R_d\,T}\right] \tag{7.3.3}$$

Therefore, the estimated saturation vapor pressure is

$$\tilde{P}_w^* = c\exp\left[-\frac{0.622\,L}{R_d\,\tilde{T}}\right] = c\exp\left[-\frac{0.622\,L}{R_d\,T\left(1 + \frac{\delta T}{T}\right)}\right]$$

$$= c \exp\left[-\frac{0.622\,L}{R_d T}\left(1 - \frac{\delta T}{T} + \frac{\delta T^2}{T^2} - \ldots\right)\right]$$

$$= c \exp\left[-\frac{0.622\,L}{R_d T}\right] \exp\left[+\frac{0.622\,L\,\delta T}{R_d T^2}\left(1 - \frac{\delta T}{T} + \ldots\right)\right]$$

$$\tilde{P}_w^* = P_w^* \exp\left[C_1 \frac{\delta T}{T}\left(1 - \frac{\delta T}{T} + \ldots\right)\right] \qquad (7.3.4)$$

where $\tilde{P}_{w0}^*$ is the saturation vapor pressure at the mean of the estimated temperatures in the region, $\bar{\tilde{T}}$ and $C_1 = \frac{0.622\,L}{R_d T}$. Therefore, $\bar{U}$ can be written as

$$\bar{\tilde{U}} = \overline{\frac{P_w}{P_w^*}\left[1 + B\left(\frac{\delta N}{N} - \frac{\delta P}{P} + \frac{\delta T}{T}\right)\right] \exp\left[-C_1 \frac{\delta T}{T}\left(1 - \frac{\delta T}{T} + \ldots\right)\right]} \qquad (7.3.5)$$

Expanding the exponential yields

$$\bar{\tilde{U}} = \overline{\frac{P_w}{P_w^*}\left[1 + B\left(\frac{\delta N}{N} - \frac{\delta P}{P} + \frac{\delta T}{T}\right)\right]\left[1 - C_1\frac{\delta T}{T} + C_1\left(1 + \frac{C_1}{2}\right)\frac{\delta T^2}{T^2} + \ldots.\right]}$$

Now separating the terms to isolate the various contributions yields

$$\bar{\tilde{U}} = \overline{\frac{P_w}{P_w^*}} + \overline{\frac{P_w}{P_w^*}B\left(\frac{\delta N}{N} - \frac{\delta P}{P} + \frac{\delta T}{T}\right)} + \overline{\frac{P_w}{P_w^*}\left[-C_1\frac{\delta T}{T} + C_1\left(1 + \frac{C_1}{2}\right)\frac{\delta T^2}{T^2} + \ldots.\right]}$$

$$+ \overline{\frac{P_w}{P_w^*}\left[B\left(\frac{\delta N}{N} - \frac{\delta P}{P} + \frac{\delta T}{T}\right)\right]\left[-C_1\frac{\delta T}{T} + C_1\left(1 + \frac{C_1}{2}\right)\frac{\delta T^2}{T^2} + \ldots.\right]} \qquad (7.3.6)$$

Temperature error contributes through both the δT and δP terms. Retaining only the first order temperature error related terms in eq. (7.3.6) yields

$$\overline{\delta U_{\delta T}} = \overline{\frac{P_w}{P_w^*}\left[\frac{\delta T}{T}\left(B - \frac{0.622\,L}{R_d\,T}\right) - B\frac{\delta P}{P}\right]} \qquad (7.3.7)$$

The temperature induced part of the pressure error depends on the average temperature error between the altitude at which use of the ECMWF temperature began and the present altitude. A positive temperature bias will cause the pressure scale height to be too large and the pressure will not increase as fast with decreasing height as it should causing pressure to be underestimated resulting in an error defined as

$$\frac{\delta P_T}{P} = -\left\langle\frac{\delta T}{T}\right\rangle\left(\frac{z_0 - z}{H}\right)$$

where $<\delta T/T>$ is the vertically averaged temperature error and δP_T refers to the temperature contribution to the pressure error as opposed to a boundary condition error. The contribution of the temperature error and pressure error to a water vapor error therefore contribute with the same sign in eq. (7.3.7). In general, however, the temperature error will have a vertical correlation and eq. (7.3.7) becomes

$$\overline{\delta U_{\delta T}} = \overline{\frac{P_w}{P_w^*}\frac{\delta T}{T}\left[B\left(1 + c\,\frac{(z_0 - z)}{H}\right) - \frac{0.622\,L}{R_d\,T}\right]} \qquad (7.3.8)$$

where c is a factor between -1 and 1 representing the vertical correlation of the temperature error. In the simplest case where the temperature error has no vertical correlation and $<\delta T/T>$ is 0, eq. (7.3.7) becomes

$$\overline{\delta U_{\delta T}} = \overline{\frac{P_w}{P_w^*}\frac{\delta T}{T}\left(B - \frac{0.622\,L}{R_d\,T}\right)} \qquad (7.3.9)$$

However, the temperature error will often be quite correlated vertically representing conditions where the atmospheric stability in the model analysis is correct but the entire temperature profile is biased and so c must be considered.

C_1 is of order 20 and changes by only 30% between 230K and 300 K whereas B changes by orders of magnitude across the troposphere. Therefore, errors in relative humidity due to temperature errors at higher altitudes are dominated by errors in the numerator (derived water vapor) whereas at lower altitudes the error in the denominator (saturation vapor pressure) dominates. The transition where the temperature error contribution to relative humidity is zero in eq. (7.3.8) occurs where

$$B = \frac{0.622\,L}{R_d\,T}\left(1 + c\,\frac{(z_0 - z)}{H}\right)^{-1}$$

or equivalently

$$q_0 = \frac{b_1\,R_d\,T^2}{b_2\,L}\left(1 + c\,\frac{(z_0 - z)}{H}\right) \qquad (7.3.8)$$

If c = 0, the resulting q_0 is ~1.3 to 1.5 g/kg. The 1.4 g/kg contour overlaps ECMWF temperature contours ranging from 260 K to 275 K. This contour is crudely about a scale height below z_0 such that if c = 1, q_0 is ~2.8 g/kg. Therefore, somewhere between the q = 1.4 and 2.8 g/kg contours, the contribution of ECMWF temperature error to relative humidity is zero and any error is due almost entirely to the occultation refractivity data. The peak height of the 2 g/kg contour occurs at the Inter-Tropical Convergence Zone (ITCZ), near 6.5 km altitude. At 80N the 2 g/kg contour is at 2 km whereas it intersects the surface between 60S and 65S. For a 1.5 K temperature error, a 10% error in relative humidity occurs near the B=40 contour. At altitudes below this contour, the temperature error contribution to relative humidity error decreases with height to the ~2 g/kg contour. At still lower altitudes, the error changes sign and grows in magnitude due to the error in saturation vapor pressure reaching a maximum value at the surface.

7.4 Individual Profile Examples

Several interesting examples of individual profile comparisons with ECMWF and radiosondes will now be presented. These profiles originally were presented at the Global Energy and Water Cycle Experiment (GEWEX) conference at the National Academy of Sciences in June 1996 [Kursinski et al., 1996b]. The examples were selected to represent a variety of climatological conditions and interesting implications.

Helsinki (63N 36.5E, 6/22/95 03:49UT)

Figure 7.4 shows an occultation which occurred over northwestern Russia to the northeast of Helsinki. The moisture content of the atmosphere is quite large, consistent with the near-continuous solar insolation at this latitude during the June-July period. The profiles of three nearby radiosondes are shown for comparison along with the interpolated ECMWF 6 hour global analysis. The ECMWF humidity profile is quite smooth except below 1.5 km altitude where it suddenly becomes iso-humid. The two radiosondes to the east are smooth while the profile from the radiosonde southwest of the occultation includes a sharp water vapor structure between 4 and 6 km altitude and is significantly drier than the other two below 6 km altitude. The occultation is closest to this last radiosonde and also includes a structure near 5 km altitude similar to but smaller in magnitude than that of the third radiosonde. Since the ECMWF temperature profile displays no hint of any structure near 5 km, the sharp vertical structure in the GEW humidity profile is due entirely to the humidity. The GEW q profile looks like a combination of the three radiosonde profiles.

Northern Germany (52.0N 11.8E, 6/22/95 12:19UT)

The occultation occurs over northern Germany in the vicinity of several radiosondes although the closest ones are 160 km away (figure 7.5). The distribution of sonde humidities indicates the variability of water vapor over regions several hundred kilometer square. Above 3 km, specific humidity varies by a factor of 5. Closer to the surface, it

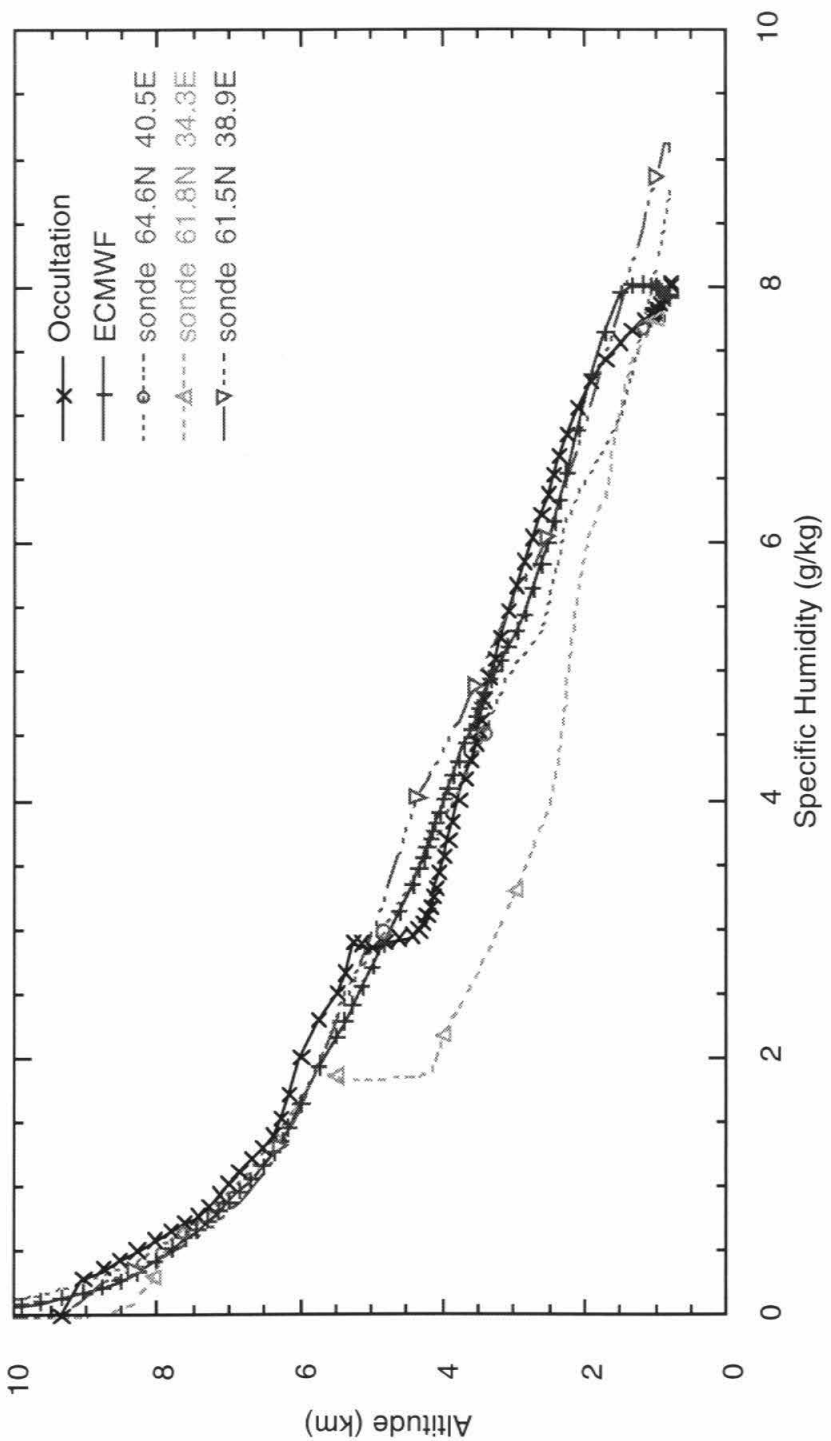

Figure 7.4 GEW specific humidity profile northeast of Helsinki at 63N, 36.5E on June 22, 1995 at 03:49 UT compared with ECMWF 6 hour global analysis interpolated to the occultation location and three nearby radiosondes

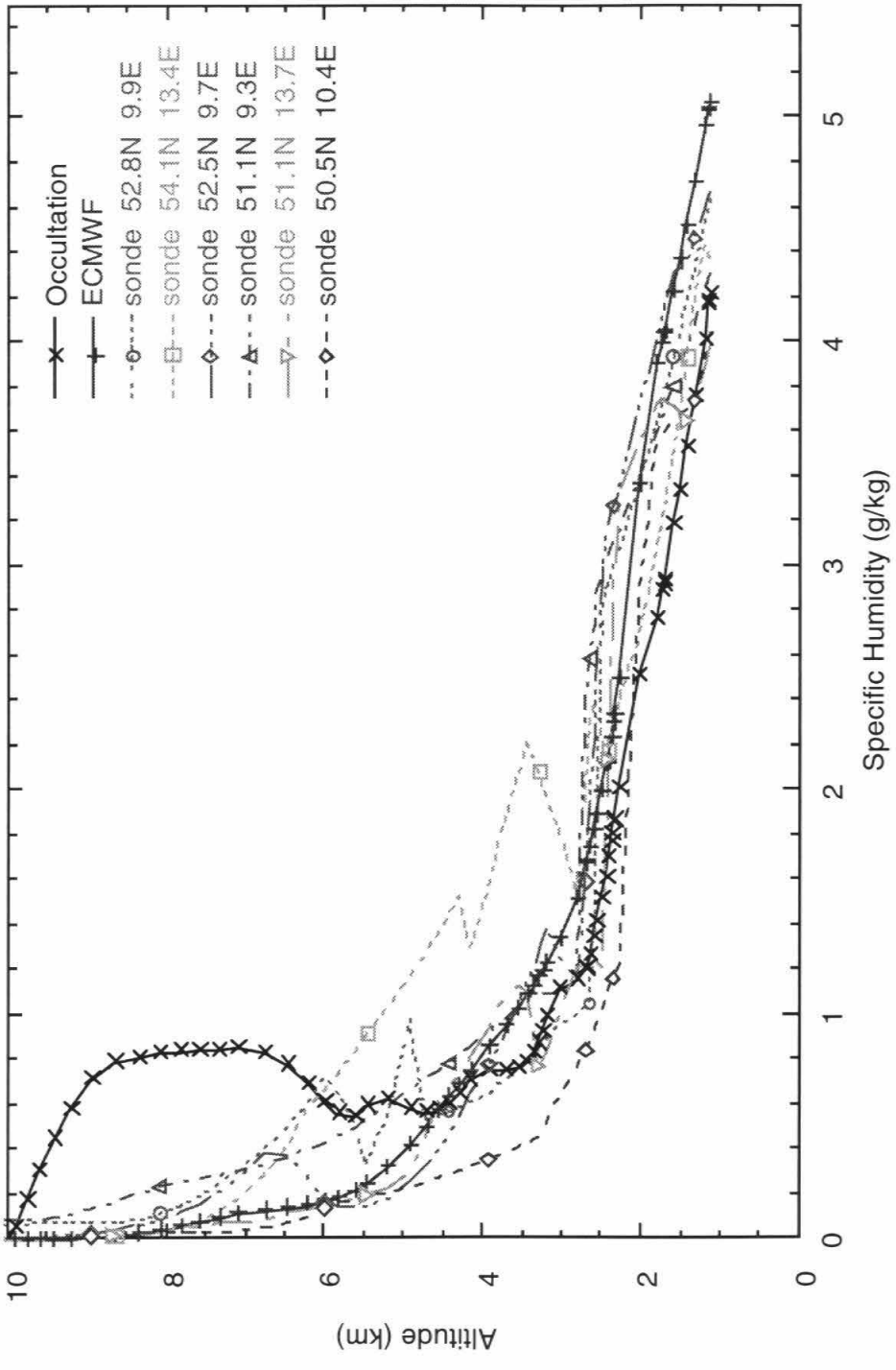

Figure 7.5 Comparison of GEW specific humidity profile over northern Germany at 52N, 11.8E on June 22 1995 at 12:19 UT with several radiosondes and the ECMWF 6 hour global analysis interpolated to the occultation location

varies by about 50%. The relationship between the GEW and sonde specific humidities varies with height. The GEW specific humidity shifts from the dry end of the sonde humidities at low altitudes to near the mean at middle altitudes to decidedly supersaturated above 6km altitude. Again the ECMWF analysis humidity is quite smooth. At most altitudes, the analysis humidity lies near the center of the radiosonde distribution which is not surprising since it assimilated these data. It is a bit surprising that the ECMWF analysis is moister than the radiosonde and GEW profiles near 1 km altitude. An overestimate of the GEW q by 0.5 g/kg can be caused by a refractivity bias of order 1.5% as well as an ECMWF temperature error of 4 K. Although 4K appears large, the six radiosonde temperature profiles vary by ~8 K near 8 km altitude so a temperature error may indeed contribute a significant fraction of the GEW moisture overestimate.

Falkland Islands (49.2S 303.2E, 6/23/95 04:07UT)

Figure 7.6 shows an occultation to the northeast of the Falkland islands along with an interpolated ECMWF profile and two radiosonde profiles. The two radiosondes were launched from a common location on the Falklands, roughly two hours before and four hours after the occultation. The occultation lies roughly 300 km to the northeast of the radiosonde location. Somewhat surprisingly, the ECMWF and occultation humidity profiles are far more similar than either of the radiosonde profiles showing a relative maximum in q near 600 mb whereas the radiosonde profiles exhibit relatively simple exponentially decay. I provided this figure to Anthony Hollingsworth, director of research at ECMWF, along with the question as to how ECMWF was able to discern the relative maximum in the mid troposphere. After consultation with his ECMWF satellite "honchos", he indicated that the occultation fell within an occluded front, a front which has lifted off the surface. The ECMWF analysis was able to identify the elevated relative maximum in humidity through assimilation of low vertical resolution humidity information from the TIROS Operational Vertical Sounder (TOVS). The ECMWF reconstruction was verified

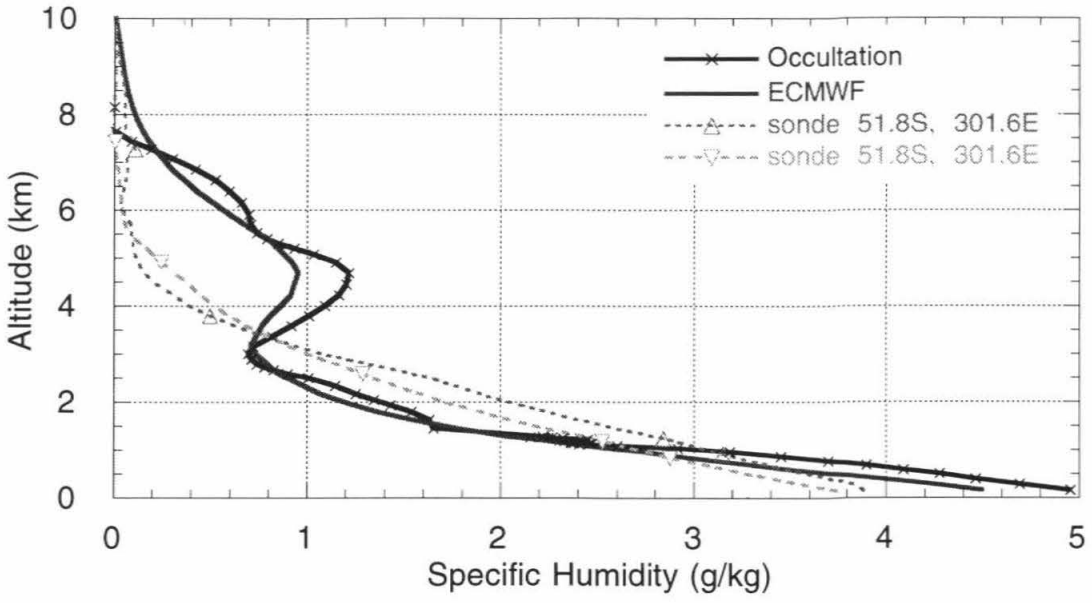

Figure 7.6 Occultation northeast of the Falkland Islands at 49.2S, 303.2E on June 23, 1995 at 04:07 UT compared with the interpolated 6 hour ECMWF global analysis and the Falkland Islands radiosonde launched at 00 UT and 06 UT on the same day.

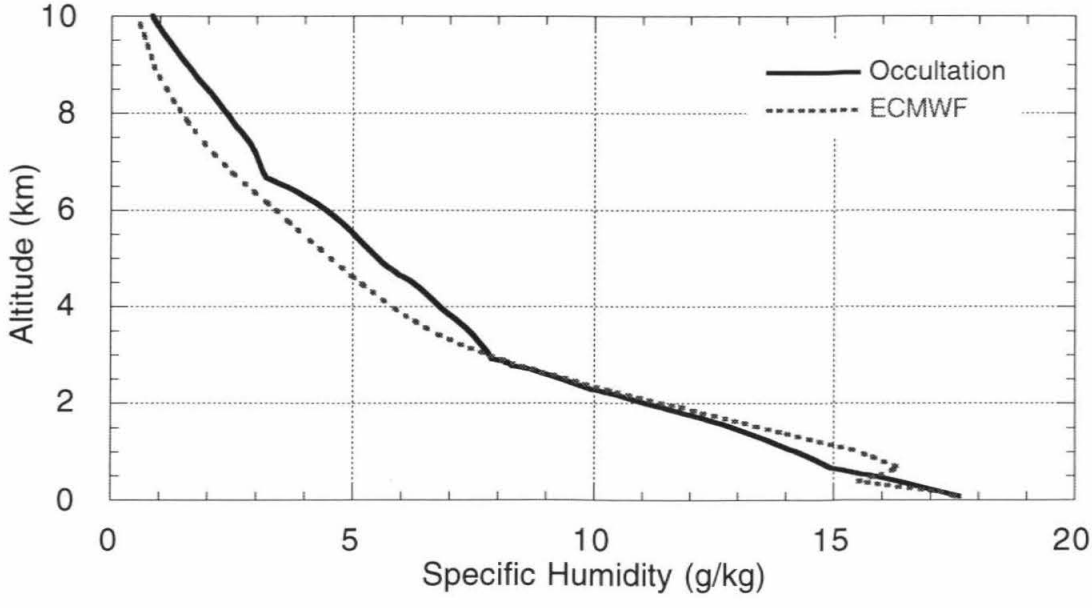

Figure 7.7 Occultation southeast of Sri Lanka at 7.1S, 88.8E on June 27, 1995 at 05:31 UT compared with the interpolated ECMWF 6 hour global analysis.

further by estimation of and comparison with the METOP IR humidity sounding channel. The agreement between the ECMWF simulated and the actual METOP image of the region was quite good. According to the analysis, the radiosondes sampled a distinctly different air mass outside the frontal region, consistent with the results of figure 7.6. The agreement between ECMWF and GEW humidity profiles and simultaneous disagreement with the two sondes is strong testament to both the accuracy and resolution of the occultation data and to the utility of assimilating TOVS data.

Indonesia (7.1S 88.8E, 6/27/95 05:31UT)

The occultation profile in figure 7.7 demonstrates the ability of occultations to retrieve water vapor structure to the surface in the presence of large specific humidities. The derived water vapor profile also exhibits cusps near ~6.5 and ~3 km altitude which have appeared in some of our retrievals near the surface. The cusps do not appear to be atmospheric in origin but rather an artifact in the retrievals. They may be associated with multipath or diffraction or an incorrect Fresnel smoothing interval and clearly require more investigation. They also may be associated with interpolation of the ECMWF temperature and moisture profiles. Using the amplitude of the occulted signals would provide an additional constraint here because the defocusing effect scales as the vertical second derivative of refractivity. The large second derivative of the cusps, if real, implies that the signal level should be dramatically defocused at the height of the cusps along with a diffraction effect associated with the short vertical extent of the cusps. I do not believe that such amplitude fluctuations are present in the occultation data indicating the cusps are not real.

Madagascar (18.7S 50.8E, 06/27/95 08:55UT)

The occultation in figure 7.8 occurs just off the east coast of Madagascar. The refractivity of the occultation profile is significantly lower in the bottom 4 km than that of

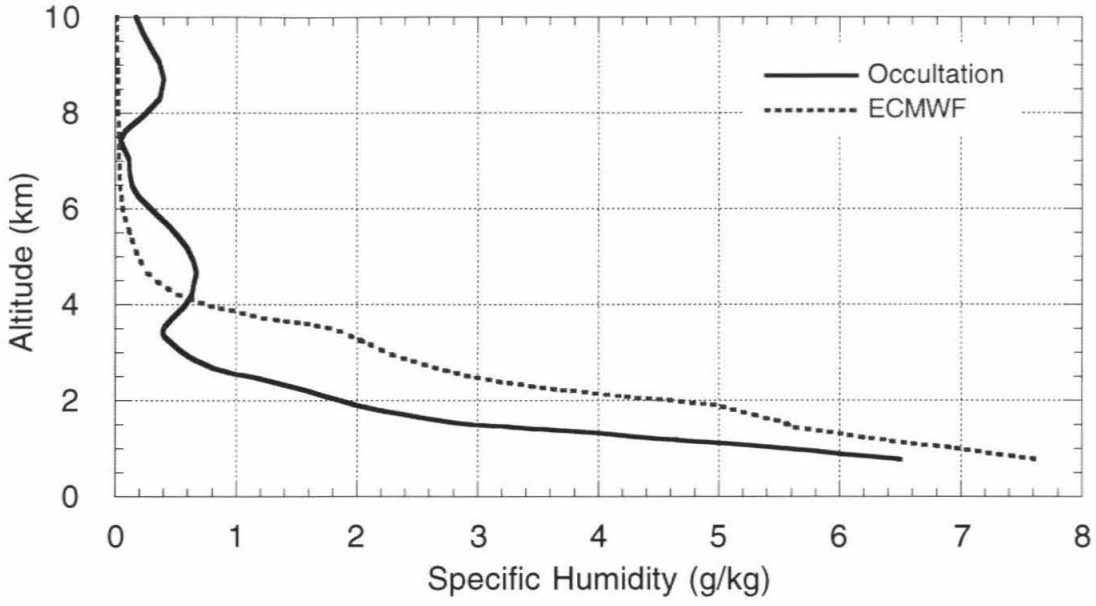

Figure 7.8 Occultation off the east coast of Madagascar at 18.7S, 50.8E on June 27, 1995 at 08:55 UT compared with interpolated ECMWF 6 hour global analysis.

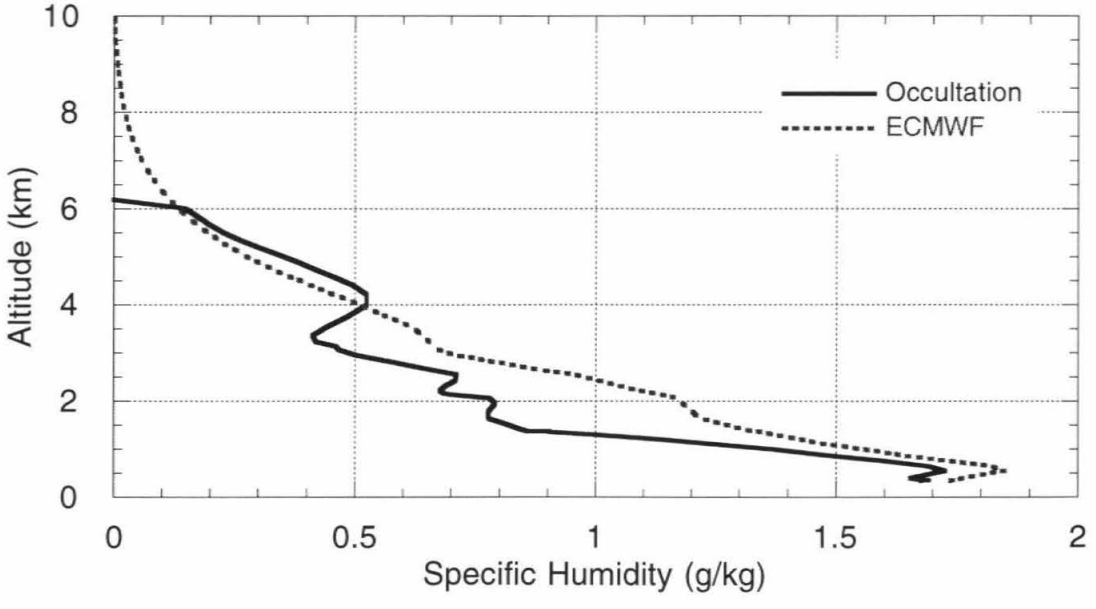

Figure 7.9 Occultation east of the Weddell Sea off the coast of Antarctica at 68.6S, 19.4W on June 23, 1995 at 00:42 UT compared with interpolated ECMWF 6 hour analysis.

ECMWF. Since it is very difficult to misplace the altitude of the high refractivity gradient in the occultation retrieval process, it is likely that the discrepancy is largely due to the ECMWF analysis. The large discrepancy is in fact representative of a systematic problem with the ECMWF analysis in the latitude belt of 10S to 20S associated with the southern descending branch of the Hadley cell during this time of year. Simply put, ECMWF has placed the altitude of the tradewind inversion too high, an issue which is discussed a bit more in Section 7.5.5. I also provided figure 7.8 to A. Hollingsworth who later indicated, again after consultation with his satellite gurus, that had the ECMWF analysis assimilated TOVS data which at the time of the analysis had only recently become available over this region, the resulting humidity profile would have been far more similar to that of the GEW profile. This conclusion again indicates the accuracy the GEW results and the importance of vertical resolution particularly near the surface.

Antarctica (69.6S 10.4W, 06/23/95 00:42UT)

Figure 7.9 provides an example of high latitude winter soundings. The 230 K altitude is ~6 km. The maximum q is only ~1.8 g/kg, roughly a tenth of that in figure 7.7. The minimum B is ~20 at which the humidity error due to a 1.5 K temperature error is 10% while a 1% error in retrieved refractivity will contribute about 20%. Errors in retrieved refractivity error should generally be smaller than 1% because synoptic weather patterns which dominate horizontal refractivity variations under these conditions are large scale structures and the 1% refractivity error in figure 3.8 is due to smaller scale humidity variations found in summer conditions. The expected error in water vapor is then ~20% near the surface growing rapidly with height with H_B ~5.5 km in the lowest 3 km. The consistency between the GEW and ECMWF humidity profiles up to 6 km is therefore somewhat unusual for high latitude winter conditions.

7.5 Latitude Versus Height Water Vapor Climatology for June 21 - July 4, 1995

Here, we discuss the latitude versus height behavior of water vapor derived from the occultation refractivity structure using temperature from the ECMWF 6 hour global analyses, and compare the results with the humidity structure from the ECMWF analyses for the same period and P&O climatology for June, July and August (JJA). In the process, we will work sequentially through refractivity, temperature, pressure and water vapor analogous to the actual processing, examining the behavior of each before moving on. Water vapor is discussed in terms of specific humidity, relative humidity and equivalent potential temperature. Precipitable water content is not estimated because the retrievals generally do not extend to the surface. Implications on accuracy will be described in the process.

7.5.1 Comments Concerning the GEW, ECMWF and P&O Climatologies

A total of 660 occultations have been combined into the GEW climatology to compare the occultation retrievals with other data sets and characterize the state of the atmosphere for the period of June 21-23 and June 27 - July 4, 1995, (unfortunately due to instrumental problems, data from June 24-26 is not used). Their distribution versus height and latitude is shown in figure 7.10a. The total number of points within these occultations is shown in figure 7.10b. The general increase in the number of points at lower altitudes is because resolution increases there (figure 2.7b). The distribution of occultations is controlled by the orbital planes of the GPS receiver and the 24 GPS transmitters, the distribution of ground receivers used to difference out clock error and the memory storage and downlink rate of the receiver's spacecraft.

The averaging inherent in a climatology reduces random errors by the square root of the number of independent points. A factor of 3 is a reasonable guess as to the

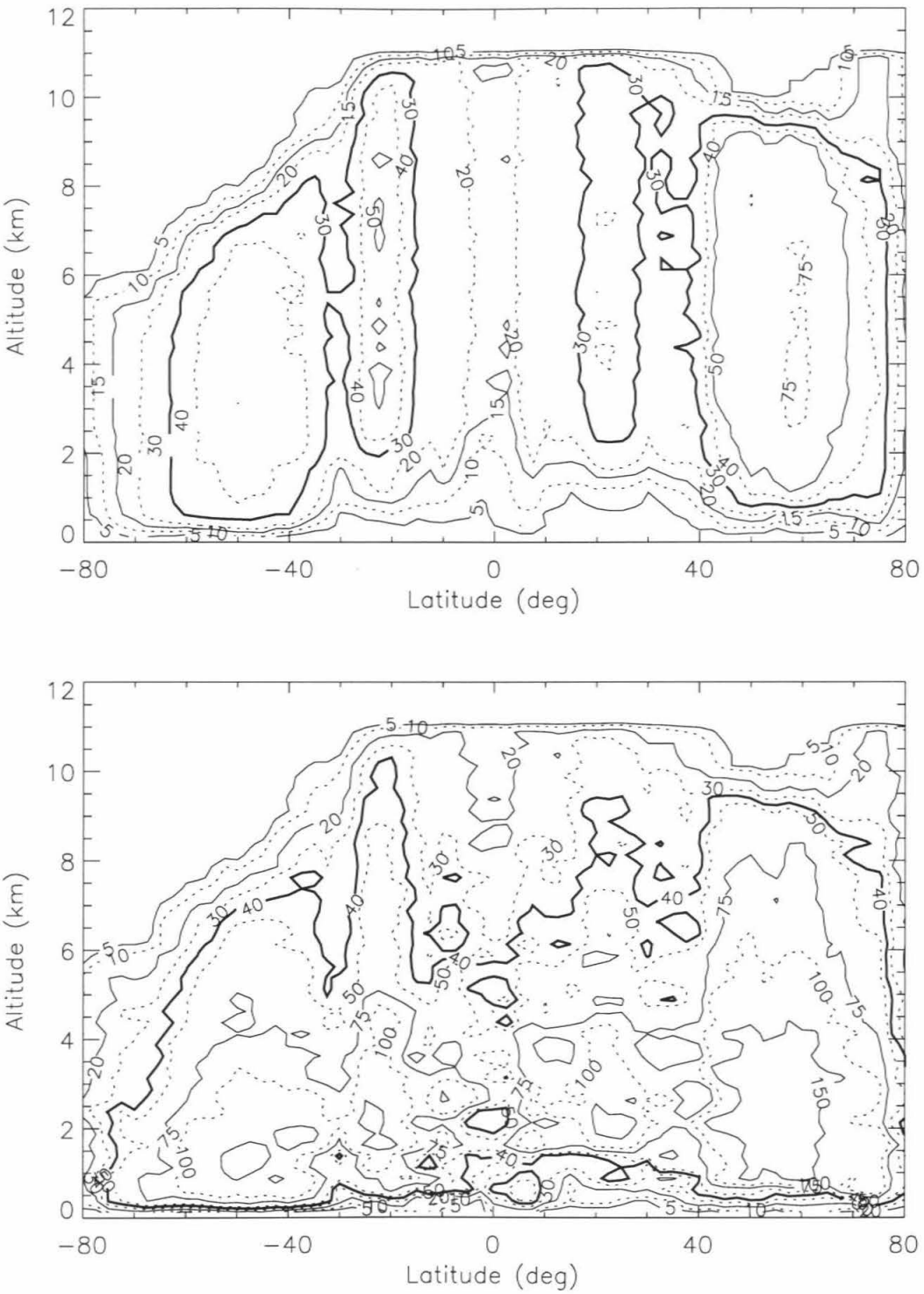

Figure 7.10 a. Number of occultation profiles per latitude bin ~ number of independent points per bin. b. Number of occultation points per latitude bin. Increase at lower altitudes is due to increase in resolution there (Figure 2.7b).

improvement in predicted accuracy of averaged versus individual profile results (Sections 4 and 7.3). In the GEW climatology, latitude bins are 10° wide separated every 5° such that every occultation profile is counted twice. Averaging the 20 to 75 occultations per latitude bin will decrease random errors in refractivity and ECMWF temperature by factors of 4 to 9 sufficient to make an accurate assessment of biases.

The ground truth against which we will compare the GEW results has its own set of the features and limitations which deserve comment. The P&O climatology is based on a network of 1093 rawinsonde stations for the ten-year period from May 1963 through April 1973 supplemented by daily surface ship reports. It suffers from the spatial biasing of the radiosonde network which is focused over continents and particularly poor in the southern hemisphere. Further, radiosonde humidities in the middle to upper troposphere are suspect due to low measurement sensitivity there. Relative humidity measurements below 20% are, in general, questionable and, in fact, some radiosonde climatological data bases simply truncate relative humidity to a minimum value of 20%.

Humidity in the ECMWF analyses is derived from available radiosonde data which is smoothed before assimilation (A. Hollingsworth, personal comment) and the TIROS Operational Vertical Sounder (TOVS) humidity channels. The TOVS IR data is in three channels sampling roughly within the 6.3 μm water absorption band. These channels, located at 6.7 μm, 7.3 μm and 8.3 μm, are sensitive to the amount of water vapor integrated over broad vertical layers centered in the upper (200-500 mb), middle (300-700 mb) and lower troposphere (600 mb - surface). ECMWF data is therefore global but suffers in remote regions from low vertical resolution and IR sounding problems associated with clouds.

We will use two versions of the ECMWF data. The first is the complete ECMWF climatology (CEC) using all the six hour global analyses available for the June 21-23 and June 27 - July 4, 1995, period. The second is a subset of 660 "profiles" which have been

extracted from the ECMWF analyses by interpolation to each point derived in each occultation profile. This data set is referred to as the interpolated ECMWF climatology (IEC). The IEC and CEC results are compared to assess the adequacy of the GEW as a climatologically valid sample of the latitude versus height structure of the atmosphere over the June 21 - July 4, 1995, period.

7.5.2 Refractivity results

Here we examine the GPS refractivity in terms of mean and standard deviation for the June 21 to July 4, 1995, period. We remove a simple exponential model to reveal the latitude versus height structure of mean atmospheric refractivity. Comparison of GPS and ECMWF refractivities reveals similar structure with a small spatial bias pattern. Variability of the two data sets is similar with GPS data exhibiting slightly larger variability.

Occultation refractivity

Figure 7.11 shows the latitude versus height structure of mean occultation, IEC and CEC refractivity for the period of June 21 - 23 and June 27 - July 4 which exhibits an exponential decrease with altitude due primarily to pressure. Superimposed on the exponential decrease is latitudinal structure which is made more evident in figure 7.12 after subtraction of a simple exponential refractivity model ($N=340\ e^{-z/7km}$) from the mean N structure. The relative maxima near 5N and relative minima near 20S and 30N are signatures of the Hadley circulation. The low latitude maximum is due to water vapor whose importance decreases with altitude such that above 6 km there is relatively little meridional contrast. Relative minima near 25S and 35N are associated with hot dry descending air in the Hadley circulation. At higher latitudes, there is relatively little change in refractivity at constant altitude because decreasing temperatures directly increase density and refractivity but also indirectly decrease pressure and therefore density and refractivity through the pressure scale height. The effect of temperature on the scale height is apparent

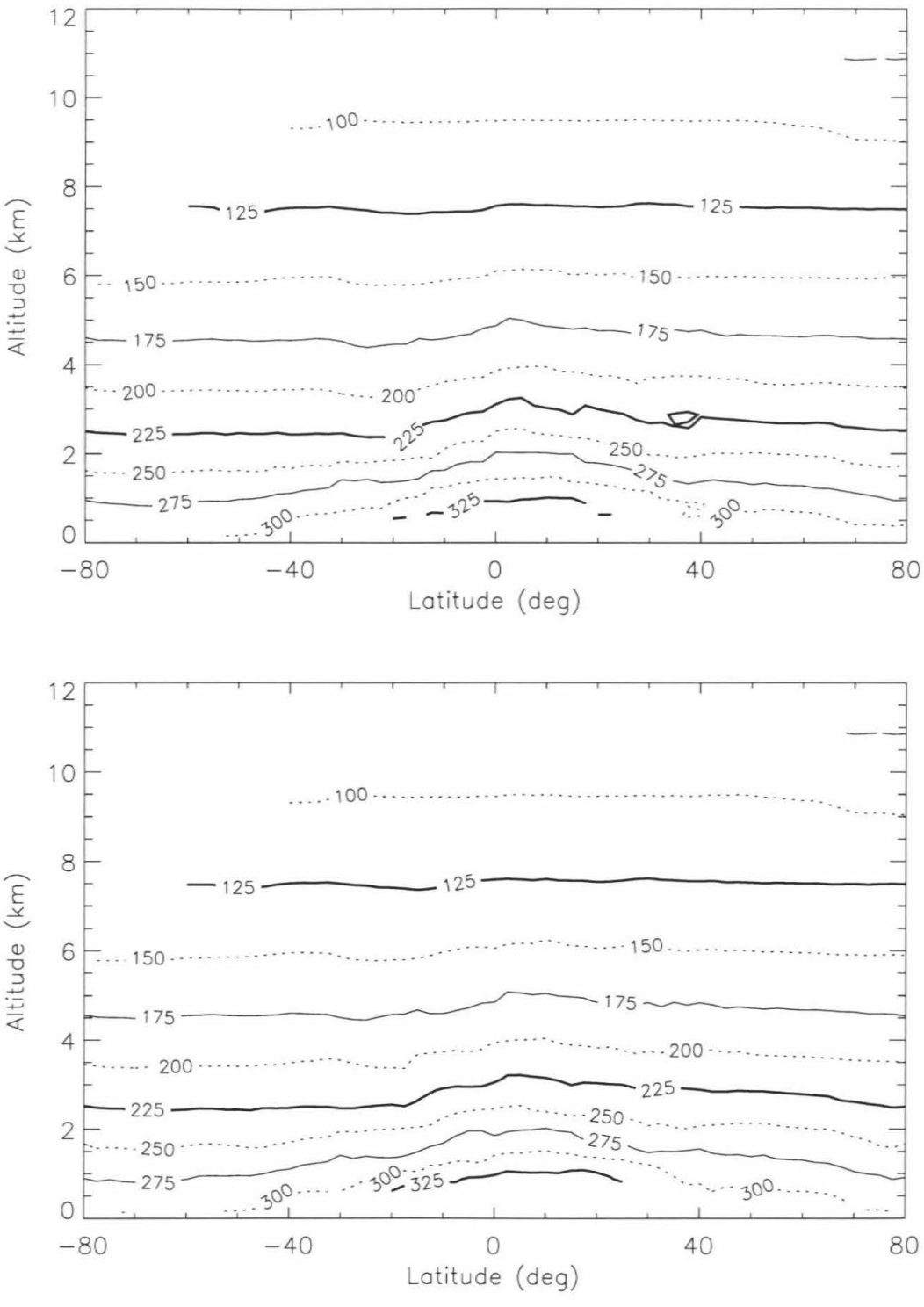

Figure 7.11 Mean refractivity derived from a. occultation data b. IEC.

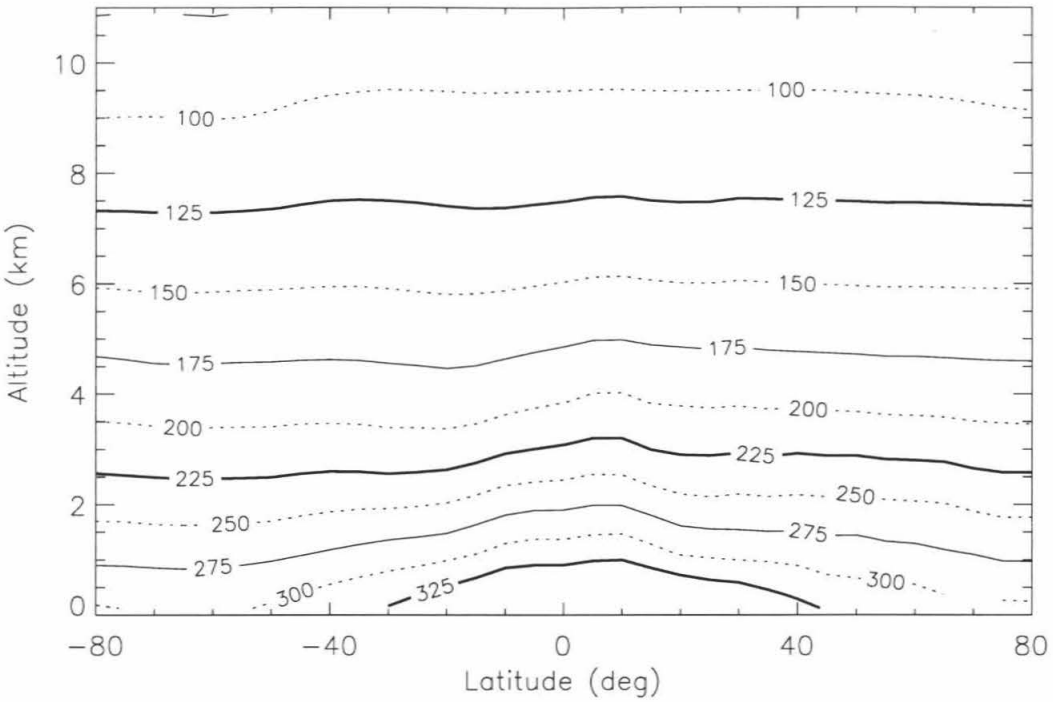

Figure 7.11 c Mean refractivity derived from CEC.

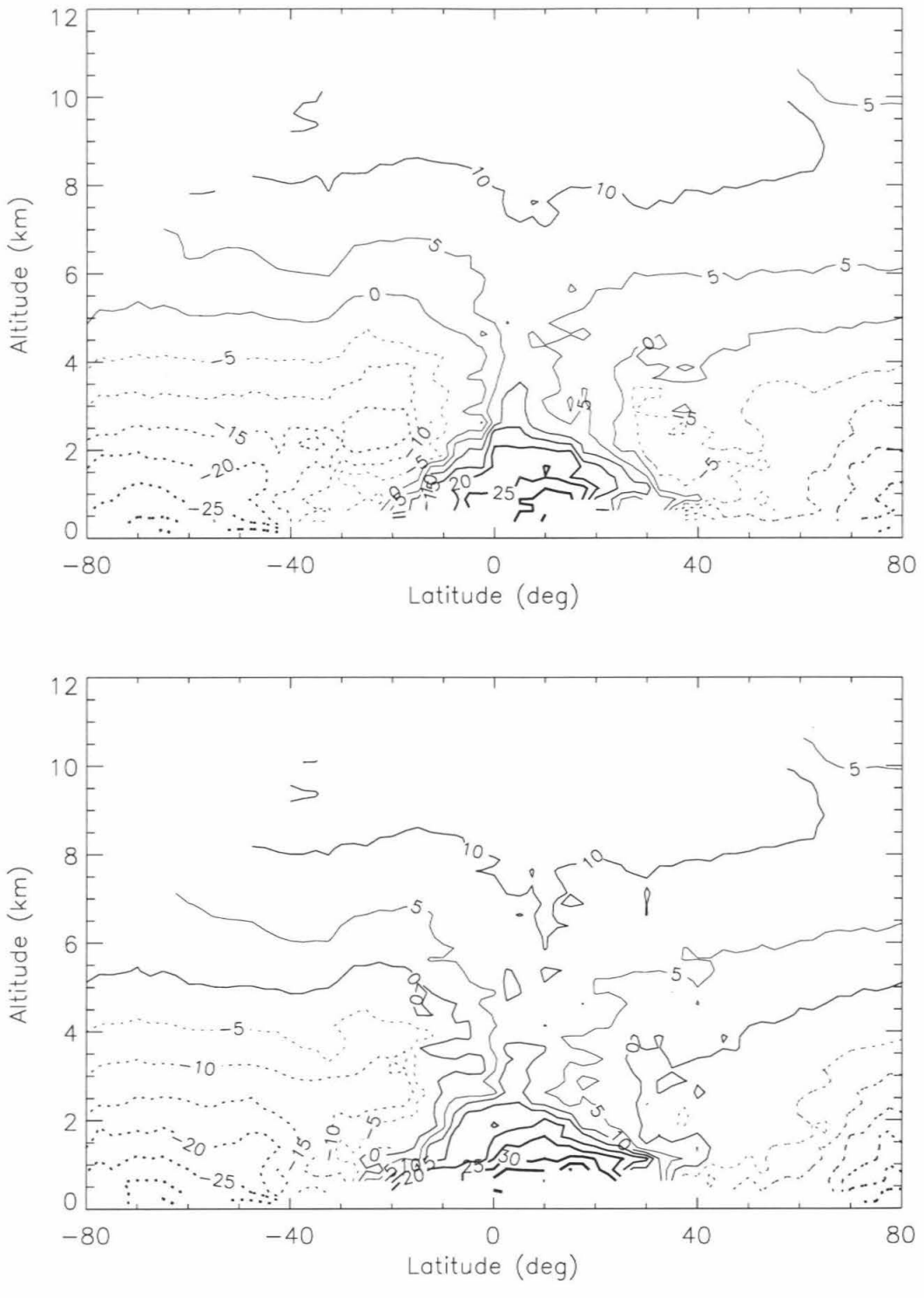

Figure 7.12 Mean refractivity minus exponential refractivity model for a. occultation and b. IEC (N-units).

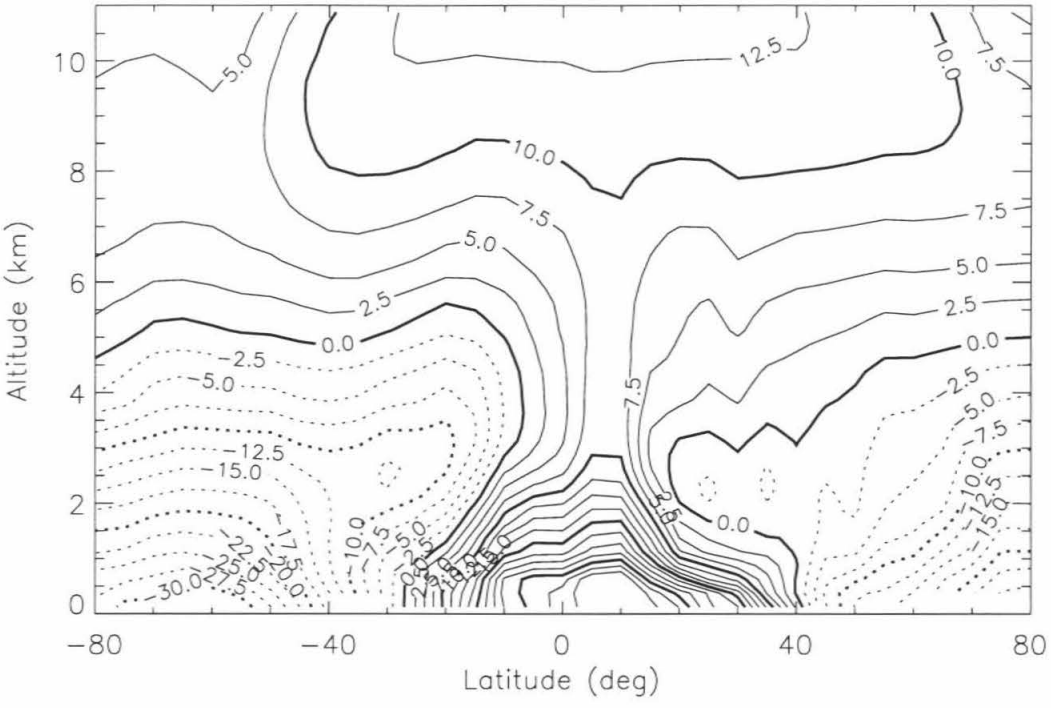

Figure 7.12 c. Mean refractivity minus exponential refractivity model for CEC (N-units).

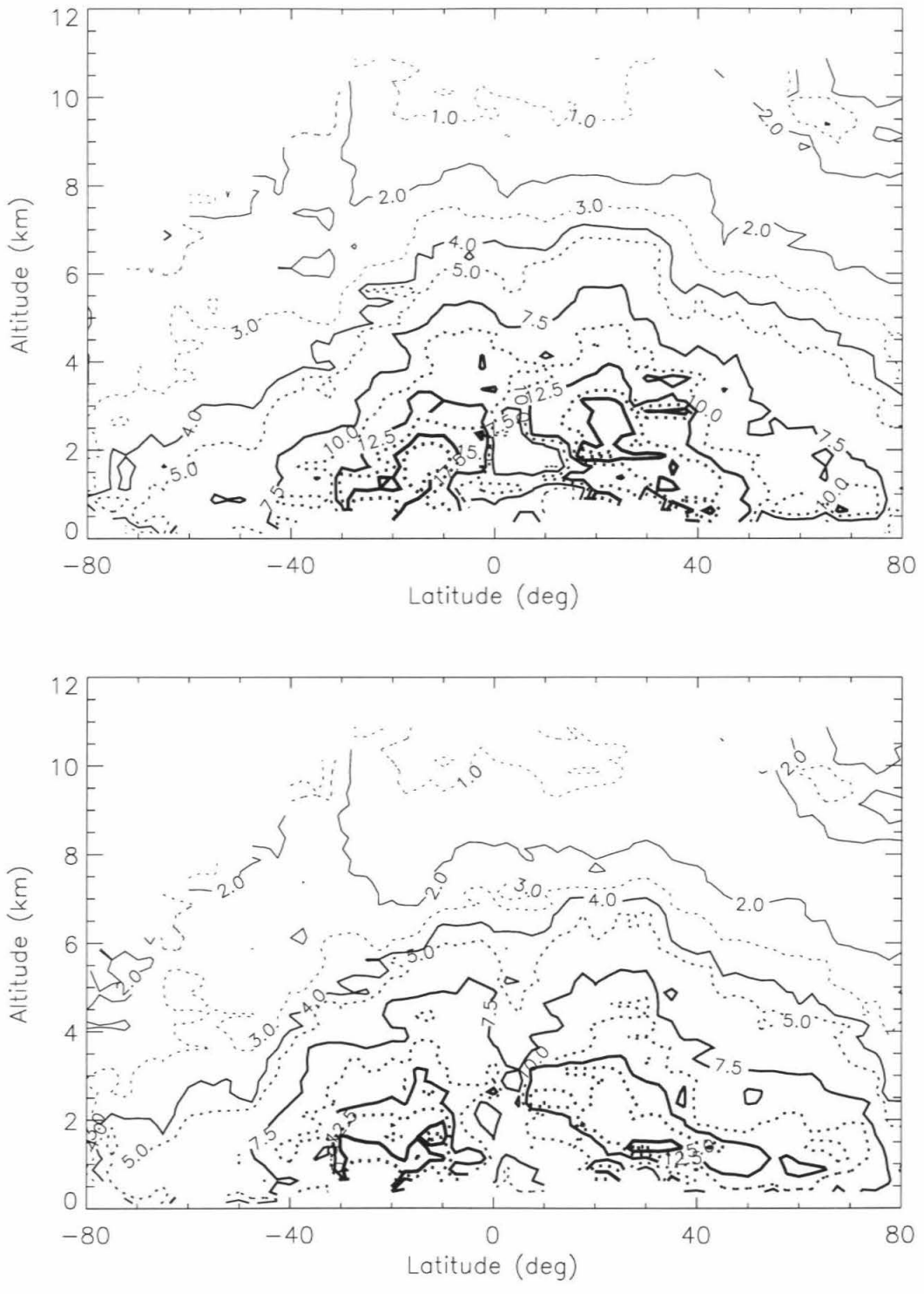

Figure 7.13 Standard deviation of refractivity from a. occultation, b. IEC, and c. CEC.

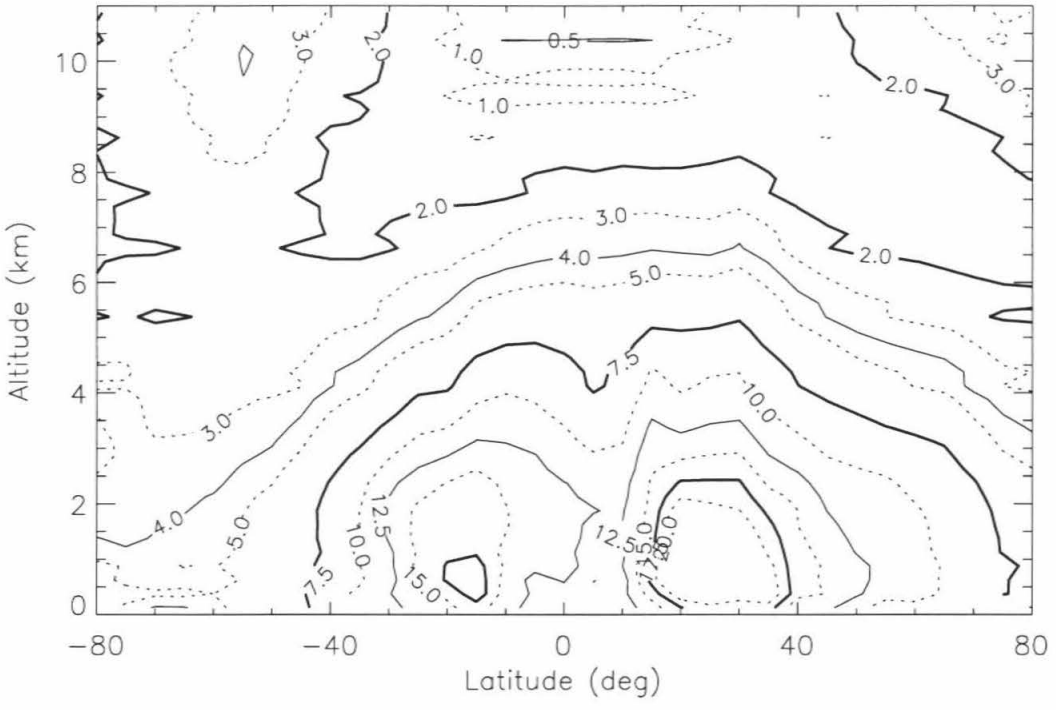

Figure 7.13 c. Standard deviation of refractivity from CEC.

in the more closely spaced contours in the southern versus northern hemisphere. A moisture increase will reduce the refractivity scale height as is apparent below 3 km altitude at low latitudes where contour spacing reaches a minimum. The minimum in surface refractivity occurs in the southern hemisphere due to lack of water.

The standard deviation of occultation, IEC and CEC refractivity (figure 7.13), shows a general increase toward lower altitudes and latitudes associated with the increase in moisture there. The low latitude variability exhibits a bimodal distribution centered near the ITCZ with peaks near 10S and 20N. There is a distinct peak in variability between 45S and 60S below 6 km. Variability is also high near 70S below 3km altitude. There are high altitude regions of high variability south of 30S and north of 30N apparent in both the occultation and ECMWF data sets

Comparison with ECMWF

The bias between occultation and IEC mean refractivities is shown in figure 7.14a and the normalized bias in figure 7.14b. Occultation refractivities are slightly larger than ECMWF values at higher altitudes. The bias is generally smaller in the southern hemisphere. Peak biases are ~8 N-units. There is a positive bias between 825 and 725 mb between the equator and 20N which stands out against a large crescent shaped region of negative bias which extends from 45S to 70N with a top near 350mb. The positive bias region lies at the bottom center of the crescent. There are three regions of negative bias greater in magnitude than 1.5% within the crescent shaped region, one each to the north, south and above the positive bias region. In the southern hemisphere the sign of the bias becomes positive south of 45S with a fractional magnitude typically between 0.2 and 0.5% with the highest values occurring between 60S and 70S.

Figure 7.15 shows the standard deviation of the refractivity difference between the occultation and IEC data. In general, the variability increases toward lower latitudes and

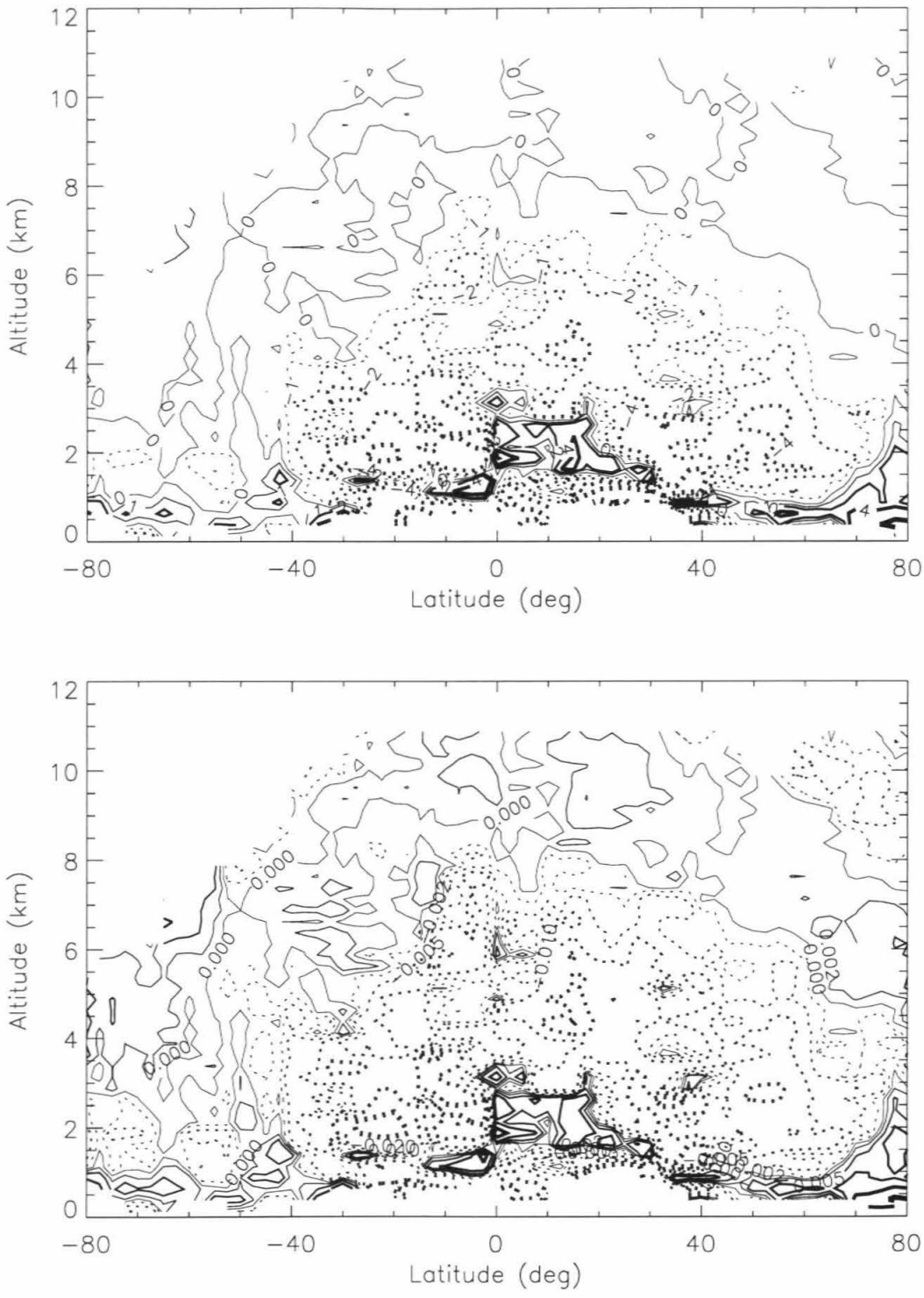

Figure 7.14 Mean of refractivity difference, occultation minus IEC
a. N-units and b. normalized.

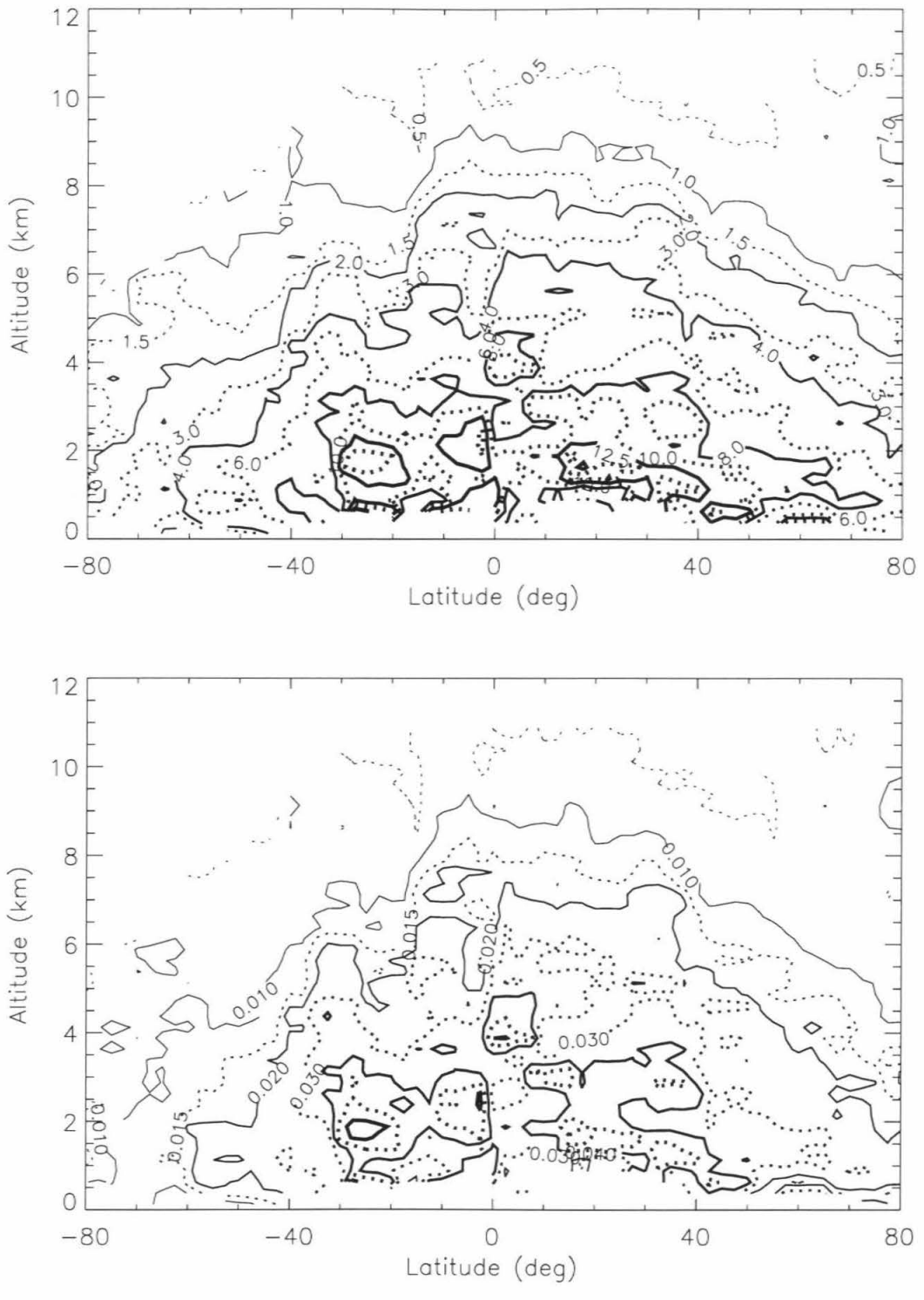

Figure 7.15 a. Standard deviation of refractivity difference, occultation minus IEC (in N-units) b. Normalized standard deviation of refractivity difference, occultation minus IEC.

lower altitudes. Fractional rms variations range from less than 1% at higher altitudes to peak values of 6% on either side of the equator (figure 7.15b). The standard deviation of the refractivity difference between the occultation and ECMWF data contains contributions from the errors in both the occultation and ECMWF data. The difference variations are smaller in general than the standard deviation of the occultation refractivities alone, indicating the structure in figure 7.13 is due to real atmospheric variability. In the future, the vertical correlation of the differences needs to be characterized and compared with the TOVS weighting functions to assess how much of the difference is due to the different vertical resolution between the occultation and TOVS soundings.

7.5.3 ECMWF temperature description

To interpret the refractivity information in terms of moisture, the dry contribution to refractivity in eq. (2.3.1) must be estimated and removed. To do so temperatures from the ECMWF global analyses have been interpolated to the location of each occultation. The nearest analysis within three hours of the occultation times is used. The mean and standard deviation of the interpolated ECMWF temperature structure are discussed below. Some additional discussion of the accuracy of ECMWF implied from the relative humidity results are discussed in Section 7.5.7.

Figure 7.16 shows the mean latitude versus height structure of the IEC and CEC temperatures. The highest temperatures and smallest latitudinal gradients are found between 20S and 30N. The lower temperatures and larger meridional gradients in the southern hemisphere are evidence of the seasonal contrast between the hemispheres. In general, the IEC and CEC mean temperature structures agree. The largest discrepancy occurs over Antarctica where the IEC temperatures are 2 to 5 K warmer than those of CEC. IEC temperatures within 1 km of the surface in the northern hemisphere are somewhat noisy. The IEC 235 K contour is also noisy and biased high. With the exception of these

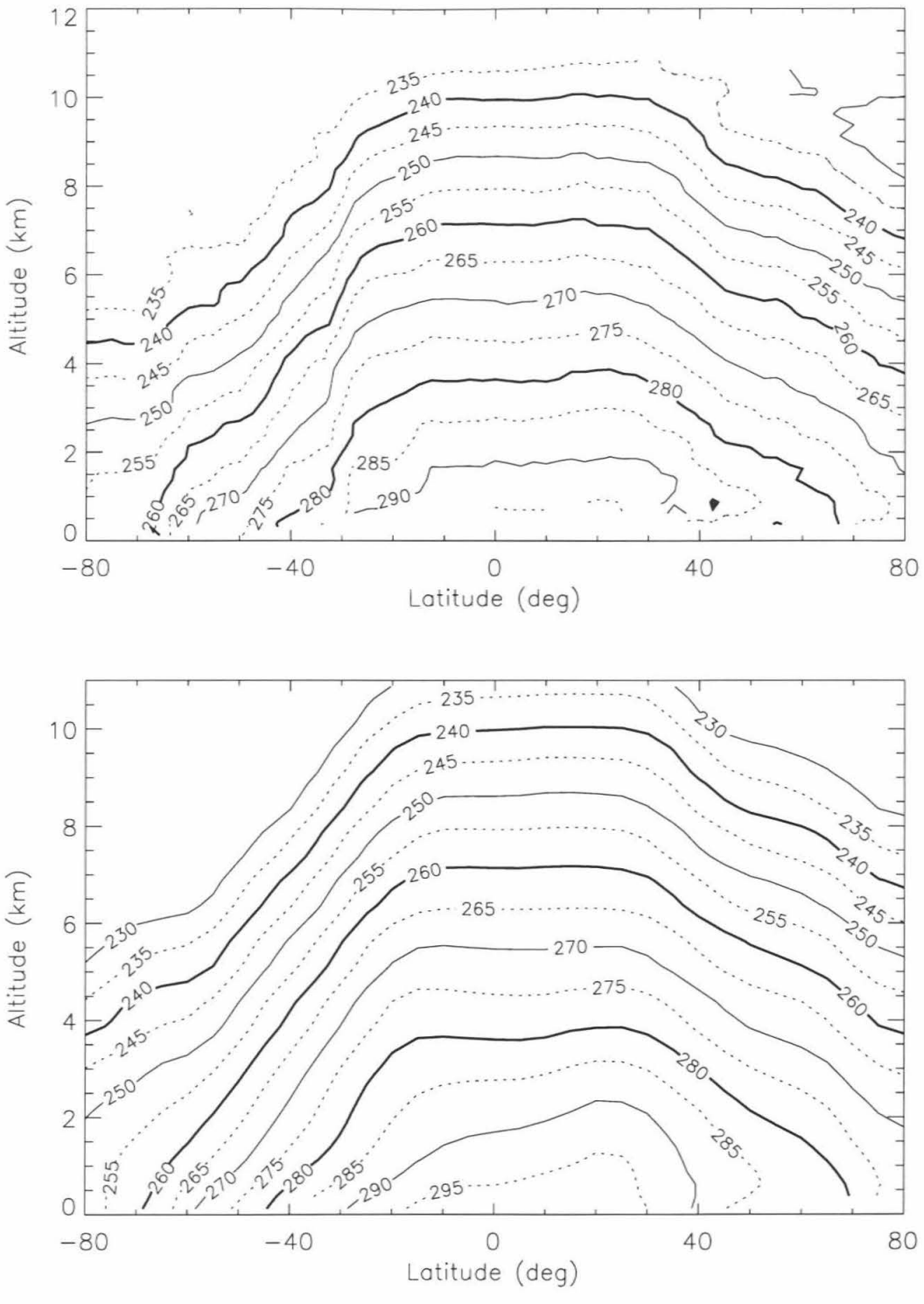

Figure 7.16 Mean temperature from a. IEC and b. CEC (in K).

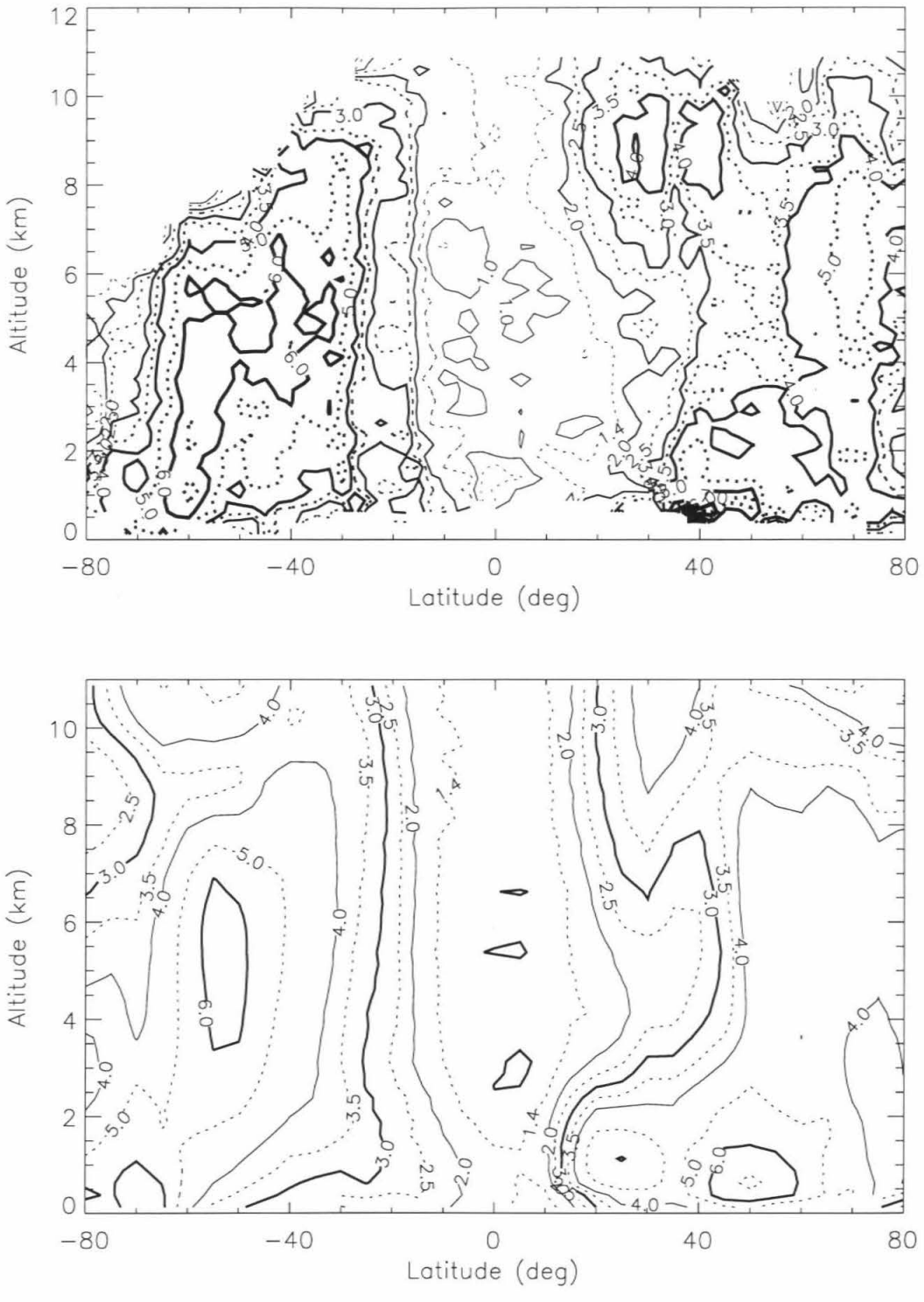

Figure 7.17 Standard deviation of temperature from a. IEC and b. CEC (in K).

few regions, the overall agreement between IEC and CEC indicate the occultation sampling is adequate to derive a representative latitude versus height climatology for the period.

At the altitude where occultation-derived temperature reaches 230K and utilization of ECMWF temperatures begins, some temperature biases are apparent. ECMWF temperatures are colder than occultation derived temperatures near 50S and north of 65N. ECMWF temperatures are warmer than occultation temperatures between 10N and 40N by as much as 5K.

Figure 7.17 shows the standard deviation of the IEC and CEC temperature data. With a few exceptions the two generally agree. A distinct minimum in the temperature standard deviation exists at low latitudes where values are between 1 and 1.4 K. Variability increases at mid-latitudes with maximum variability of 6K in the high baroclinicity region between 30S and 60S. Northern hemisphere variability is generally smaller. An unusual structure exists between 15N and 40N with maximum variability near 9 km altitude, lower variability near 5 km and somewhat higher variability near the surface. The similarity of the CEC σ_T structure above 2 km indicates the GEW climatology has captured behavior above 2 km. The smaller variability within 1.5 km of the surface exhibited in the IEC data probably indicates the reduced number of occultations reaching that altitude regime (figure 7.10a) are not entirely representative of behavior there. As suggested previously, low temperature variability implies that refractivity, when added to the existing knowledge base, will yield improved understanding of moisture at low latitudes in the lower half of the troposphere.

7.5.4 Pressure

We will now examine the pressure difference between the GEW and IEC results. It turns out that the pressure difference between the two is the dominant source of difference between the GEW and IEC humidities at higher altitudes.

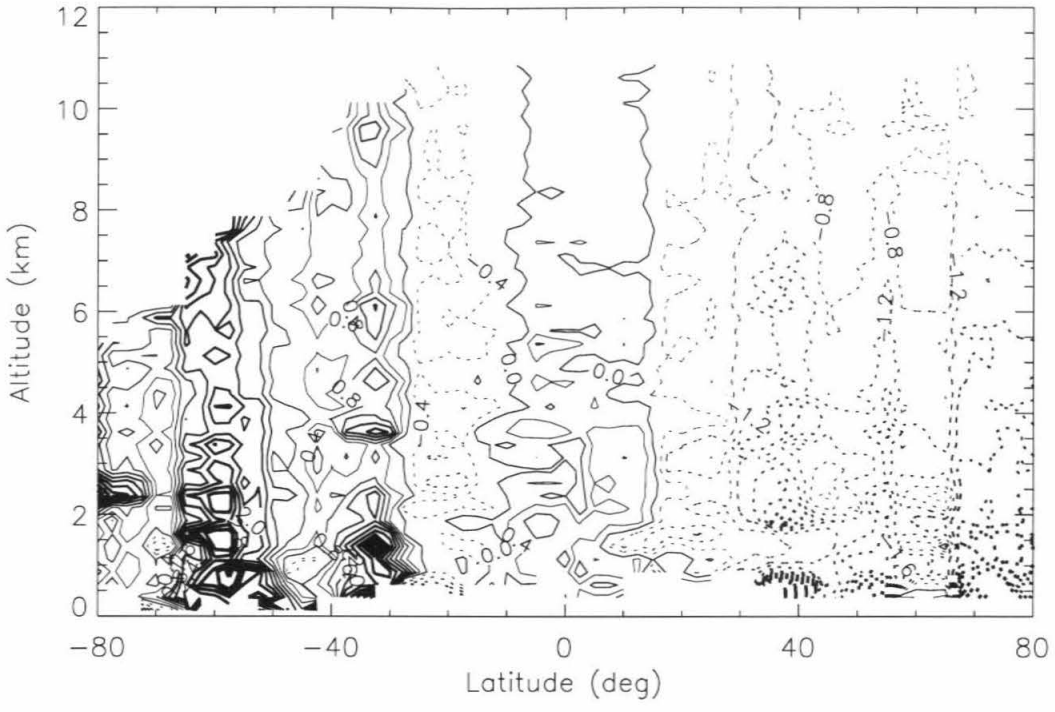

Figure 7.18 Mean of pressure difference, GEW minus IEC (in mbar).

Above the occultation-derived 230 K temperature level in the troposphere, pressure is derived directly from the occultation observations. Below this height, the ECMWF temperatures are used and basically define changes in pressure below the 230K height. At lower altitudes, the pressure bias due to the pressure bias at 230K scales in proportion to the overall increase in pressure. Refractivity does influence pressure below the 230K height through the mean molecular mass term but the impact is quite small. The difference between the occultation and ECMWF pressure structures (figure 7.18) primarily reflects the pressure difference at the 230K level.

Contrast across 30S

A sharp contrast in the pressure bias is evident across 30S where pressure differences to the south are positive ($P_{OC} > P_{EC}$) and negative to the north. 30S corresponds to the transition between the Hadley circulation and winter baroclinic zone. Upper tropospheric pressure derived from occultation observations is a measure of stratospheric mass which has large meridional scales and should produce little latitudinal dependence on occultation accuracy. Therefore, the latitudinal dependence of the pressure bias probably reflects a latitudinal bias in the ECMWF analyses due to a combination of surface pressure and tropospheric temperature biases. Since data assimilated by ECMWF south of 30S is primarily low vertical resolution satellite passive soundings, the change in bias across 30S may reflect the influence of radiosondes on the analysis. The effect of the latitudinal variation in the pressure bias on the derived humidities is discussed below.

Northern hemisphere bias

In the northern hemisphere, where the ECMWF analyses should be most accurate, a -1.2 to -1.6 mb bias exists near 300 mb whose cause is unclear. Occultation error sources include the ionosphere correction and the signal detection. Boundary conditions in initiating the Abel and hydrostatic integrals should not produce errors of this magnitude at

these altitudes (Section 3.10). Another possibility is the asymmetrical distortion of the vertical structure of the Fresnel zones due to the exponential structure of the atmosphere may systematically shift the vertical location of the phase center of the diffraction pattern relative to geometric optics solution. Such an effect would increase at low altitudes as refractivity increases. This possibility can be explored in a future study using the diffraction representation discussed in Sections 2.5 and 3.12.

The contribution of the ECMWF model to the observed pressure and refractivity bias structure can be considered with a simple error model consisting of a temperature bias and surface pressure bias. The fractional pressure error resulting from a surface pressure bias, δP_s, and a temperature bias, δT, is

$$\frac{\delta P}{P} = \frac{\delta P_s}{P_s} + \frac{z}{H_s}\frac{\delta T}{T}$$

The fractional dry refractivity error is then

$$\frac{\delta N}{N} = \frac{\delta P_s}{P_s} + \frac{\delta T}{T}\left(\frac{z}{H_s} - 1\right)$$

The observed refractivity bias ($\delta N = N_{OC} - N_{EC}$) north of 45S is very slightly positive near $z=H_s$ and generally negative below. Over this same latitude range, the pressure bias near $z = H_s$ is -0.3%. Taking $\delta N/N \sim 0$ at $z \sim H_s$ requires $\delta P_s/P_s \sim 0$. The pressure bias at $z \sim H_s$ requires that $\delta T/T \sim -0.3\%$. At lower altitudes, the zero surface pressure bias and negative temperature bias would cause ECMWF refractivities to be *smaller* than the occultation refractivities opposite to observed behavior indicating that a simple temperature plus surface pressure bias in the ECMWF model cannot explain the observed bias structure.

Adding a lapse rate error provides the additional degree of freedom to match observed behavior. For a constant temperature gradient, pressure and temperature are related as

$$\frac{P_1}{P_0} = \left(\frac{T_1}{T_0}\right)^{-\frac{gm}{R\dot{T}}}$$

where $\dot{T}$ is the vertical temperature gradient. The fractional temperature error due to biases in surface temperature, T_s, and temperature gradient is

$$\frac{\delta T}{T} = \frac{\delta T_s}{T} + \frac{z\,\delta\dot{T}}{T}$$

The fractional pressure error due to errors in surface pressure, surface temperature and temperature gradient is

$$\frac{\delta P}{P} = \frac{\delta P_s}{P_s} + \frac{z}{H_s}\left(\frac{\delta T_s}{T} + \frac{\delta\dot{T}}{\dot{T}}\right)$$

Combining the temperature and pressure contributions to dry refractivity yields

$$\frac{\delta N}{N} = \frac{\delta P_s}{P_s} + \frac{\delta T_s}{T}\left(\frac{z}{H_s} - 1\right) + \frac{z\,\delta\dot{T}}{T}\left(1 + \frac{T}{H_s\dot{T}}\right)$$

Taking the following conditions as representative: $\delta N/N(z=H_s) \sim 0$, $\delta N/N(z=0) \sim 1\%$ and $\delta P/P(z=H_s) \sim 5\text{x}10^{-3}$ yields $\delta T_s \sim -1.4$ K, $\delta\dot{T} \sim 0.04$ K/km and $\delta P_s \sim 4.4$ mb. While these errors are certainly possible, the pressure error in particular is larger than surface pressure bias is generally believed. Dividing these magnitudes by two might be more representative of ECMWF biases in the northern hemisphere. The implication is the occultation refractivities are slightly low perhaps for one of the reasons mentioned earlier. Further investigation is clearly warranted.

P and N Bias variation south of 50S

In the region south of 50S, between 800 and 500 mb, the fractional pressure and refractivity biases exhibit latitudinal bands which vary little with height which will create similar structure in derived water vapor. Given the very cold temperatures and small scale height of water, the near constancy with height suggests a latitudinally varying temperature

bias structure. Assuming the small amount of water is either negligible or estimated correctly by ECMWF, the fractional refractivity, pressure and temperature biases are related as

$$\delta N/N = \delta P/P - \delta T/T$$

where δ means occ - ECMWF. The pressure bias can be decomposed into contributions from refractivity bias and the pressure bias at 230K such that

$$\frac{\delta T}{T} = \frac{\delta P_{230} + \int_z^{z_{230}} \delta \frac{\partial P}{\partial z} dz}{P} - \frac{\delta N}{N} = \frac{\delta P_{230} + \int_z^{z_{230}} c\, \delta N\, dz}{P} - \frac{\delta N}{N}$$

If $\delta N/N$ varies linearly with height, the implied temperature error is

$$T_{EC} - T = T\left[-\frac{\delta P_{230}}{P} + a\frac{P_{230}}{P} - b\frac{\Delta z}{2}\left(1 + \frac{P_{230}}{P}\right)\right]$$

where $\delta N/N = a + b\ (z - z_{230}) = a - b\ \Delta z$ and $\Delta z = z_{230} - z$ and $a = \delta N_{230}/N_{230}$. For constant $\delta N/N$ (b=0), temperature error grows exponentially with altitude in proportion to pressure. $\delta T/T$ is constant when

$$b \cong \frac{1}{H}\left(\frac{\delta P_{230}}{P_{230}} - a\right) \cong \frac{1}{H}\left(\frac{\delta P_{230}}{P_{230}} - \frac{\delta N_{230}}{N_{230}}\right)$$

At 60S between 800 and 550 mb, $\delta N/N \sim 0$ and $\delta P/P \sim 0.4\%$ implying T_{ECMWF} is low by about 1K (0.4%) relative to the temperatures implied by the occultation N and P. At 70S between 700 and 550 mb, $\delta N/N \sim 0.3\%$ and $\delta P/P \sim 0.1\%$ implying T_{ECMWF} is biased high by ~0.6K relative to the occultation N and P. Since the absolute biases of the occultation versus ECMWF refractivity and pressure are suspect based on the northern hemisphere results and since the bias in the occultation results probably does not shift dramatically with latitude, it is likely that the temperature *change* between zones is correct such that bias in

T_{ECMWF} at 70S is roughly 1.6K higher than at 60S. In fact, adjusting ECMWF temperature across the entire region south of 30S in terms of a bias and lapse rate to match the pressure geopotential bias at T=230K matching that observed in the northern hemisphere may yield the best results.

7.5.5 Specific Humidity

We begin our examination of derived water vapor with specific humidity, q, the mass mixing ratio of water vapor. The GEW and ECMWF results are more similar to one another than to the P&O JJA climatology generally agreeing at the 0.1 g/kg level. $\bar{q}_{GEW}$ and $\bar{q}_{IEC}$ are both drier in the southern hemisphere and reveal a more peaked $\bar{q}$ structure near the ITCZ than the P&O results. $\bar{q}_{GEW}$ results tend to be drier than $\bar{q}_{IEC}$ below 500 mb in the crescent-shaped region centered at 10N near 800 mb discussed in Section 7.5.2. The tripole refractivity bias pattern along the bottom of the crescent pattern near 800 mb has also been converted to a moisture bias pattern where low latitude q_{GEW} is moister at the center and drier to either side and above. The location and symmetry of the crescent and the tripole suggest a relation with the Hadley circulation. Bias magnitudes are similar in the two hemispheres with fractional errors being two to three times smaller on average in the north because of the larger mixing ratios there. Accuracy south of 50S appears limited by the occultation versus ECMWF refractivity and pressure bias pattern discussed in Section 7.5.4. The spatial pattern of variability reveals a bimodal pattern with a maximum on either side of a minimum near the ITCZ generally similar to that of ECMWF and P&O for JJA but somewhat larger in magnitude. GEW results also indicate regions of high variability at high southern latitudes not indicated in the other data sets.

Average specific humidity

To first order, $\bar{q}_{GEW}$ has a well-defined peak centered in the vicinity of the ITCZ (figure 7.19). The altitude of contours of constant $\bar{q}$ decrease roughly linearly as a

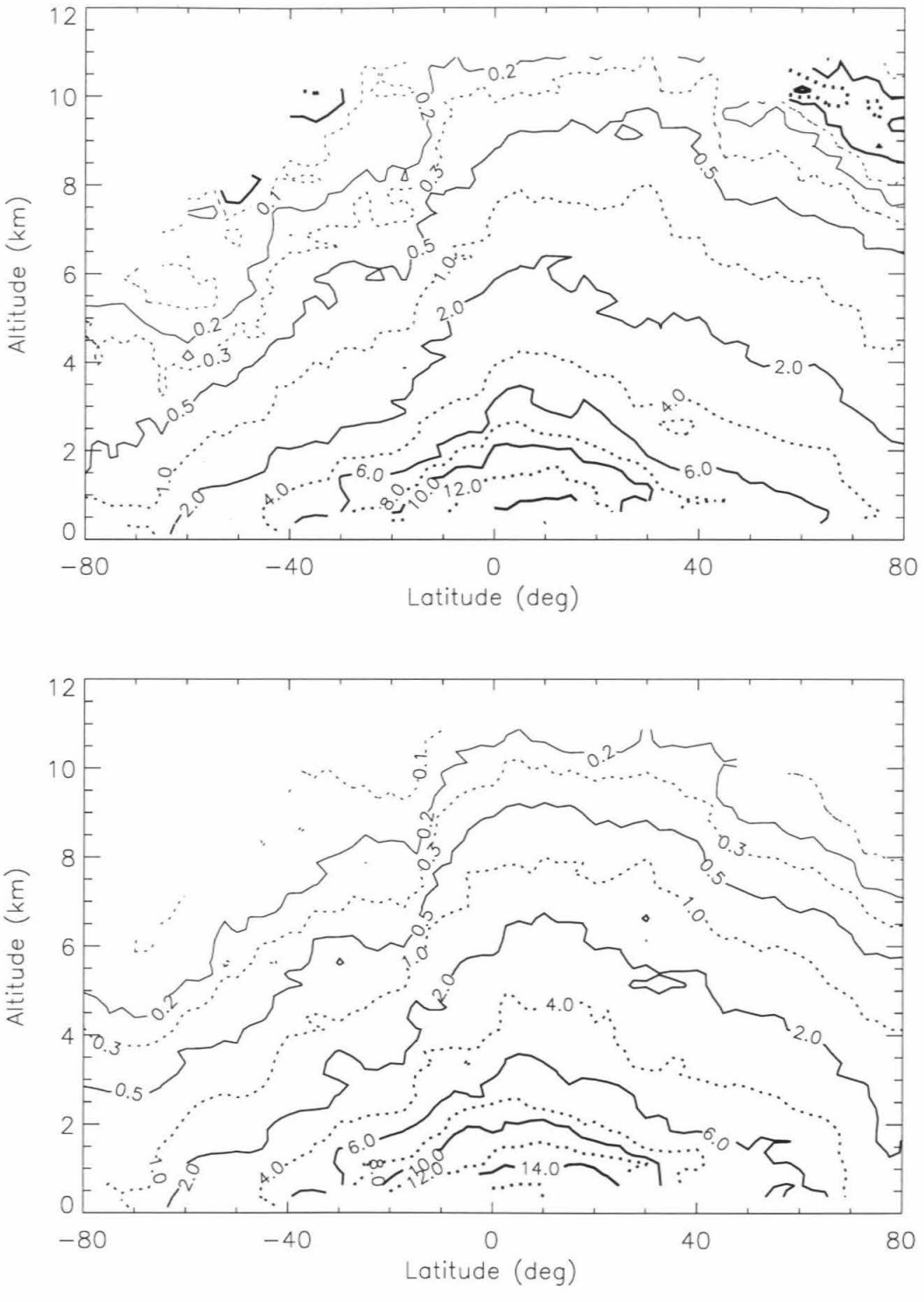

Figure 7.19 Mean specific humidity (g/kg) from a. GEW and b. IEC.

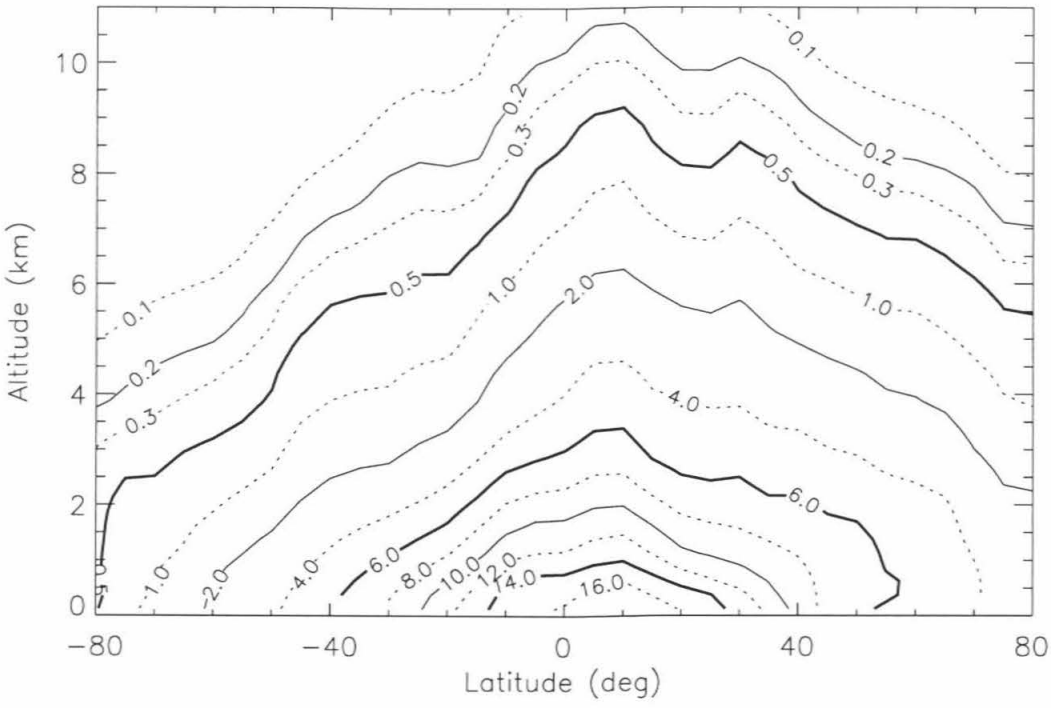

Figure 7.19 c. Mean specific humidity (g/kg) from CEC.

function of latitude away from the peak. The decrease toward higher latitudes falls off more rapidly in the southern hemisphere than the northern hemisphere qualitatively consistent with the more rapid decrease of temperature with increasing latitude in the southern (winter) hemisphere. Consistent with summer-winter hemisphere difference, $\bar{q}$'s in the northern hemisphere are much larger than their southern hemisphere counterparts. For example, at 750 mb at 50° latitude, $\bar{q}$ is ~4 g/kg in the north, whereas in the south it is ~1 g/kg. A second peak in $\bar{q}$ exists near 30N above 400 mb. There may be a much smaller peak near 30S near 600 mb. Over Antarctica above the 650 mb level there is a relative maximum in $\bar{q}$ which is supersaturated w.r.t. ice which may be due to a positive bias in ECMWF temperatures.

Comparison with P&O and ECMWF climatologies

In general $\bar{q}_{GEW}$ and $\bar{q}_{IEC}$ are more similar to one another than either is to the P&O JJA climatology. $\bar{q}_{GEW}$ and $\bar{q}_{IEC}$ are both significantly more peaked near the ITCZ. The $\bar{q}_{GEW}$ and $\bar{q}_{IEC}$ contours exhibit secondary bumps near 30N and 30S and valleys near 20N and 20S which are either slightly or not visible in the P&O JJA climatology. The longer estimation period of the P&O climatology may have smeared the more instantaneous behavior represented in the GEW and ECMWF data. The general coincidence of bumps and valleys in the GEW, IEC and CEC data indicate the features are real, climatologically representative. Further, as we will discuss in Section 7.6, these features appear to be fundamentally related to the control of atmospheric water vapor. Since the P&O data does not include the globally distributed data assimilated by ECMWF or the globally distributed, high vertical resolution information of the GEW dataset, it may suffer from limited coverage particularly at low latitudes and in the southern hemisphere. The ability of a sparse 11-day global sampling by GPS to reveal basic features not apparent in 10 years of radiosonde data indicates the importance of global coverage.

Both the GEW and IEC $\bar{q}$'s in the southern hemisphere are distinctly lower than those of P&O above ~1 km altitude. Since ECMWF temperatures are less accurate in the southern hemisphere, these biases could in theory be due to biases in ECMWF temperatures, the impact of which can be assessed by considering the following situation. Given a true $\bar{q}$ of 2 g/kg and B of order 17, a derived $\bar{q}$ which is 30% low requires a temperature bias of -4.5 K which is very large and unlikely relative to the occultation vs. ECMWF temperature comparisons (figure 6.3). Further, a constraint can be placed on the vertically averaged value of any ECMWF vs. occultation temperature bias between the surface and 300 mb level by comparing the occultation and ECMWF geopotentials at 300 mb. The results of Leroy (1996) and figure (7.18) indicate that the geopotential heights of pressure levels are consistent to within ~20-30 m on average such that the average temperature bias is < 0.7-1.0 K assuming the bias is not due to surface pressure error. However, this conclusion may be flawed given the systematic difference in pressure bias between the southern and northern hemispheres. The fact that south of 30S southern hemisphere occultation versus ECMWF pressure biases near 400 mb are lower than their northern hemisphere counterparts by 1.4 to 3.8 mb may indicate southern hemisphere ECMWF temperatures are actually biased low by 0.9 to 2.5 K. Still, this range of biases is smaller than that required to explain ~30% difference between the GEW and P&O $\bar{q}$'s in the southern hemisphere except at higher altitudes. For an error of 0.7 K, the water vapor accuracy at 750 mb near 40S (B ~ 25) is ~7%. The accuracy estimated by extrapolating downward from the altitude of 100% error yields an accuracy of ~5% at 750 mb near 40S.

Biases in derived refractivity, another source of water vapor error, should be much less than 1% (figure 3.20a), according to simulations, except perhaps within 1 km of the surface where the number of retrievals diminishes. Further, there is little reason to expect different accuracy of the refractivity retrievals in the southern hemisphere in comparison with the northern hemisphere. In the southern hemisphere, ECMWF and occultation mean

refractivities are generally consistent to better than 0.5% outside the southern pole of the tripole bias region. For the case above with B~17, a 0.5% refractivity error translates to an 8% water vapor error. Therefore, the GEW $\bar{q}$ near 750 mb and 40S appears to be correct at the ~5-10% level implying that the southern hemisphere during the June 21 - July 4, 1995, period is indeed significantly drier than the JJA climatology of P&O.

The similarity between the GEW and IEC $\bar{q}$ in figure 7.19 is striking particularly in contrast to differences between the GEW and P&O JJA $\bar{q}$ climatologies. Notable similarities between $\bar{q}_{GEW}$ and $\bar{q}_{IEC}$ include

A distinct peak in $\bar{q}$ near the mean ITCZ

Similar falloffs of $\bar{q}$ with latitude away from ITCZ

A relative maximum at 30N above 400 mb

A small peak near 30S between 650 and 450 mb

A peak in 6 g/kg contour at 20N

A relative maximum over Antarctica above the 700 mb level

Notable differences include

The crescent and tripole bias structure (Section 7.5.2)

A relative maximum in $\bar{q}_{ECM} = 2$ g/kg contour at 30S which does not appear in $\bar{q}_{GEW} = 2$ contour

The vertical structure of the rapid change in $\bar{q}$ across the 10 to 15S band differs

The magnitude of the maximum over Antarctica differs

Crescent Bias: Figure 7.20 shows the mean of the differences between the ECMWF and GEW q's indicating GEW $\bar{q}$'s are lower on average than those of ECMWF below the 550 mb level between 65S and 65N more or less in the same crescent shaped pattern as the refractivity bias in figure 7.14. The negative bias is as large as -1 g/kg. In the northern hemisphere, the fractional bias is typically 5 to 10% but can be greater than 20% in places. In the southern hemisphere fractional biases are larger more typically 10 to

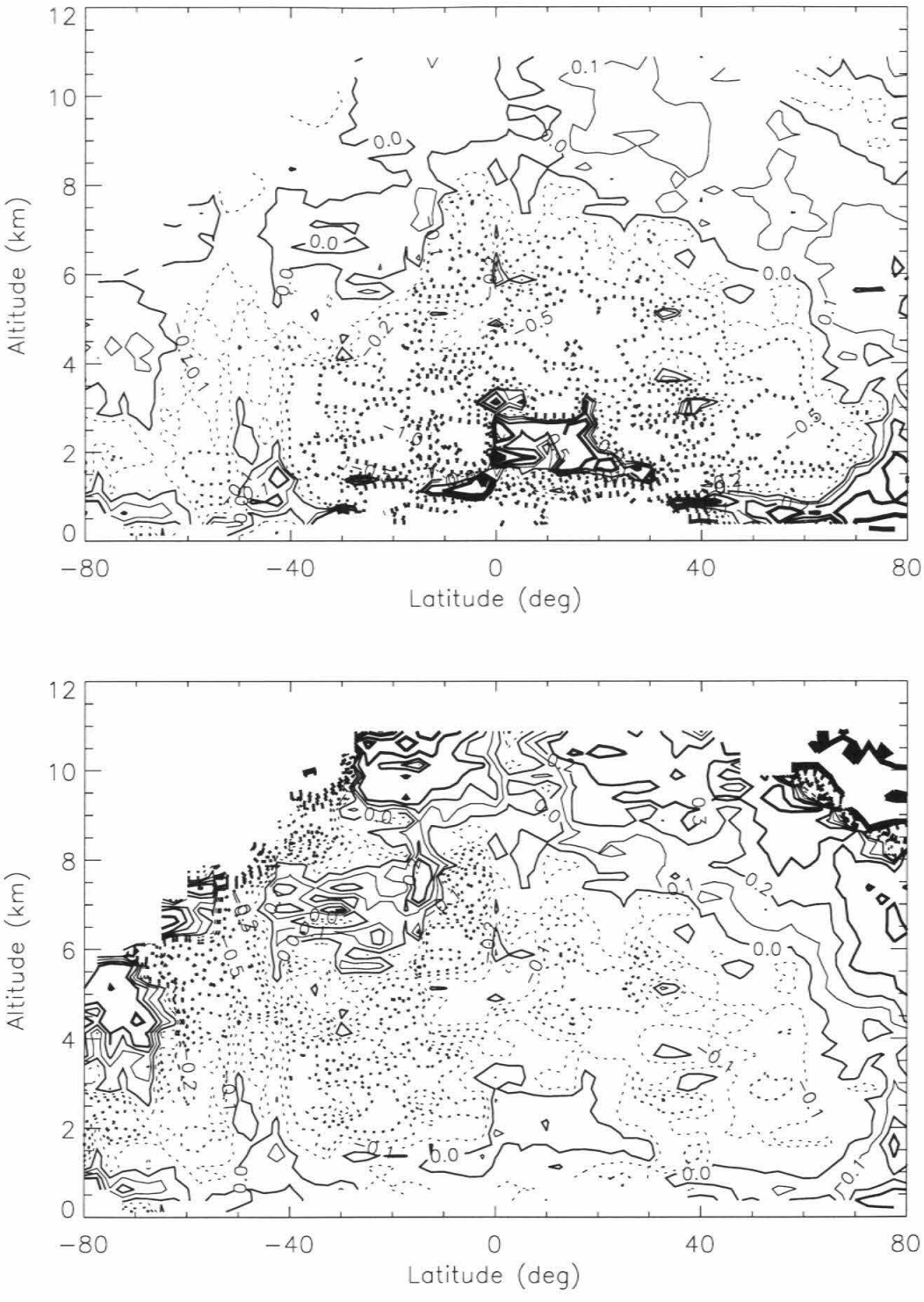

Figure 7.20 Mean of specific humidity difference, GEW minus IEC a. g/kg and b. normalized.

20% because of lower mean q's. At high latitudes in both hemispheres, biases are somewhat similar in structure such that above the 650 to 700 mb level, GEW $\bar{q}$'s are higher than ECMWF mixing ratios, and in the lowest part of the troposphere (<1.5 km in the south and <3 km in the north) GEW $\bar{q}$'s are larger than ECMWF. In the region between 550 and 400 mb and 65N to 80N, occultation humidities exceed ECMWF humidities by 20 to 30%.

To the north of the ITCZ, the region between 500 and 750 mb in the GEW climatology is drier than ECMWF with peak biases larger than -0.7 g/kg. GEW derived mixing ratios are about 20% lower than ECMWF humidities in this region. The difference between GEW and ECMWF <q>'s changes sign quite rapidly near 750 mb in the latitude band from 0 to 20N.

Tri-pole bias: After compensating for the dry term, the tri-pole refractivity bias has been translated to a bias between the GEW and IEC q's centered near 10N between 1 and 3.5 km altitude. GEW $\bar{q}$'s are distinctly higher than ECMWF in the center region and lower to either side. The region of higher mean GEW q lies between the equator and 20N between 850 and 750 mb and extends farther north to 30N at 850 mb and to the south to 15S at heights between 900 and 850 mb. Maximum bias in this region is greater than 0.5 g/kg and has a horseshoe pattern about the ITCZ.

To the north side of the positive bias region, a smaller region exists centered at 35N and 750 mb in which the GEW climatology is noticeably drier than ECMWF. The region lies at the lower end of the low relative humidity region defining the northern descending branch of the Hadley circulation.

In the region near the southern descending branch of the Hadley cell, GEW q's are systematically lower than IEC values. The largest bias of > 1 g/kg exists between 850 and 750 mb between 30S and the equator. The bias of -0.5 g/kg near 30S between 650 and

700 mb is a fractional bias of 30% corresponding to the peak in the IEC 2 g/kg contour which is absent in both the GEW and CEC climatologies (figure 7.19). It may be an artifact of assimilation given the limited vertical resolution of TOVS. Biases in this region are due in part to the altitude of the tradewind inversion being too high in the analyses. The average position of the inversion seems to be near 850 mb based on vertical gradients of q_{GEW} in figure 7.19. The bias is actually worse than figure 7.20 indicates because a number of profiles from this region were excluded from the GEW climatology because large ECMWF temperature errors caused derived humidities to be negative at low altitudes which failed the selection criteria. On the other hand, since boundary layer temperatures are colder than temperatures in the overlying subsidence region, the bias between GEW and ECMWF humidities could be overestimated because ECMWF temperatures are too cold between the correct and overestimated inversion height leading to underestimated GEW humidities there. Examples of GEW derived-humidities near 20S are given in figure 7.21 demonstrating the vertical structure of the humidity retrieval error in this latitude band is distinct and easily recognizable.

The tri-pole structure is consistent with an analysis indicating the TOVS vertically integrated water content is systematically overestimated in the subtropics and underestimated in the tropics with respect to that of SSM/I [Wittnmeyer and Vonder Haar, 1994]. The GPS comparison indicates this tropical-subtropical bias structure is largely associated with the inversion height and seems to be associated with errors in the tropical Hadley circulation as represented in the ECMWF model. On the southern side, the tradewind inversion is too high in the model such that in the altitude interval above the true inversion but below the model inversion temperatures are too low and model humidities are too high. In the 1.5 to 3.5 km altitude range, the information assimilated by ECMWF comes largely from the lowest TOVS moisture channel which spans the 1000 to 700 mb interval. The bias may represent real atmospheric structure which the analyses have smoothed over because of limited vertical information from TOVS. The bias may also be

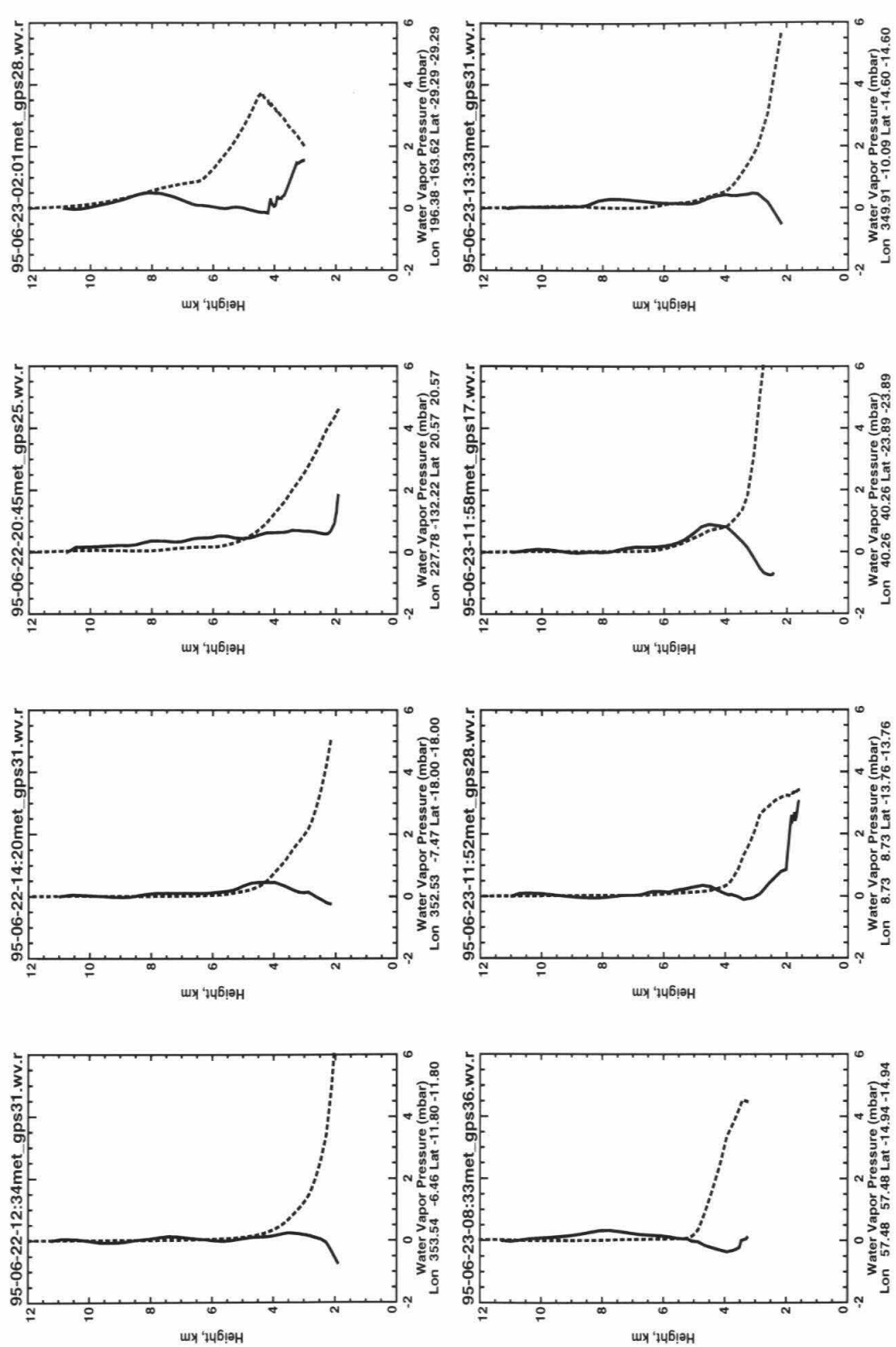

Figure 7.21 Examples of water vapor partial pressure profiles from June 22 and 23, 1995, where the tradewind inversion height is overestimated in the ECMWF analyses. Solid lines are GEW profiles. Dashed lines are from interpolated ECMWF analyses

related to surface emission in the lowest TOVS moisture channel although the surface of most of this latitude band is water whose emissivity is better known than that of land. It may also be related to clouds. Unfortunately, since the GPS data does not extend to the surface, we cannot determine the behavior of the q bias over the entire altitude range of the lowest TOVS channel.

High southern latitude bias structure: q bias structure south of 50S exhibits a latitudinally alternating bias pattern related to the refractivity and pressure bias patterns discussed in Section 7.5.4. Over Antarctica, the GEW and ECMWF $\bar{q}$'s both exhibit a relative maximum above the 700 to 650 mb level although the GEW values are significantly larger in a fractional sense reaching supersaturated w.r.t. ice above 600 mb south of 70S. As discussed in Section 7.5.3, the refractivity and pressure biases exhibit an oscillatory dependence on latitude in this area producing the latitudinal $\bar{q}$ bias pattern in figure 7.20. Since occultation refractivity accuracy should vary little with latitude under such dry conditions, the latitudinal changes are more likely due to ECMWF particularly given the region's sparse radiosonde sampling. Table 7.1 shows the effect of the biases at 500 mb at 60S and 70S over which the largest fractional change in GEW $\bar{q}$ occurs indicating the ECMWF temperature bias seems to increase by 1.3K from 60S to 70S. The temperature bias necessary to cause a 30% supersaturation near 500 mb is about +0.5 K which is not large particularly in the southern hemisphere. Combining the various pieces of information, the most likely scenario is ECMWF temperatures are biased slightly high above Antarctica causing GEW and ECMWF q's to be slightly overestimated implying that the true q's are slightly lower than q_{ECMWF}.

The maximum does exist in both the IEC and GEW results as well reflecting near saturation values there for the occultation sampling. However, the CEC results do not show the relative maximum over Antarctica indicating the occultation results are not climatologically representative, consistent with the difference between IEC and CEC mean

temperatures over Antarctica. In this region mean occultation refractivity is higher than that of ECMWF by 0.2 to 0.5%. The pressure of the occultation is high by 0.6 to 2.8 mb (0.13 to 0.6%). The dramatic rise in q over Antarctica in the GEW results is due to the latitudinal decrease in the pressure bias between 60S and 70S of ~0.4% causing an increase in q from 60S to 70S of ~0.1 g/kg. The pressure bias between the occultation results and the ECMWF analyses (figure 7.18) will bias $\bar{q}$ derived in the GEW results. From eq. (4.2.1) and the definition of q, the change in q due to a change in P is

$$\frac{\partial q}{\partial P} = -\frac{m_w b_1 T}{m_t b_2 P} \tag{7.5.1}$$

From eq. (7.5.1), a 0.3% error in P will produce an error in q of 0.1 g/kg. Since 0.3% is representative of the magnitude of occultation versus ECMWF pressure bias, the absolute accuracy of the occultation q's are limited near the 0.1 g/kg level

Table 7.1

	60S	70S	70S-60S
ΔN/N(%)	0.15	0.3	0.15
ΔP/P(%)	0.5	0.1	-0.4
equivalent ΔT (K)	-0.8	0.5	1.3
Δq (g/kg)	-0.11	0.06	0.17

Upper troposphere: Above the 500 to 400 mb level at latitudes north of 35S, GEW $\bar{q}$'s are systematically higher than those of IEC possibly indicating a systematically moister upper troposphere than that in the ECMWF analyses. Any bias at these levels is important because of the radiative influence of upper level water vapor on surface temperature.

Although the occultation and IEC refractivities agree to within +0.2% in this region, in general (figure 7.14), the pressure bias between the two data sets of order -0.5% causes the $\bar{q}_{GEW}$ to be higher than $\bar{q}_{IEC}$ by ~1.5 g/kg. As discussed in Section 7.5.4, the cause of the pressure bias is not clear but may be in the occultation data. If $\bar{q}_{GEW}$ were correct, the mean air would be close to or at saturation w.r.t. ice at higher latitudes in the northern hemisphere which seems unlikely because it would require that σ_U be small which is not indicated in the CEC data (see Section 7.5.6). In the future skewness in the distribution of relative humidities could be checked as an indication sign of condensation effects.

Climatological representativeness: Examining differences between IEC and CEC behavior should reveal biases associated with the limited GEW sampling. In the southern hemisphere, overall agreement is good. IEC may be very slightly drier on average. In the northern hemisphere, between 15N and 40N above 6km altitude, the IEC data is definitely moister than CEC by 0.1 to 0.2 g/kg.

Accuracies implied by negative and supersaturated values of $\bar{q}$.

Regions exist in the upper troposphere near the 230 K contour where $\bar{q}_{GEW}$ is negative (figure 7.19) implying that occultation-derived temperatures are systematically lower than ECMWF temperatures in this region which is consistent with the southern and northern hemisphere temperature comparisons for this period. Further, 230 K contours are occasionally visible in the IEC mean temperature (Figure 7.16a). Since GEW data is selected only when occultation temperatures exceed 230 K, occultation-derived temperatures must reach 230 K at higher altitudes than the ECMWF temperatures in regions where 230K contours are visible.

Based on the CEC results, extrapolation of the P&O data and the occultation data itself, the true $\bar{q}$ near 230 K is ~0.1 g/kg so the accuracy of $\bar{q}_{GEW}$ is ~ 0.1 g/kg. Combining eq. (4.2.1) and the definition of q, the fractional error in q is

$$\frac{dq}{q} \cong (B+1)\frac{dN}{N} + (B+2)\frac{dT}{T} - (B+1)\frac{dP}{P} \qquad (7.5.2)$$

where the approximate equality is because the small effect of water on the mean molecular mass has been ignored. When B > 10, eq. (7.5.2) can be approximated as

$$dq \sim \frac{m_w T b_1}{m b_2}\left(\frac{dN}{N} + \frac{dT}{T} - \frac{dP}{P}\right) \qquad (7.5.3)$$

where dq is unitless. Given that the accuracy of $\bar{q}_{GEW}$ is ~ 0.1 g/kg near the 230 K level, we would like to try to estimate its accuracy at lower altitudes via extrapolation. The increase in T from 230 K to surface values will increase the scale factor in eq. (7.5.3) by 10 to 30% at the surface. The pressure error changes across the troposphere because above the 230K altitude, pressure is derived entirely from the occultation-derived densities whereas the pressure increase at lower altitudes is derived using the ECMWF temperatures in the hydrostatic relation. Therefore, the pressure error in the lower troposphere will generally be dominated by the mean ECMWF temperature error. If the ECMWF temperature error were constant with altitude, the pressure error would equal -δT/T one scale height below the 230 K level (Section 7.5.7). Therefore, the contribution of the temperature error might be doubled near the surface. The refractivity error is harder to estimate. Simulations show refractivity biases are very small (Figure 3.20). However, there are signs of a refractivity bias in the occultation results (Sections 7.5.2 and 7.5.4). Taking these scaling arguments into account and assuming the refractivity bias is of order 1% or less, the accuracy of $\bar{q}_{GEW}$ in the lower troposphere is probably between 0.2 and 0.4 g/kg.

Variability of q

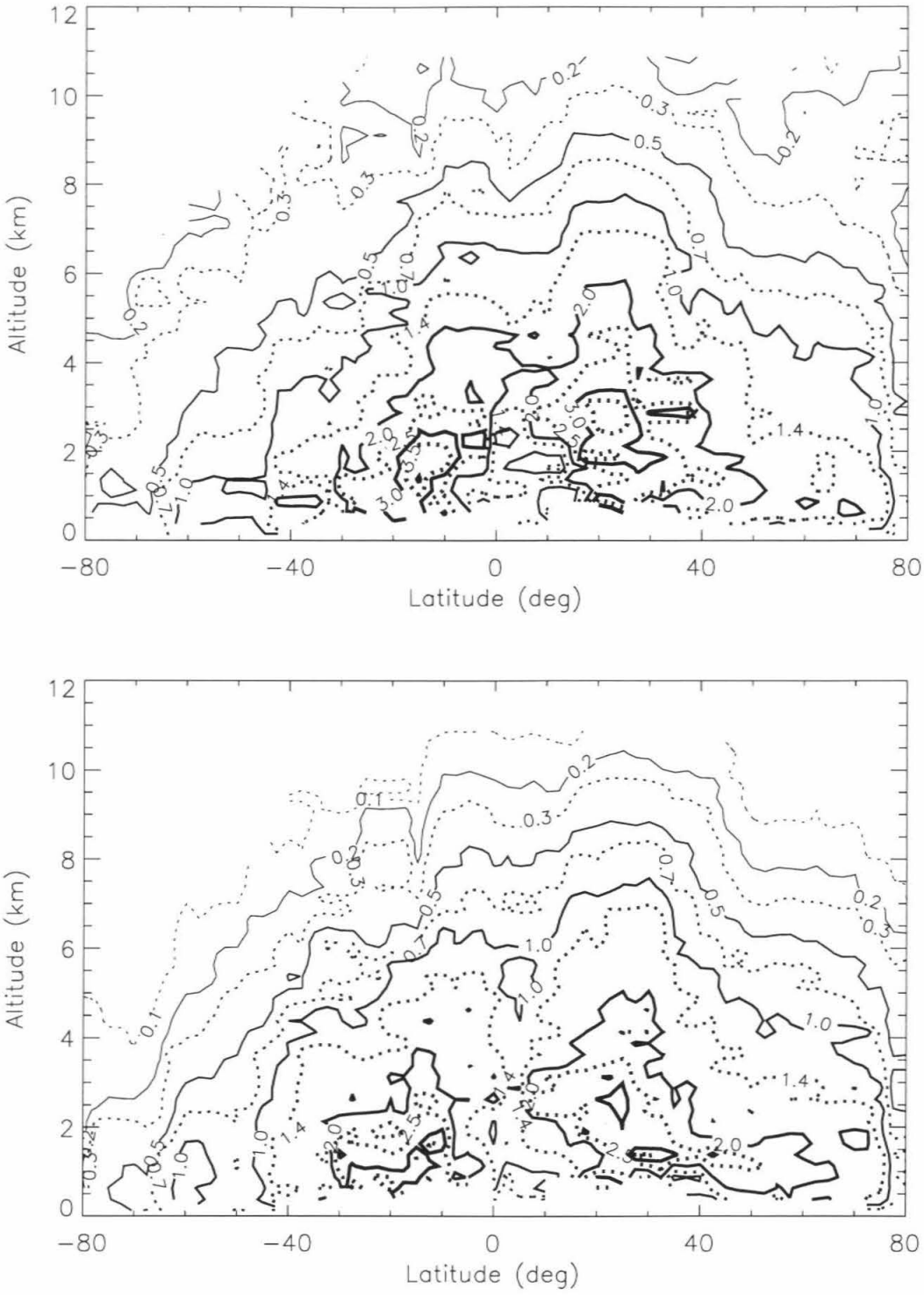

Figure 7.22 Standard deviation of specific humidity from a. GEW and b. IEC (g/kg).

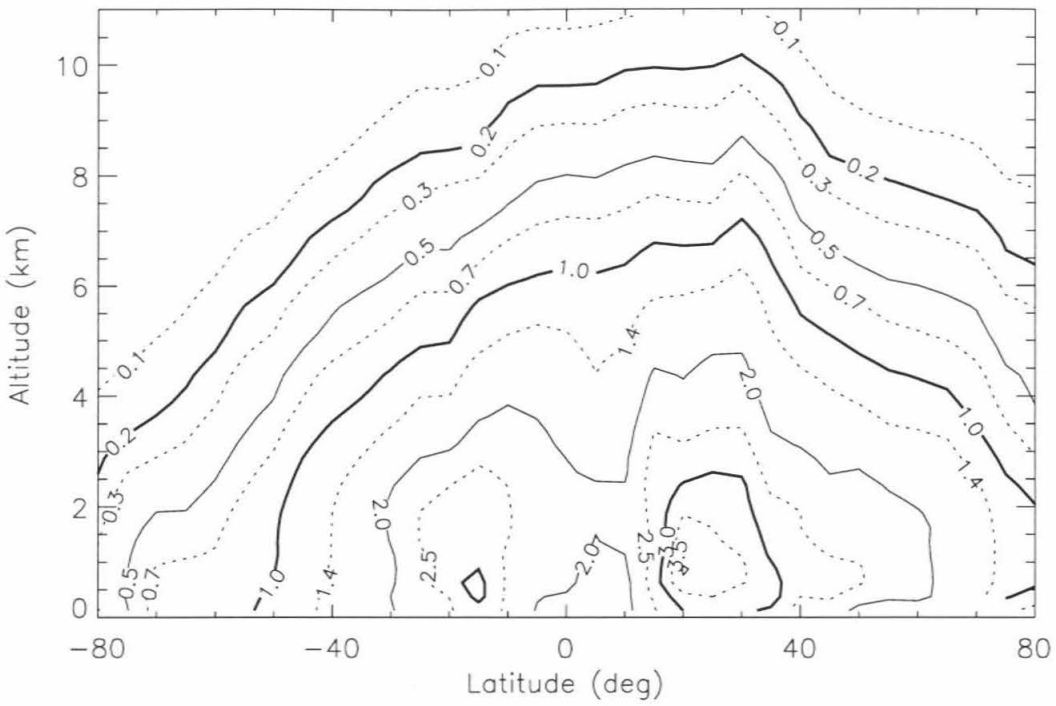

Figure 7.22 c. Standard deviation of specific humidity from CEC (g/kg).

The latitude versus height dependence of the variability of q as represented by the standard deviation of q, σ_q, is similar to the standard deviation of refractivity (figures 7.22 and 7.13). The general GEW behavior is similar to ECMWF and P&O results. At low latitudes a peak sits to either side of the equator straddling a minimum at the equator. Figure 7.22 demonstrates that most, although not all, of the bimodal structure of refractivity variability (figure 7.13) is due to moisture. The northern variability peak extends up above the 300 mb level where it has shifted south from ~25N at 750 mb to 15N at 300 mb. GEW variability is somewhat larger than ECMWF variability consistent with underestimated TOVS-derived humidity variability noted by Wittmeyer and Vonder Haar [1994] associated with systematic problems in recovering extremely humid events.

The sampling of the P&O climatology is able to distinguish between two sources of the double-peaked variability of q, namely stationary eddies which are steady atmospheric features such as the subtropical anticyclones, and transient eddies which come and go in time. The density and duration of the GEW data set are not sufficient to distinguish between these two eddy types. The peak values observed here are significantly larger than the variations of the combined eddy types of P&O (> 3.5 vs. ~2.5 g/kg). Further, the σ_q peaks in GEW occur at somewhat higher altitudes. The latitudes of the GEW σ_q peaks correspond more closely with those of the stationary eddy peaks which are only characterized for annual mean conditions. The magnitude of the variations in the P&O climatology associated with each type is estimated in terms of day-to-day variations. The peak magnitude on each side of the equator of the P&O stationary eddy variations is ~1.25 g/kg which could be underestimated if the cycle of variability is longer than a day. The ratio between the maximum and the minimum near the equator is similar in both climatologies at 2.5. The northern hemisphere peak is also the larger maximum at higher altitudes in all three climatologies.

A distinct peak in q variability exists between 45S and 60S associated with the refractivity variability peak observed there. Much of the refractivity variability is due to the high variability of T in the region but roughly 40% of the variance is due to moisture assuming the ECMWF temperatures are correct. IEC moisture variability exhibits a weak peak near the same location which is not present in the CEC results.

Standard deviation of differences and accuracy

We will now attempt to sort out accuracy of individual moisture profiles. Errors are contributed by occultation refractivity, ECMWF temperature and ECMWF humidity. The standard deviation of the difference between GEW and IEC specific humidity, $\sigma_{qO\text{-}E}$ (figure 7.23) structure differs from that of the GEW σ_q (figure 7.22) exhibiting peaks as large as ~2.5 g/kg with the discrepancy at the σ_q maxima ranging from 1.4 to 2 g/kg. The fractional differences in q can be large particularly in the upper southern hemisphere where standard deviations of the q differences near the 230 K contour are greater than 200% (figure 7.23). As will be discussed in Section 7.5.7, negative mean relative humidities in this area imply ECMWF temperatures are particularly noisy here contributing much, if not most, of this error. A 3 K temperature standard deviation based upon the temperature error inference in Section 7.5.7 combined with $\bar{q}$ ~ 0.1 g/kg in this region (B~300) yields a 400% standard deviation of specific humidity via eq. (4.2.1), consistent with observed inconsistencies.

In Section 4, assuming individual profile rms refractivity accuracy of 0.2% in the upper regions decreasing to 1% near the surface and rms temperature accuracy of 1.5K, the predicted 100% rms humidity error of individual GEW profiles is roughly at $\bar{q}$ = 0.2 g/kg. The $\sigma_{qO\text{-}E}$ behavior (figure 7.23) varies significantly relative to this accuracy prediction. In the worst region between 60S and 10S, the difference rms ranges from a factor of 1.4 to 5 worse than predicted with the biggest discrepancies occurring in altitudes spanning the 0.3 and 1 g/kg contours. North of 25N, agreement ranges from 40 to 70% near the 0.2 g/kg level, which is better than predicted indicating the standard deviation of deviations between

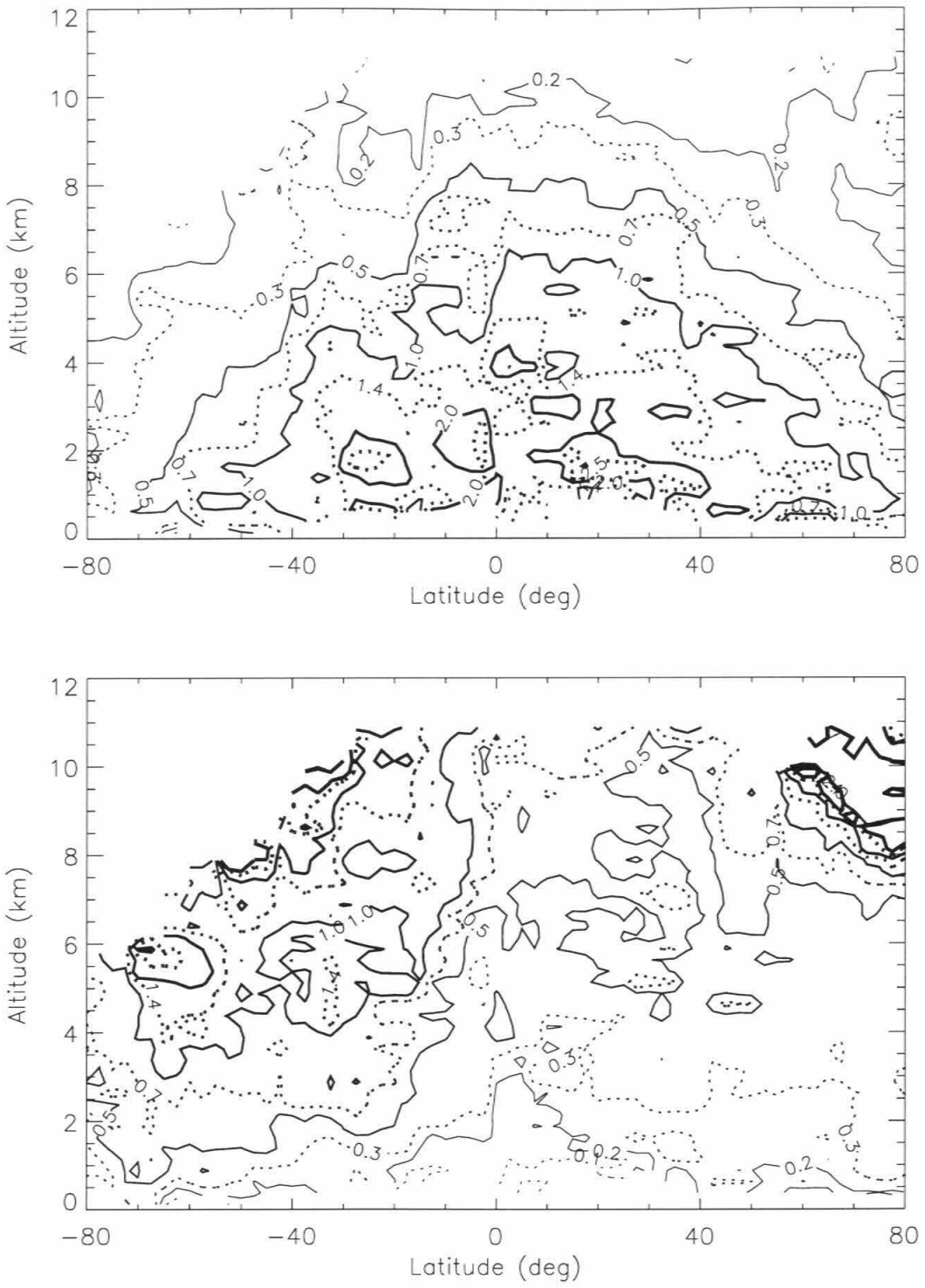

Figure 7.23 Standard deviation of specific humidity difference, GEW minus IEC.
a. g/kg and b. normalized.

ECMWF and occultation temperatures is ~ 1 K (0.4%) in the northern hemisphere. This may also reflect some correlation between GEW and ECMWF humidities via the common ECMWF temperature. Since refractivity accuracy in dry conditions should depend little on latitude, the difference between northern and southern hemisphere consistency near the 0.2 g/kg level probably reflects a hemispheric difference in ECMWF temperature and humidity errors. A fractional rms difference of 70% at 0.2 g/kg representative of the upper northern hemisphere region bias represents consistency between N, P and T at the 0.5% level.

According to eq. (4.2.1), if the N, P and T error magnitudes remain approximately constant with altitude, the fractional GEW accuracy should improve approximately in proportion to q. However, the agreement between ECMWF and GEW humidities generally improves more slowly at lower altitudes implying consistency is limited by a combination of increasing refractivity error at lower altitudes and finite ECMWF humidity accuracy. For example, at high northern latitudes agreement improves from 50% to 30% while specific humidity increases by a factor of 8 such that the two extreme explanations are either that refractivity error has grown by a factor of ~5 at lower altitudes or ECMWF is contributing a large portion of the difference or some combination of the two. The best agreement at the 20 to 30% level in the lower troposphere is comparable to quoted accuracies for ECMWF humidity accuracy and may indicate the ECMWF humidity error dominates at low altitudes. Simulated occultation retrieval accuracy does degrade by a factor of 5 between 7 km and 2 km altitude and occultation refractivity error might also be a significant source of error. It is not clear at this point what the relative contributions are in the northern hemisphere.

In the particularly noisy regions of the southern hemisphere near $\bar{q}$=0.5 to 1 g/kg, the standard deviation of q differences is equivalent to a 3% difference in some combination of N, P and T if the ECMWF humidities were perfect. It is hard to imagine that errors in either the ECMWF temperatures or the occultation refractivities could approach this

amount. At the other extreme, if ECMWF humidities were entirely to blame, they would have to be in error by ~100%. If we assume the northern hemisphere errors such that the refractivity error is ~1% and assume 3 K rms temperature error based on negative mean relative humidities discussed in Section 7.5.7, the ECMWF humidity error would be ~70% assuming the errors are independent. This is preliminary and better separation of the contribution may be achieved with a Bayesian approach utilizing the individual characteristics of each noise source such as the vertical correlation of ECMWF humidity due to the TOVS weighting functions.

Even when the GEW-ECMWF discrepancy is conservatively assumed to be due entirely to occultation retrieval error, the residual σ_q, after subtracting the contribution of σ_{qO-E} = 2 g/kg to the total variance is 2.9 g/kg, which is larger than the largest P&O σ_q values. Therefore, the GEW results indicate subtropical humidity variability is larger than that of P&O.

7.5.6 Relative Humidity

Relative humidity represents the water vapor abundance relative to the maximum vapor abundance at the point of condensation. All relative humidities have been estimated over liquid water as is generally done. Figure (7.24) shows the saturation vapor pressure over liquid and solid water for reference. The correct form of the saturation vapor pressure is somewhat tricky because at temperatures below but within 40K of freezing, the condensed form of water may be either ice or water.

Because the ECMWF temperatures are noisy at high southern latitudes, relative humidities estimated there are particularly noisy. In fact, mean relative humidities can be negative there even when q is positive, a fact which is used in Section 7.5.7 to derive the rms error in the ECMWF temperatures in that region. Therefore, the mean relative humidity will be estimated as

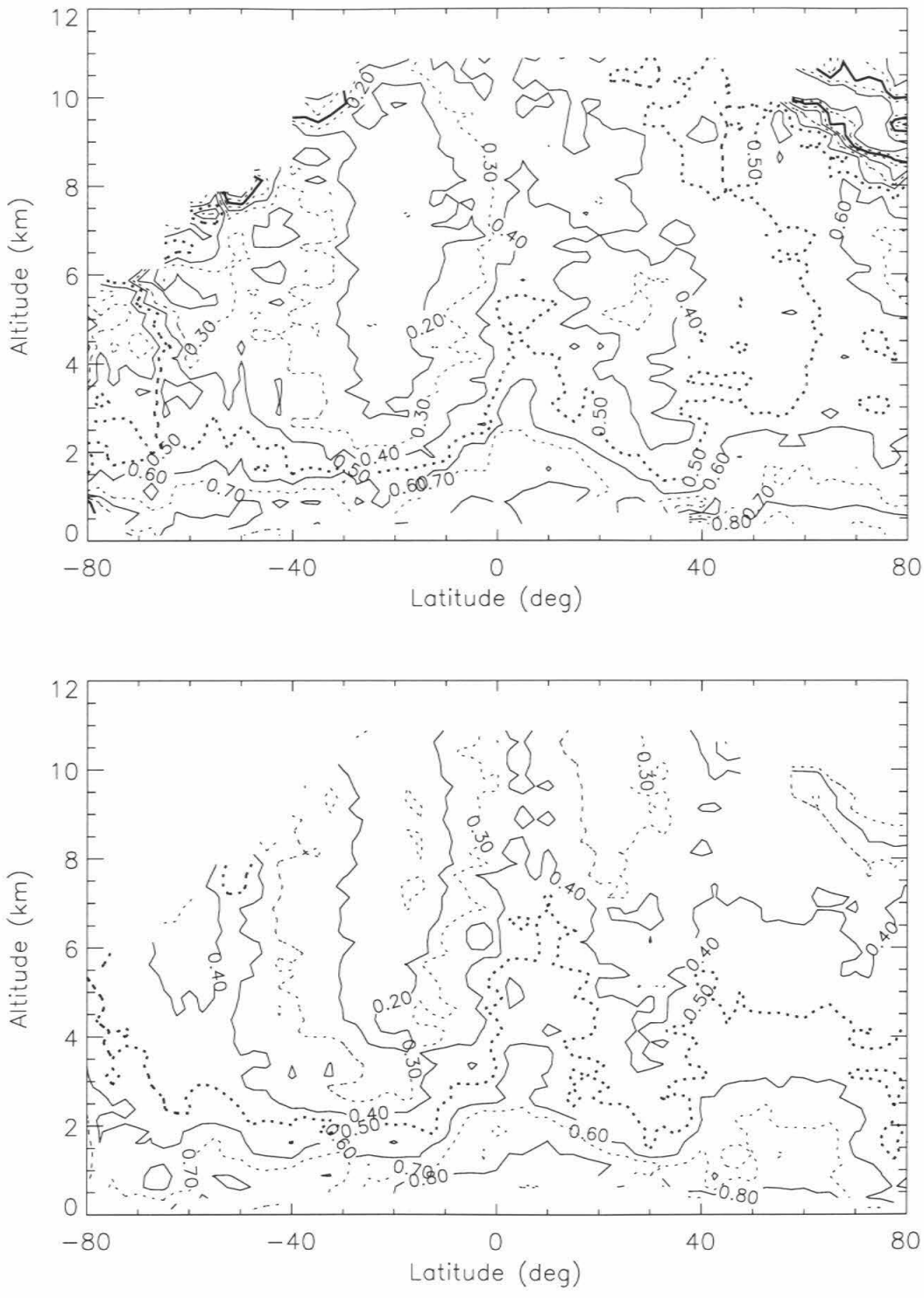

Figure 7.24 Mean relative humidity over liquid from a. GEW and b. IEC.

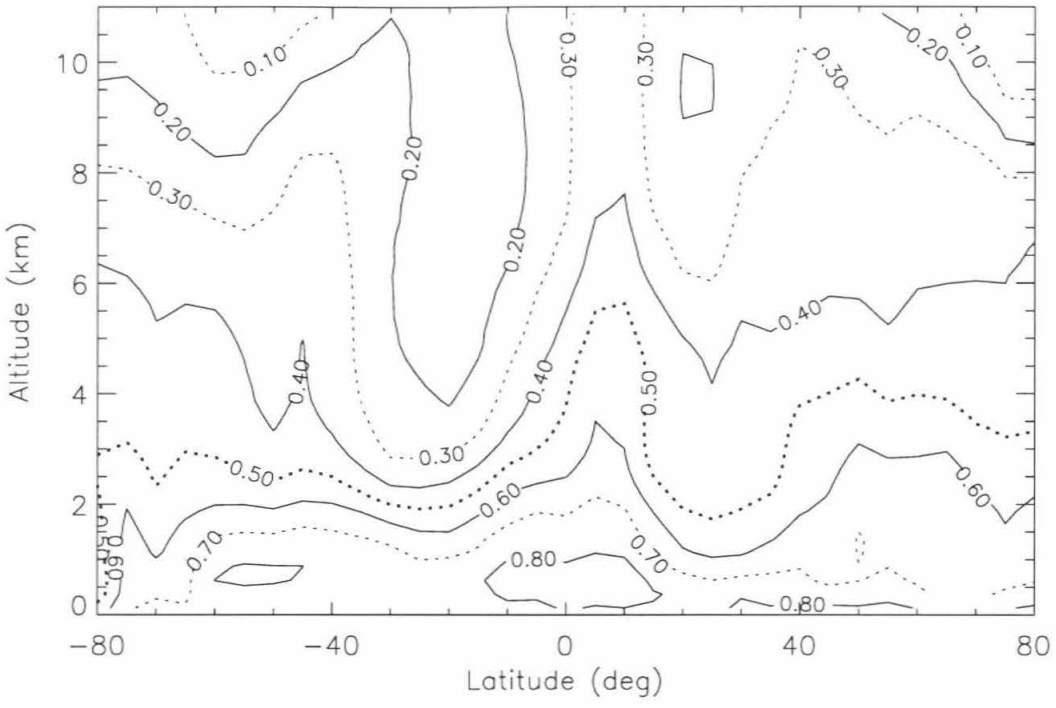

Figure 7.24 c. Mean relative humidity over liquid from CEC.

$$\bar{U} = \langle P_w \rangle / \langle P_w^* \rangle \qquad (7.5.4)$$

rather than the mean of eq. (7.3.1) because eq. (7.5.4) is better behaved.

The upwelling zone of the Hadley cell at the ITCZ is well defined as a maximum in the GEW $\bar{U}$ contours similar to the IEC and CEC results (Figure 7.24). Since meridional temperature variations are small within 30° of the equator, the meridional $\bar{U}$ structure is due almost entirely to moisture. The latitudinal definition of $\bar{U}$ is sharper than the P&O JJA climatology possibly due to smearing of climatological behavior over the much longer period of the P&O data. Any sampling bias in the occultation data is not too large because the CEC data exhibits a similar latitudinal definition. Since the ITCZ over land follows the sun which is at its northern-most location in the sky during this period and the ITCZ over the oceans is generally found on the northern side of the equator [Waliser and Gautier, 1993], the ITCZ is perhaps best defined latitudinally during the June and July interval.

The downwelling zones of the Hadley circulation are apparent as regions of minimum $\bar{U}_{GEW}$ on either side of the ITCZ. Asymmetry in the downwelling zones is apparent as minimum $\bar{U}_{GEW}$ in the southern hemisphere (<20%) is much less than that in the northern descending branch (<40%). The asymmetry in $\bar{U}_{GEW}$ is consistent with the climatological asymmetry in the Hadley circulation according to P&O with a far stronger downward circulation in the winter hemisphere than in the summer hemisphere (figure 7.24). Minimum relative humidities in the southern hemisphere exist over a broad region spanning roughly 700 mb to 300mb vertically and perhaps 15 degrees of latitude centered at ~20S. The northern and southern hemisphere tradewind inversions in the $\bar{U}_{GEW}$ are vertically sharper and occur at lower altitudes than in the $\bar{U}_{CEC}$ results such that column integrated moisture at these latitudes will be larger in the ECMWF results.

Reconstruction of the minimum relative humidities has stirred some debate and is relevant to understanding the control of water over a large fraction of the globe. As previously mentioned radiosonde humidities less than 20% are untrustworthy and climatological databases often assign a minimum value of 20%. SAGE, a solar occultation instrument, has made observations of water vapor above 5 km which show relative humidities near 10%, far drier than radiosonde-based climatologies. However, because SAGE requires very clear conditions, SAGE humidities set a lower climatological bound. According to Soden and Bretherton [1996] [SB96 hereafter], minimum relative humidity in the 400 to 800 mb interval during the June 21- July 4 period is between 20 and 25% between 15S and 20S. The driest period spans the end of July through the beginning of August. Minimum $\bar{U}$ on the north side of the equator in SB96 in the 400-800 mb interval are ~40% between about 25 and 35N.

The GEW $\bar{U}$ in the southern hemisphere descending branch (Figure 7.25) fall between the sonde and SAGE-based climatologies. GEW $\bar{U}$ indicate that most of the descending zone has relative humidity less than 18% with a minimum between 10 and 12% in a small region centered at 22.5S and 550 mb. The southern edge is near 30S at the northern edge of the baroclinic zone where decreasing temperatures cause relative humidity to increase. The latitude band of $\bar{U} \leq 20\%$ is close to that of CEC. Occultation $\bar{U}$ are smaller at lower altitudes and larger up high. The bias up high is due to the pressure bias discussed in section 7.5.4 probably indicating the GEW values are too high. The bias at lower altitudes may be real reflecting better vertical resolution in the GPS results.

The minimum $\bar{U}_{\text{GEW}}$ and $\bar{U}_{\text{CEC}}$ are both significantly drier than SB96 which is also derived from TOVS data. The difference may represent variability in the climate. The mean of the middle troposphere TOVS channel (400-800 mb) in this zone varies with season from a high of 35-40% in February to a low of 20-25% during the July-August

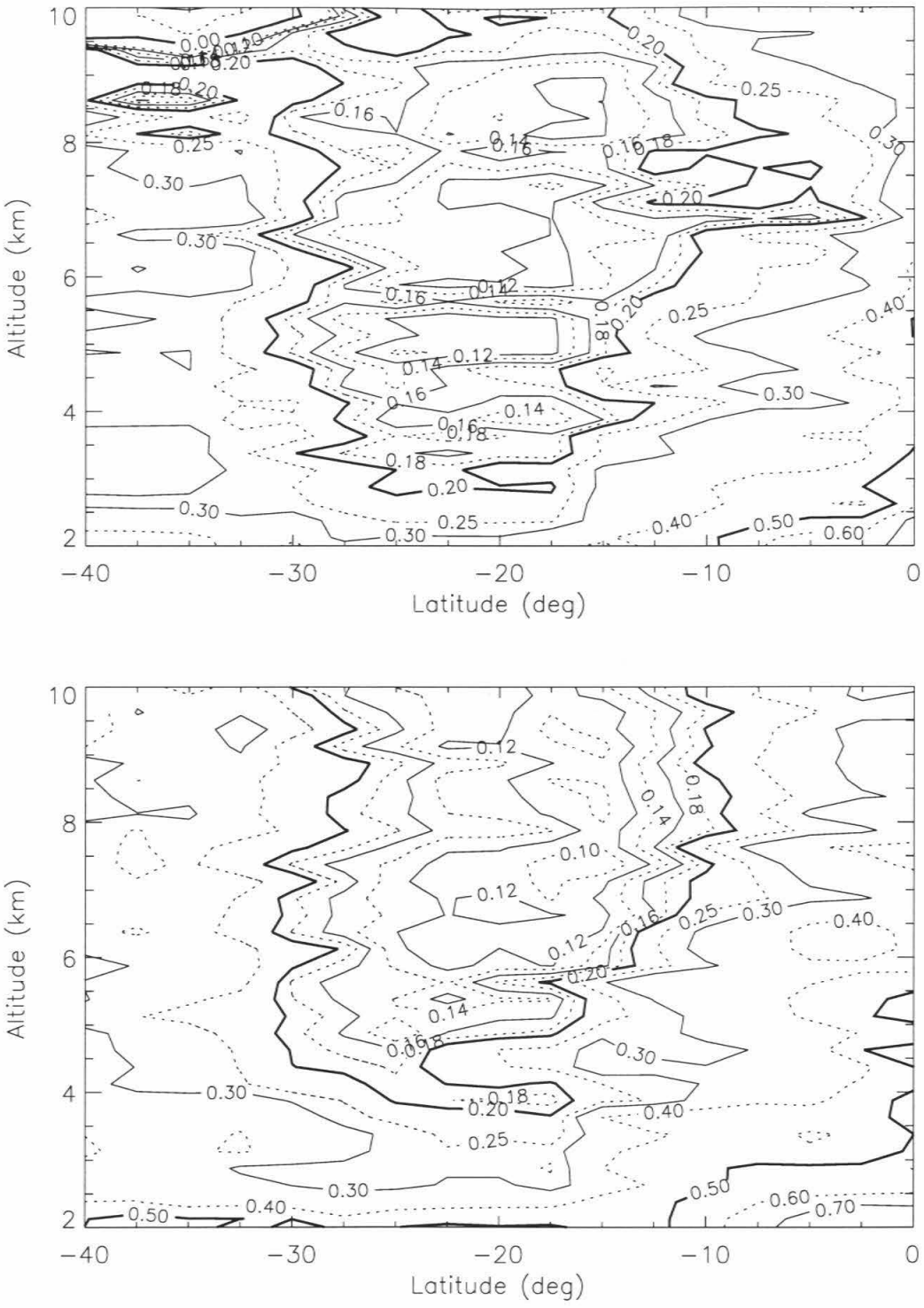

Figure 7.25 Mean relative humidity over liquid in the southern descending branch of the Hadley circulation a. GEW and b. IEC.

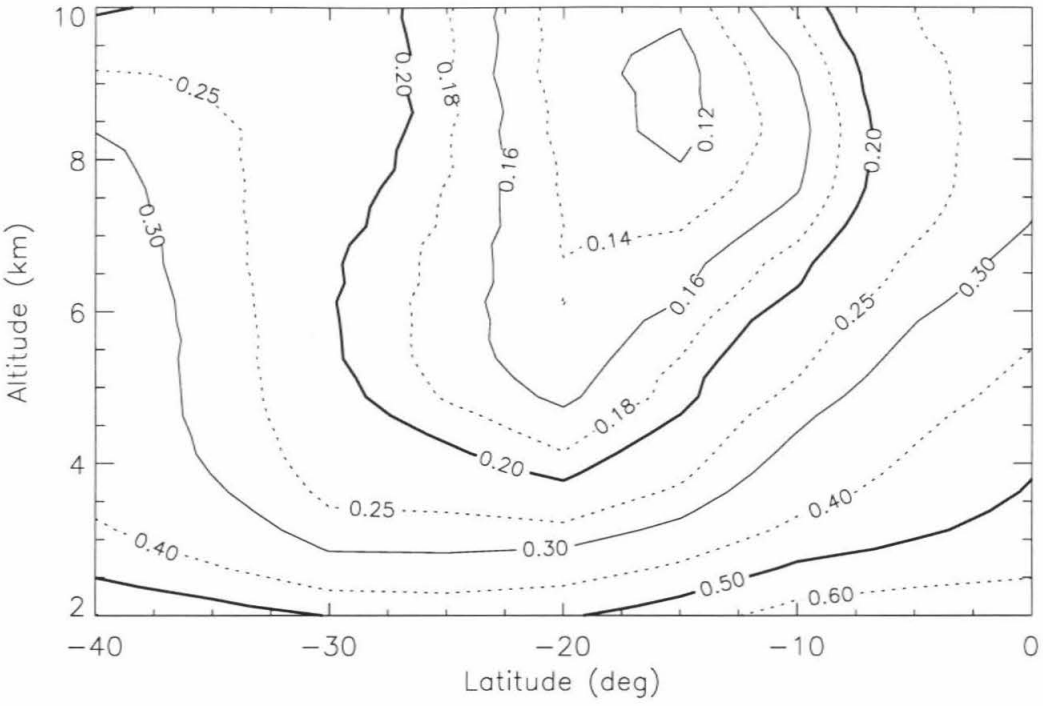

Figure 7.25 c. Mean relative humidity over liquid in the southern descending branch of the Hadley circulation from CEC.

period. They do not give an indication of the variability of mid-troposphere during the June-July period.

Northern Hemisphere biases In the northern hemisphere $\bar{U}_{GEW}$ are substantially higher than $\bar{U}_{CEC}$ above 6 km altitude due largely to a bias between the occultation and ECMWF pressures at the 230 K altitude (Section 7.5.4). Near the $\bar{q}$ = 1.4 and 2.8 g/kg contours, where the contribution of the temperature error to relative humidity cancels approximately, the agreement is much better with the GEW values being slightly higher. GEW relative humidities in the northern descending branch of the Hadley circulation near 30 N are drier than CEC values between 1 and 4 km altitude.

Biases at high southern latitudes South of 50S above 2 km altitude, $\bar{U}_{GEW}$ is so noisy that correction of the ECMWF biases are needed before further interpretation can be made. Below 2 km, GEW results are similar but slightly moister.

Variability

Figure 7.26 shows the standard deviation of relative humidity. GEW results are similar to those of IEC and CEC but somewhat larger in magnitude. At high southern latitudes, the impact of large ECMWF temperature errors is apparent as the variability becomes so large as to be unphysical.

7.5.7 Accuracy of ECMWF temperature implied from U and q

Noisy temperatures when $\bar{U} < 0$ and $\bar{q} > 0$

Near 55S at 400 mb, $\bar{U}$ is negative over a region whose extent is significantly larger than that where $\bar{q}$ is negative. $\bar{U}$ can be negative when $\bar{q}$ is positive because temperature errors affect the numerator (derived vapor pressure) and the denominator (the saturated vapor pressure) in the same direction. A negative temperature error reduces the

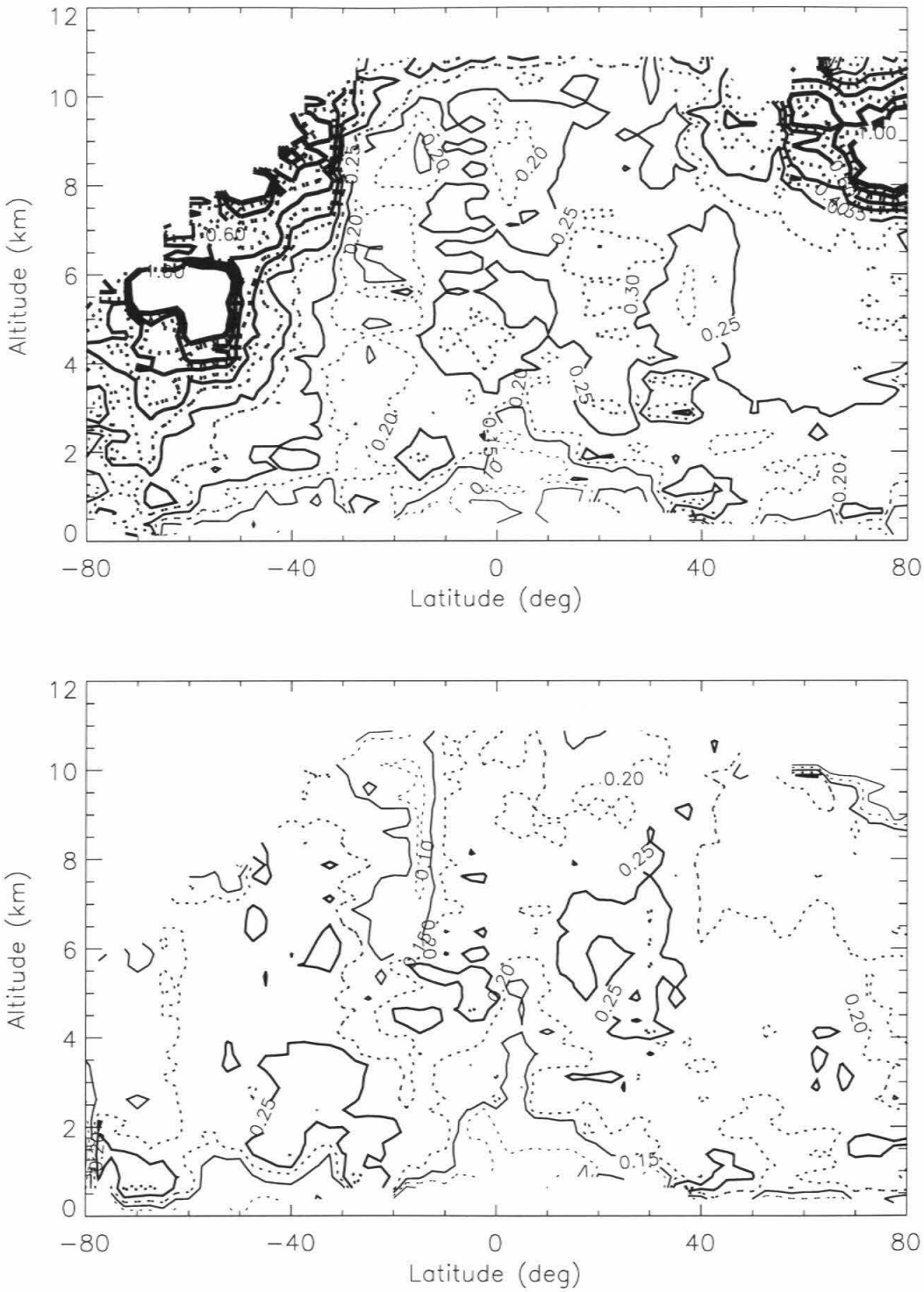

Figure 7.26 Standard deviation of relative humidity over liquid from a. GEW and b. IEC.

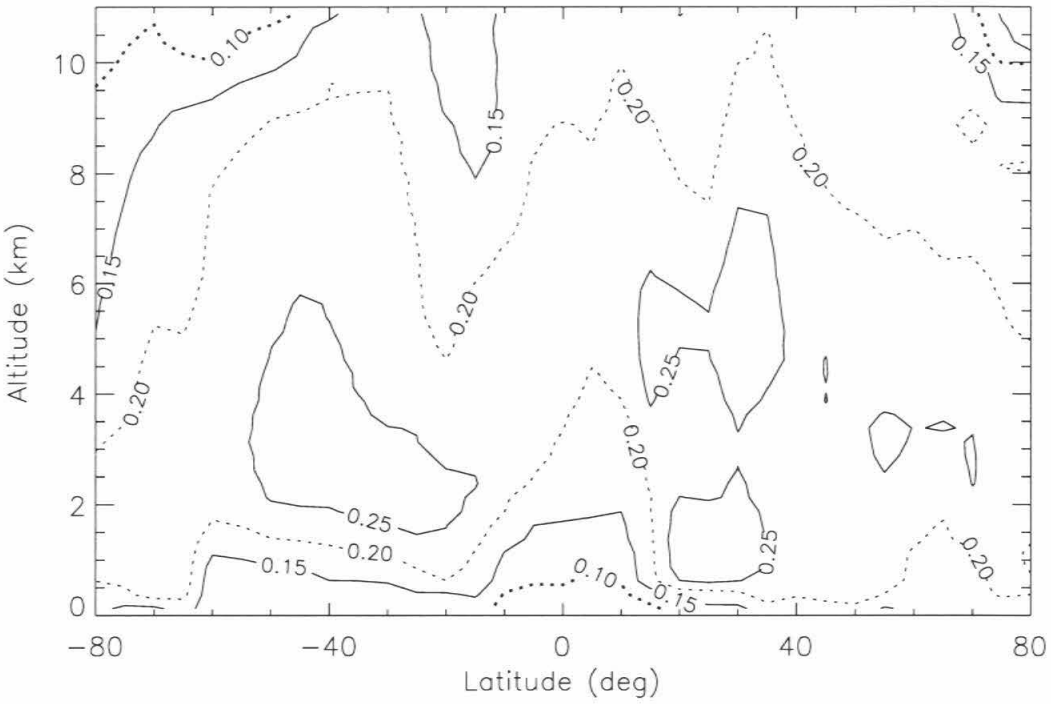

Figure 7.26 c Standard deviation of relative humidity over liquid from CEC.

denominator increasing the weighting of negative errors in the numerator whereas positive temperature errors increase the denominator reducing the weighting of positive errors in the numerator resulting in a shift of $\bar{U}$ toward negative values leading to negative $\bar{U}$ when the range of temperature errors is sufficiently large.

As mentioned, under sufficiently noisy conditions,

$$\bar{U}_2 = \langle P_w \rangle / \langle P_w^* \rangle \tag{7.5.4}$$

is a better estimate than the mean of Eq. (7.3.1). The mean relative humidity estimated as the mean of Eq. (7.3.1), which we will denote as $\bar{U}_1$, is sensitive to and can therefore provide a measure of the mean square error of ECMWF temperatures.

Consider the terms contributing to $\bar{U}_1$ estimated via eq. (7.3.1). From eq. (4.2.1), the estimate of P_w ($\equiv \tilde{P}_w$) is the true P_w plus error

$$\tilde{P}_w = P_w\left(1 + \frac{\delta P_w}{P_w}\right) \cong P_w\left(1 + B\left[\frac{\delta N}{N} - \frac{\delta P}{P} + \frac{\delta T}{T}\right]\right) \tag{7.5.5}$$

where we have assumed that $B >> 1$ in the regions of interest to simplify eq. (7.5.5). δN is the error in the retrieved refractivity, δT is the error in the ECMWF temperature and δP is the error in the hydrostatic pressure estimate. At the altitude where the water vapor retrieval begins, δP is due entirely to refractivity error at higher altitudes. At lower altitudes, δP accumulates error from errors in both refractivity and the ECMWF temperature.

Now to express the saturation vapor pressure, P_w^*, we will write the ECMWF temperature as the true temperature, T, plus an error, δT. The true temperature can be written as a mean temperature plus a perturbation, T', as can the temperature error such that

$$\tilde{T} = T + \delta T = \bar{T} + T' + \overline{\delta T} + \delta T' = \bar{\bar{T}} + T' + \delta T'$$

where overbar denotes a mean and ' denotes perturbations relative to the mean. From the Clausius-Clapeyron equation,

$$P_w^* \;\alpha\; \exp\left[-\frac{0.622\,L}{R_d\,T}\right]$$

Therefore,

$$\tilde{P}_w^* = c\exp\left[-\frac{0.622\,L}{R_d\,\tilde{T}}\right] = c\exp\left[-\frac{0.622\,L}{R_d\,\tilde{\bar{T}}\left(1+\dfrac{\left(T'+\delta T'\right)}{\tilde{\bar{T}}}\right)}\right]$$

$$\cong c\exp\left[-\frac{0.622\,L}{R_d\,\tilde{\bar{T}}}\left(1-\frac{\left(T'+\delta T'\right)}{\tilde{\bar{T}}}\right)\right] = c\exp\left[-\frac{0.622\,L}{R_d\,\tilde{\bar{T}}}\right]\exp\left[+\frac{0.622\,L}{R_d}\left(\frac{T'+\delta T}{\tilde{\bar{T}}^2}\right)\right]$$

$$\tilde{P}_w^* \cong \tilde{P}_{w0}^*\exp\left[+\frac{0.622\,L}{R_d}\left(\frac{T'+\delta T'}{\tilde{\bar{T}}^2}\right)\right]$$

where $\tilde{P}_{w0}^*$ is the saturation vapor pressure at the mean of the estimated temperatures in the region, $\tilde{\bar{T}}$. Therefore $\bar{U}_1$ can be written as

$$\bar{U}_1 = \overline{\frac{P_w}{\tilde{P}_{w0}^*}\left[1+B\left(\frac{\delta N}{N}-\frac{\delta P}{P}+\frac{\delta T}{T}\right)\right]\exp\left[-C_1\left(T'+\delta T'\right)\right]}$$

where $C_1 = \dfrac{0.622\,L}{R_d\,\tilde{\bar{T}}^2}$. Expanding the temperature errors in the exponential yields

$$\bar{U}_1 = \overline{\frac{P_w}{\tilde{P}_{w0}^*}\left[1+B\left(\frac{\delta N}{N}-\frac{\delta P}{P}+\frac{\delta T}{T}\right)\right]\exp\left[-C_1T'\right]\left[1-C_1\delta T'+\frac{C_1^2}{2}\delta T'^2+\ldots.\right]}$$

Now separating the terms to isolate the various contributions yields

$$\bar{U}_1 = \overline{\frac{P_w}{\tilde{P}_{w0}^*}\exp\left[-C_1T'\right]} + \overline{\frac{P_w}{\tilde{P}_{w0}^*}B\left(\frac{\delta N}{N}-\frac{\delta P}{P}+\frac{\delta T}{T}\right)\exp\left[-C_1T'\right]}$$

$$+ \overline{\frac{P_w}{\tilde{P}^*_{w0}} \exp\left[-C_1 T'\right]\left[-C_1\delta T' + \frac{C_1^2}{2}\delta T'^2 + ...\right]}$$

$$+ \overline{\frac{P_w}{\tilde{P}^*_{w0}} \exp\left[-C_1 T'\right] B\left(\frac{\delta N}{N} - \frac{\delta P}{P} + \frac{\delta T}{T}\right)\left[-C_1\delta T' + \frac{C_1^2}{2}\delta T'^2 + ...\right]} \qquad (7.5.6)$$

Assuming no correlation between the variations in the ECMWF temperature error and the true temperature variations and the water vapor variations, Eq. (7.5.6) can be simplified to

$$\bar{U}_1 = \overline{\frac{P_w}{\tilde{P}^*_{w0}}\left[1 - C_1 T' + \frac{C_1^2}{2}T'^2 + ...\right]} + \overline{\frac{P_w}{\tilde{P}^*_{w0}}\left[1 - C_1 T' + \frac{C_1^2}{2}T'^2 + ...\right] B\left(\frac{\delta N}{N} - \frac{\delta P}{P} + \frac{\delta T}{T}\right)}$$

$$+ \frac{\overline{P_w}}{\tilde{P}^*_{w0}}\left[\frac{C_1^2}{2}\overline{\delta T'^2} + ...\right] - \frac{\overline{P_w}}{\tilde{P}^*_{w0}} B\, C_1 \frac{\overline{\delta T'^2}}{T} + \overline{\frac{P_w}{\tilde{P}^*_{w0}} B\left[C_1^2 \frac{T'\,\delta T'^2}{T} + ...\right]} \qquad (7.5.7)$$

The first term on the RHS of Eq. (7.5.6) is the true $\bar{U}$ and must be positive. The second term is the contribution of <q> which can be negative. However, we are looking for terms which can make $\bar{U} < 0$ when $\bar{q} > 0$. The first two subterms in term 3 in eq. (7.5.6) are zero when there is no correlation between ECMWF temperature errors and the true temperature and water vapor variations and are not in eq. (7.5.7). The remaining subterm in term 3 in eq. (7.5.6) is small and always positive and so is not the term of interest. The fourth term in eq. (7.5.7) is negative and is the term of interest. The fifth term is smaller than the fourth term. From the fourth term, a crude estimate of the magnitude of $\overline{\delta T'^2}$ can be made. Because $\overline{P_w}/\tilde{P}^*_{w0} \sim \bar{U}$, $\overline{\delta T'^2}$ B C_1/T must be ≥ 1 in order to make $\bar{U}_1 < 0$. Assuming T~ 230 K and $\bar{q}$ is 0.1 g/kg, then C_1 is ~0.1 K^{-1} and B is ~300, such that $\overline{\delta T'^2} \geq$ T/(B C_1) = 7.7 and $(\overline{\delta T'^2})^{1/2} \geq 2.8$ K. This estimate is a lower bound on the temperature error and in fact is only about 50% larger than the standard deviation of the difference between southern hemisphere ECMWF and occultation derived temperatures at this

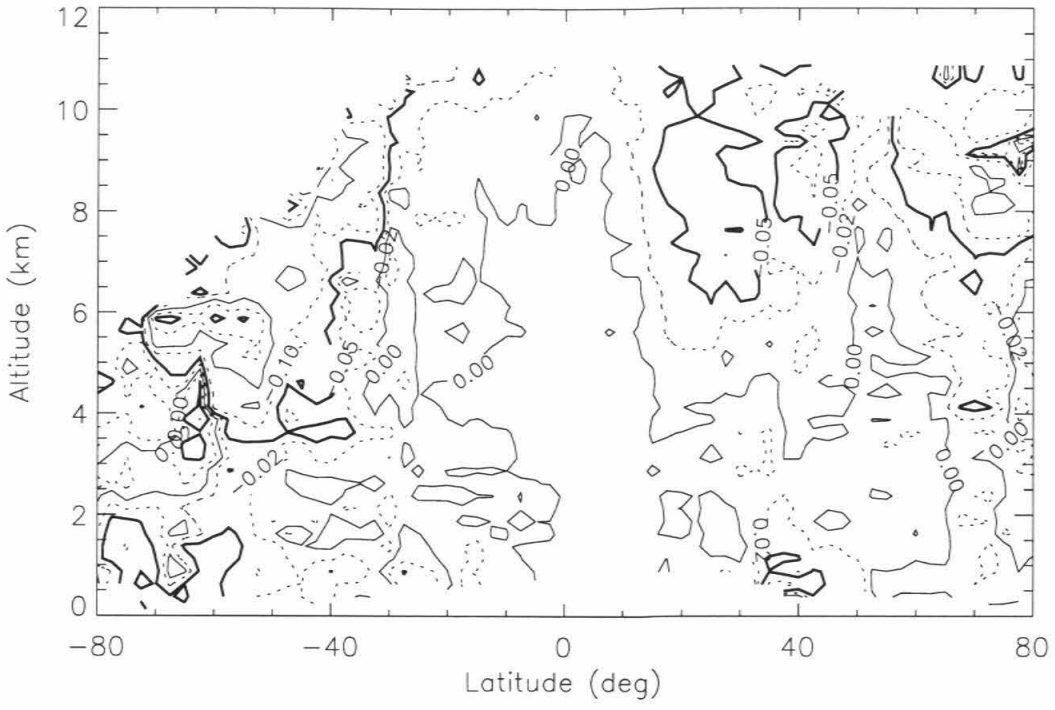

Figure 7.27 Difference between GEW mean relative humidity estimated as $\overline{P_w / P_w^*}$ and as $\overline{P_w} / P_w^*(\bar{T})$.

pressure level (figure 6.3). It is an underestimate because we have ignored the contribution of the fifth term.

Figure 7.27 shows the difference between the $\bar{U}_1$ and $\bar{U}_2$ and indicates the negative bias problem in $\bar{U}_1$ is confined mostly to the region above the 600 mb level between 70S and 40S. Most regions are consistent to 2% or better.

7.5.8 Equivalent Potential Temperature

Equivalent potential temperature, θ_E, is a measure of entropy which includes the contribution of the latent heat of water in air parcels. Its vertical gradient reflects the conditional stability of the atmosphere, that is its tendency to overturn. It is conditional because it depends on whether the water vapor reaches condensation. It is very important in the Hadley circulation and convection in general. Here we combine ECMWF temperatures with GEW water vapor data to compute a latitude versus height snapshot of θ_E for the June 21 - July 1995 period.

Description of $<\theta_E>$

Equivalent potential temperature is the final temperature a parcel of air attains when lifted dry adiabatically to its lifting condensation level and then pseudo-wet adiabatically (with respect to water saturation) to a great height (dropping out condensed water as it forms) and then finally brought to the 1000 mb level along a dry adiabat. The expression for θ_E used here is

$$\theta_E = T\left(\frac{1000}{P}\right)^{0.2854(1-0.28q)} \exp\left[\left(\frac{3.376}{T_L} - 0.00254\right) q\left(1 + 0.81\, q\right)\right]$$

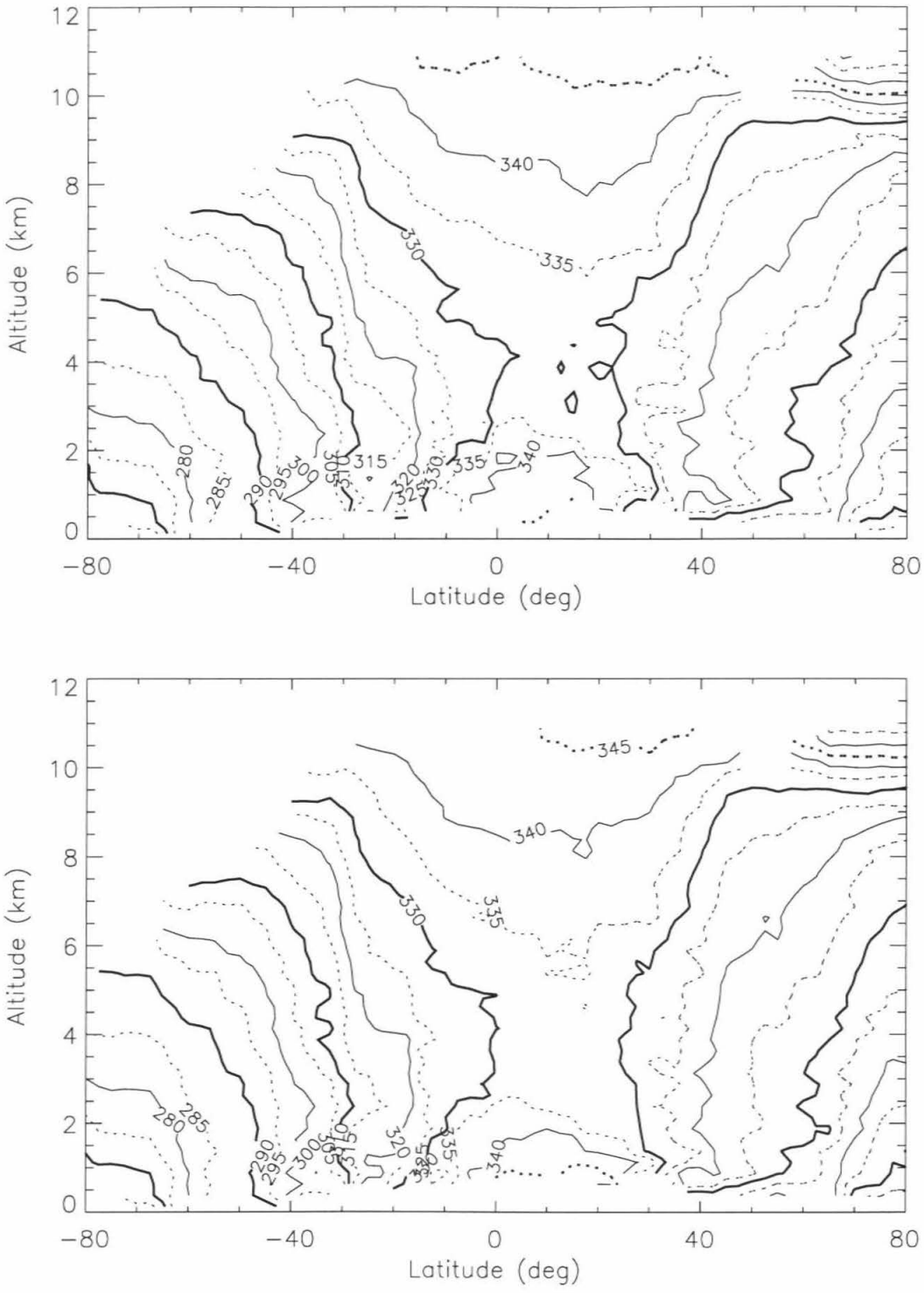

Figure 7.28 Mean of equivalent potential temperature estimated from a. GEW and b. IEC (K).

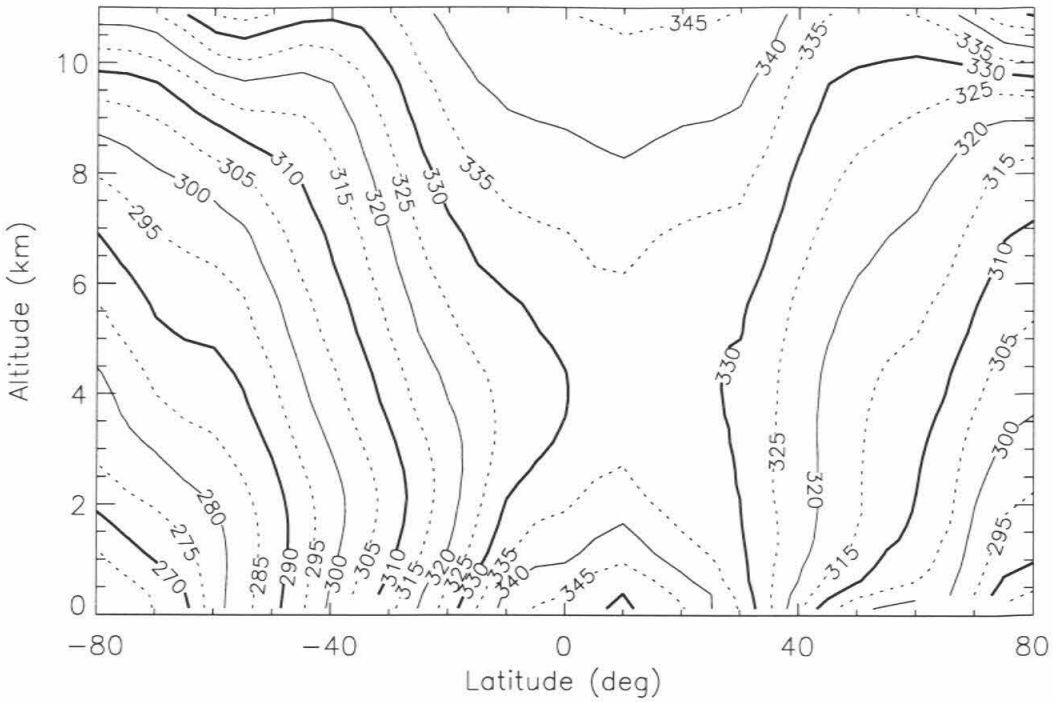

Figure 7.28 c. Mean of equivalent potential temperature estimated from CEC (K).

from Bolton [1980] where q is unitless and T_L is the temperature in Kelvin which a parcel would obtain if lifted adiabatically to its condensation level defined as

$$T_L = \frac{2840}{3.5 \ln T - \ln P_w - 4.805} + 55$$

also from Bolton (1980).

The largest values of θ_E in the lower troposphere are greater than 350K near the tropical surface although one must be careful of low sampling densities at altitudes below the 850 mb level between 5S and 35N (figure 7.28). At high northern latitudes, the minimum surface value is 285K whereas at high southerly latitudes it is less than 265K.

The vertical gradient of θ_E defines the conditional stability of the atmosphere. The vertical gradient of θ_E is negative near the ITCZ up to about 600 mb where θ_E is < 330 K. The altitude where θ_E = 350K in the upper troposphere is above the 250 mb level, and is completely defined by the ECMWF temperature structure. The negative vertical gradients and corresponding instability are of course due to the latent heat of water vapor in the lower troposphere and the associated instability drives the Hadley cell. Between 10S and 30S, the atmosphere is unstable up to ~ 800 mb above which it is stable. The lowest 2 km of the atmosphere are approximately neutrally stable between 30S and 60S. Stability increases at higher latitudes in both hemispheres. In the northern hemisphere at 30N, the atmosphere between 900 mb and 600 mb is unstable. At lower altitudes, the GEW climatology is probably unreliable due to a low sampling density there. The high stability of the stratosphere is apparent at high northern latitudes near 250 mb where the potential temperature gradient increases dramatically across the tropopause into the lower stratosphere.

Comparison with P&O

P&O gives an contour plot for θ_E for annual mean conditions whereas our mini-climatology covers just the short period from end of June through beginning of July, 1995. The minimum θ_E at tropical latitudes is 330 K at ~650 mb, very close to the GEW result. According to P&O, the vertical gradient of θ_E is negative at the surface at all latitudes which does not agree with the GEW results. GEW values are also warmer in the north and colder in the south at high latitudes than P&O because P&O is an annual climatology.

7.6 Discussion of Water Vapor Results

7.6.1 Introduction

In this section we will make some preliminary interpretation of gross features of the water vapor results. Ultimately my goal is to characterize the present behavior of water vapor and develop an understanding of the mechanisms controlling its distribution to the point that behavior may be extrapolated to a changing climate. While the GPS data is almost assuredly insufficient to reach these goals, it is the beginning of a longer process and will hopefully yield some new insight on the state and control of moisture in Earth's atmosphere. Concerning the quality of the data set, based on comparison of the IEC and CEC results, the GEW data has captured the general latitude versus height behavior of the atmosphere with a few exceptions for the June 21-July 4, 1995, period. Another point is that although it is quite variable, the lifetime of atmospheric water vapor based on global average atmospheric content divided by the global average precipitation rate is of order 10 days. Therefore, over the 2 week span of the GEW snapshot, atmospheric water vapor has recycled ~once on average globally. Having established the data as representative at least in the meridional versus height, the particular topics we will discuss are the length scales of water vapor, moisture balance in the descending branch of the Hadley circulation, the signature of the monsoons and Walker circulation near 30 N and a possible moisture source near 30S.

7.6.2 Vertical and meridional scales of $\bar{q}$

On a latitude vs. height map, contours of constant $\bar{q}$ appear to first order as a skewed pyramid with a peak at the ITCZ (figure 7.19) giving a significantly different impression than the more rounded latitude versus height $\bar{q}$ structure of P&O. The pyramid structure suggests a source at the ITCZ which maintains a maximum atmospheric moisture level balanced against meridional diffusional transport to downward sinks at higher latitudes which define moisture levels throughout the atmosphere. The locations of secondary peaks near 30S and 30N latitude are consistent with the CEC data and indicate possible other source regions.

Since q decreases approximately exponentially with height, the approximately linear decrease in the altitude of constant $\bar{q}$ contours with increasing latitude away from the ITCZ imply that, at constant altitude, q decreases exponentially with latitude away from the ITCZ peak. The latitude-height dependence of $\bar{q}$ can be described crudely by a double exponential of the form

$$\bar{q}(y,z) = q_0 \exp\left[-z / H_w\right] \exp\left[-\left[y - y_0\right] / L_w\right] \qquad (7.6.1)$$

using 5 free parameters: q_0, the surface specific humidity at the ITCZ, y_0, the average latitude of the ITCZ, H_w, the vertical scale height of q and two values for L_w, one for the meridional length scale in each hemisphere. Figure 7.29 shows the $\bar{q}_{CEC}$ structure with the eq. (7.6.1) model removed. q_0 is 22 g/kg, H_w is 2.5 km, y_0 is 9N and L_{wS} and L_{wN}, the southern and northern hemisphere length scales, are 2700 and 5000 km respectively. The similarity in vertical temperature gradient allows a single H_w to be used although it is clear that H_w is changing with latitude. The latitudinal length scale, L_w, is larger than that predicted using the CC relation and the baroclinic latitudinal temperature gradient. For instance in the southern (northern) hemisphere, L_w is ~2700 (5000) km while L_{CC} is substantially smaller at ~1200 (3800) km. L_w is larger in the northern hemisphere as the

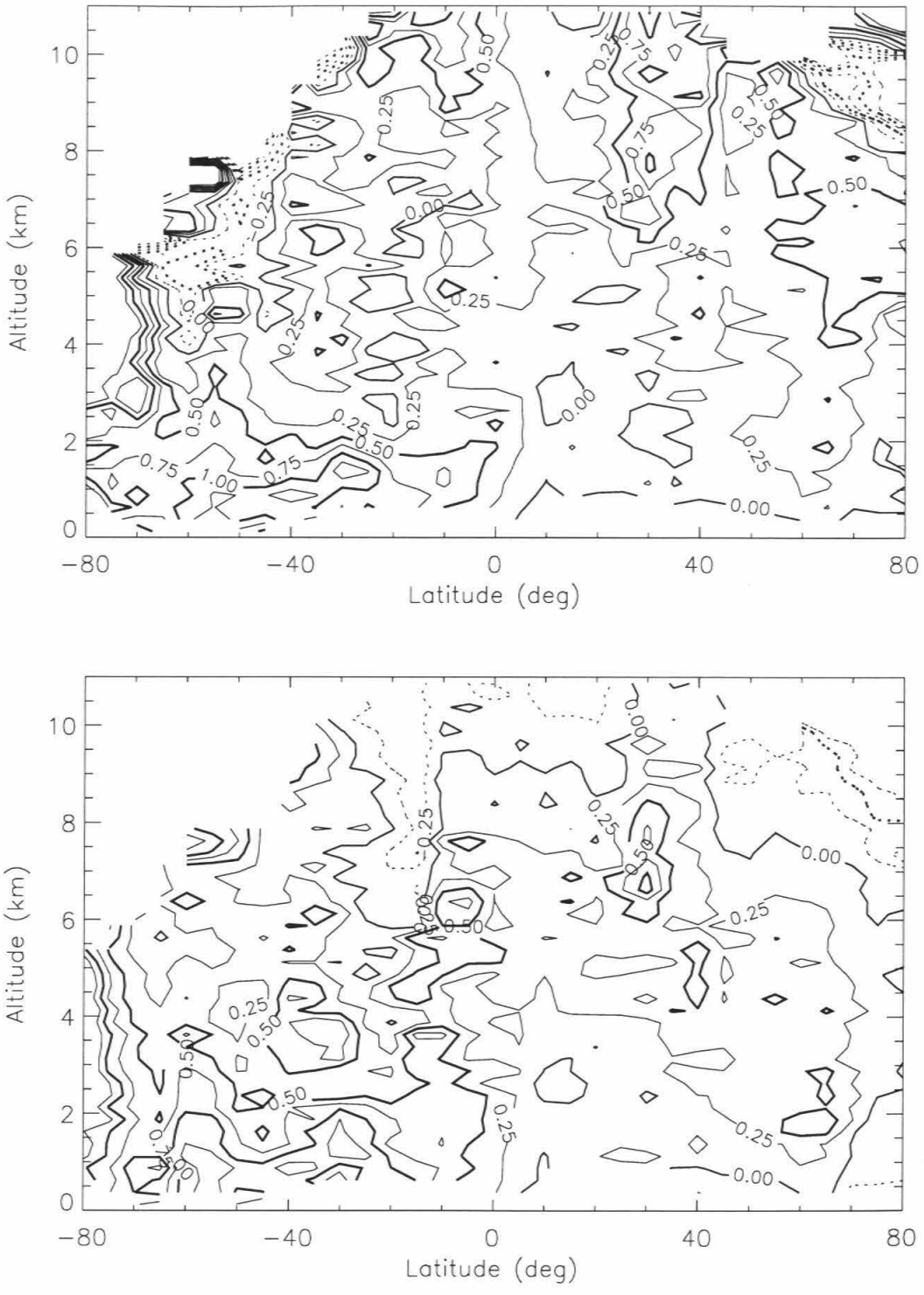

Figure 7.29 Difference between mean specific humidity and meridional-vertical exponential model normalized by mean specific humidity. a. GEW b. IEC See text for details.

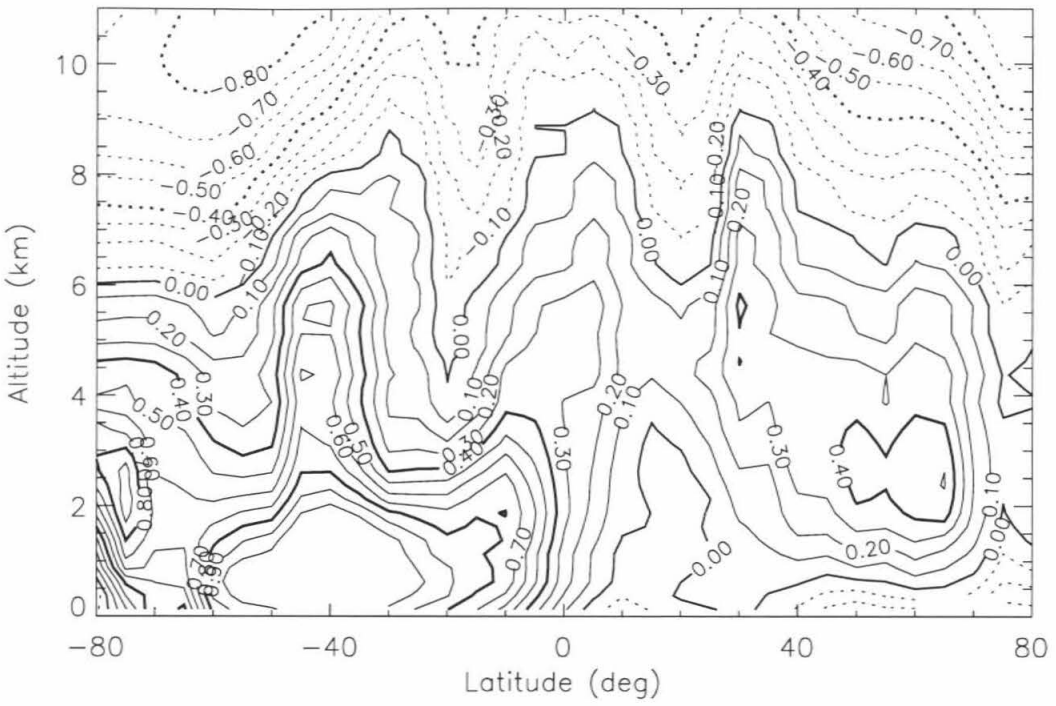

Figure 7.29 c. Difference between mean specific humidity from CEC and meridional-vertical exponential model normalized by mean specific humidity. See text for details.

larger meridional temperature gradient in the southern hemisphere suggests indicating that these horizontal length scales are undoubtedly coupled to temperature. Nonetheless eq. (7.6.1) is simpler than the horizontal temperature structure raising the question as to the mechanisms controlling the latitudinal length scales. The exponential behavior is suggestive of simple diffusion from the ITCZ to higher latitudes.

The very simple double exponential representation of q is accurate to about a factor of 2. The overall positive bias below 7 km altitude indicates the true q scale height is larger than 2.5 km whereas at higher altitudes, the true scale height is smaller than 2.5 km The largest residual meridional structure is the dry zones near 20N and 20S apparently associated with the Hadley circulation and relative maxima at 30N and 40S (shifting to 30S at higher altitudes). The maximum near 30N is associated with monsoonal behavior and the Walker circulation (Section 7.6.4). Interestingly, this structure is not apparent in the relative humidity (figure 7.24). The maximum near 40S is within the ~30S to ~60S baroclinic zone and is apparent in the relative humidity. It is clear that moisture near the surface in the southern hemisphere decreases more slowly toward the pole than the 2700 km scale length whereas above 6 km the 2700 km scale length is too long in comparison with real structure. The structure between 40S and 70S may be indicative of an upwelling moisture source near 40S and a downward sink between 50S and 60S with a possible upwelling source over Antarctica. The weak maximum between 60N and 65N may be an upwelling source.

7.6.3 Dryness of southern descending branch of the Hadley cell

The lowest relative humidities in the middle troposphere exist roughly between 10S and 30S between 200 and 800 mb coinciding with the southern descending branch of the meridional Hadley circulation. The Hadley circulation pattern according to P&O is decidedly asymmetrical during JJA with air moving predominantly southward from the

region of upward flow divergence near 200 mb. The asymmetry in the observed relative humidity structure is consistent with the asymmetrical flow. We will focus on the southern descending branch because of its simplicity which improves our chances of gaining new insight into moisture behavior and control using a limited data set. In particular, we will explore whether simple meridional diffusion of moisture from the ITCZ into the descending Hadley branch is consistent with observed q and climatological vertical velocity in the region.

There are several key points:

The pattern of relative humidity is decidedly asymmetric about the ITCZ with very low southern hemisphere minimum relative humidity consistent with the climatologically asymmetric meridional flow of the Hadley cell.

The location of minimum U is near the region of climatologically most intense downwelling.

The increase in q at lower altitudes in the downwelling region implies moisture is supplied to the downwelling air as it descends.

A zone of high latitudinal gradient in both $\bar{q}_{GEW}$ and $\bar{q}_{CEC}$ coincides approximately with downwelling.

Approximate coincidence of downwelling with positive meridional curvature implies source supplying moisture to downwelling zone may be meridional diffusion.

Order of magnitude estimate for eddy diffusivity indicates large scale eddies (~500 km) crudely consistent with satellite water vapor imagery.

The downwelling required in a diffusive balance is inferred and lies ~5^{o} south of the location of the most intense climatological downwelling.

The misalignment can be checked with ECMWF vertical winds for the period to determine whether the climatology is not representative or the diffusive model is too simplistic.

Observed dryness

The IEC relative humidity for the entire period is below 20% above 4 km altitude between 5S and 25S. The meridional structure of the region of U < 20% overlap pretty well in the GEW, and CEC data with the GEW region extends about a 1 km closer to the surface. Both ECMWF and GEW results indicate air is extremely dry, significantly below the 20 to 25% climatological minima of the largely radiosonde based P&O JJA climatology and the TOVS-based SB96 climatology for the June-July period. Both the ECMWF and GEW results show a southward migration of minimum U as the air descends.

Implications for moisture flow

Air rising near the ITCZ warms due to latent heat release and cools adiabatically which removes moisture by condensation and precipitation. While a small amount of condensed moisture is carried aloft, air is extremely dry near 200 mb as it begins its descent. In a simple Hadley circulation, q of the descending air would remain constant as the air descends leading to enormous latitudinal gradients in q at lower altitudes. During JJA, the climatological southern descending branch of the Hadley cell is defined by downward flow extending from between 5 and 10S to beyond 30S with the most intense downward flow between 8 and 20S peaking near 12S (P&O). In the GEW, IEC and CEC data, q in air which has descended the vertical extent of the GEW 20% relative humidity contour, has increased from 0.2 to 2 g/kg, an order of magnitude, requiring a supply of moisture to the air as it descends.

The steady state moisture budget in the meridional plane can be written as

$$0 = \frac{\partial\left[\overline{w}\overline{q_t}\right]}{\partial P} + \frac{\partial\left[\overline{v}\overline{q_t}\right]}{\partial y} = \frac{\partial\left[\overline{w}\overline{q_t}\right]}{\partial P} + \frac{\partial\left[\overline{w'q_t'}\right]}{\partial P} + \frac{\partial\left[\overline{v}\overline{q_t}\right]}{\partial y} + \frac{\partial\left[\overline{v'q_t'}\right]}{\partial y} \qquad (7.6.2)$$

where q_t is the total moisture mixing ratio including condensed moisture, w is vertical velocity in mb/s, v is the meridional velocity and $\bar{x}$ and x' refer to the mean and varying

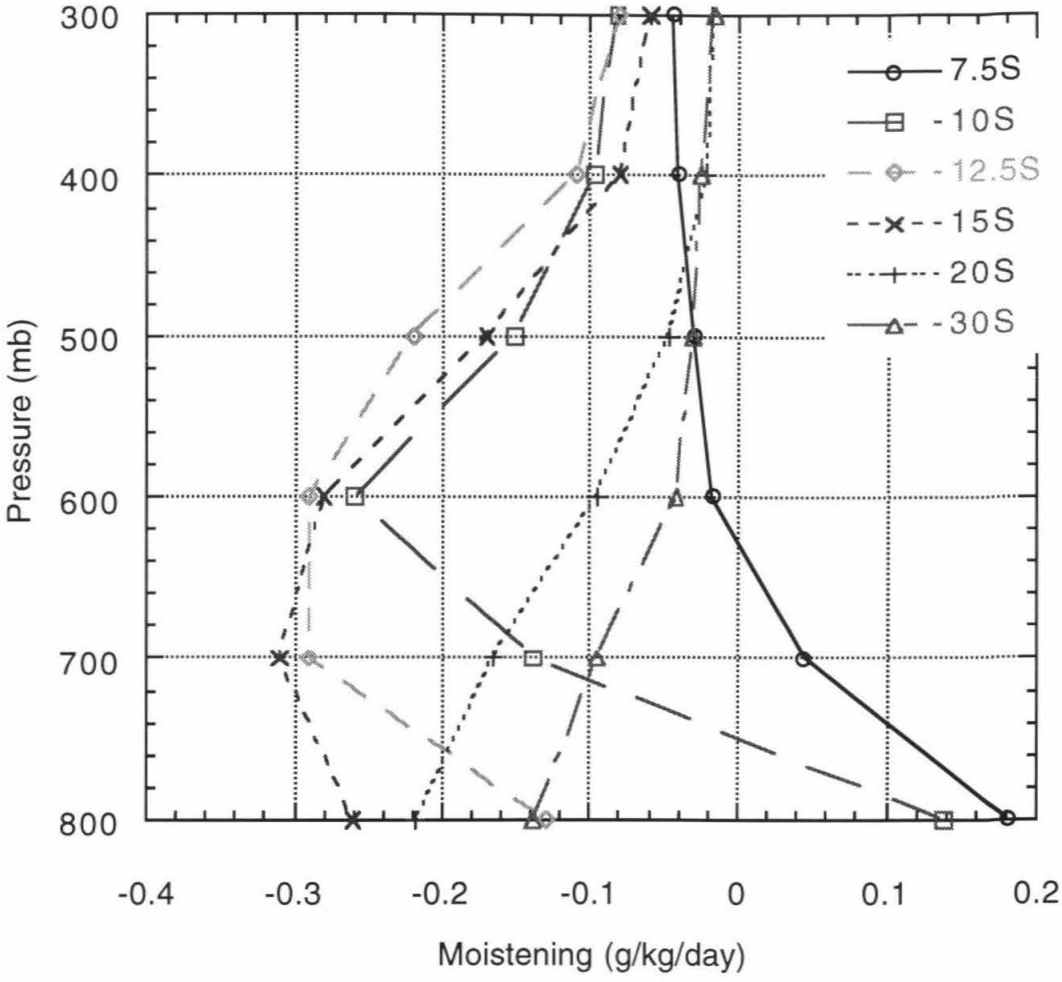

Figure 7.30 Zonal mean cross sections of moistening in the southern descending branch of the tropical Hadlley circulation inferred using eq. 7.6.3 and $\bar{q}$ from GEW and vertical velocity for JJA from P&O.

part of x. In the downwelling region, the $\bar{q}$ from GEW, IEC or CEC can be combined with the climatological $\bar{w}$ from P&O to estimate drying due to the divergent vertical water vapor flux in the downwelling regions. Under these conditions and assumptions, the divergence of the downwelling moisture flux is

$$\frac{\partial[\bar{q}\,\bar{w}]}{\partial P} = \bar{q}\,\frac{\partial[\bar{w}]}{\partial P} + \bar{w}\,\frac{\partial[\bar{q}]}{\partial P} \qquad (7.6.3)$$

where w is vertical velocity in mb/s. Figure 7.30 shows the drying associated with descending air between 7.5S and 30S. Maximum drying of order -0.3 g/kg/day occurs between 600 and 700 mb between 10S and 15S where $\bar{q}$ is ~2 to 4 g/kg indicating a decay time of ~10 days in the absence of a compensating source. The second term on the RHS of eq. (7.6.3) generally dominates the drying except near the surface.

In steady state, drying due to downward divergence must be balanced by a convergent moisture source. Four possible methods could supply moisture to the downwelling region: hygrometeors supplying moisture from above, convection bringing moisture from below, and meridional flow supplying moisture laterally either by mean flow or via eddies. In the Sun and Linden (1993) hygrometeor scenario, moisture is supplied to the descending air via condensed water phases falling and evaporating from above. The difference in downward moisture flux at 800 mb vs. 200 mb is ~ 6×10^{-4} versus 6×10^{-5} g/kg mb/s which must be balanced by a hygrometeor flux. This essentially implies that throughout the tropical and subtropical upper troposphere, there is an order of magnitude more ice (~2 g/kg) than vapor (~0.2 g/kg) implying the upper troposphere is continually saturated. While there is evidence via limb sounding of a very thin veil of cirrus in the tropical upper troposphere, it seems unlikely that the required mixing ratio of the Sun and Lindzen ice cloud which is 4 times the high liquid water content of optically thick marine stratus (0.5 g/kg) could be up there and not have been noticed. The Sun and Lindzen scenario further requires that this enormous amount of upper troposphere ice remain aloft

during its ~2000 km southward migration from the ITCZ to the region of descent and then fall so rapidly that it supply moisture via sublimation while leaving air unsaturated at higher altitudes.

Concerning convection, the capping tradewind inversion inhibits convection from penetrating into the free troposphere in general consistent with personal comments from N. Renno and Dave Raymond. The large decrease in relative humidity from the ITCZ to 25S is due to a decrease in q as temperature remains ~constant implying condensation does not control mixing ratios in the descending branch. Further, both the mean and standard deviations of relative humidity are ~20% in this region indicating descending region air seldom, if ever, reaches saturation. Therefore, convection appears unlikely as a source, although it cannot be ruled out.

Mean meridional flow is significant in the P&O Hadley circulation at the top (100 to 300 mb) and bottom (1000 to 800 mb). Between 200mb and 800 mb, where we are interested, mean meridional flow is insignificant according to P&O.

Given the unlikely nature of the other scenarios, the best candidate is a meridional eddy flux. Meridional moisture transfer from lower latitudes via eddies is suggested by the general decrease in q from the ITCZ zone to the south. Meridional eddy terms will carry moisture away from a relative maximum in $\bar{q}$. The ITCZ provides the best example. Since $\bar{q}$ decreases away from the ITCZ, air leaving the ITCZ has higher q on average than replacement air arriving from higher latitudes. Therefore, at low latitudes, v' is correlated with q' such that $\overline{q'v'}$ is < 0 and > 0 to the south and north of the ITCZ respectively. In steady state, the divergence of the meridional eddy flux must balance that of the vertical flux:

$$\frac{\partial\left[\bar{q}\,\bar{w}\right]}{\partial P} + \frac{\partial\left[\overline{q'v'}\right]}{\partial y} = 0 \qquad (7.6.4)$$

Using eq. (7.6.4), the vertical moisture flux divergence can be integrated meridionally to establish the total change in the meridional flux across the downwelling which establishes the minimum magnitude of the meridional flux at the equatorward edge of the downwelling region. For instance, near 600 mb where the flux divergence is ~-0.25 g/kg/day between 10S and 15S, the meridional flux at 10S near 600 mb must be at least 210 g/kg km/day. Unfortunately, one cannot compare this estimate with $\overline{qv}$ from P&O because the P&O results are given only for annual conditions which are not representative of June-July behavior because of the seasonal reversal of low latitude vertical velocity washes out in an annual mean. In the future I will get the ECMWF wind field analyses so I can directly determine whether eq. (7.6.4) is satisfied. For the moment we will use the information available to us and leave consistency checks using ECMWF derived winds for the future.

The diffusive flux concept: The net flux of moisture in the exchange of high q air with low q air may be represented as a downgradient diffusion process with a diffusivity directly related to the properties of the eddies responsible for the exchange and transport process. Diffusive moisture fluxes are related to the first and second meridional derivatives of the GEW moisture climatology. The flux is

$$F_{qy} = -\kappa_e \frac{\partial \bar{q}}{\partial y} = -v_e \lambda_e \frac{\partial \bar{q}}{\partial y} = \overline{q'v'} \qquad (7.6.5)$$

where κ_e is eddy diffusivity and v_e and λ_e are a characteristic velocity and length scale of the eddies. The flux divergence is

$$\frac{\partial F_{qy}}{\partial y} = \frac{\partial \overline{q'v'}}{\partial y} = -\kappa_e \frac{\partial^2 \bar{q}}{\partial y^2} - \frac{\partial \kappa_e}{\partial y} \frac{\partial \bar{q}}{\partial y} \qquad (7.6.6)$$

In a region where κ_e varies little, a positive meridional specific humidity curvature is a region of meridional convergence and moistening. To balance drying in the Hadley downwelling region, positive meridional curvature is expected and is indeed observed.

Order of magnitude estimates: Assuming eq. (7.6.5) holds and moistening is dominated by the second derivative term in eq. (7.6.6) yields

$$\frac{\partial F_{q\,P}}{\partial P} \sim \kappa_e \frac{\partial^2 \bar{q}}{\partial y^2} \sim \lambda_e v' \frac{\partial^2 \bar{q}}{\partial y^2} \tag{7.6.7}$$

Balancing a vertical moisture divergence drying rate of ~0.26 g/kg/day with a diffusive meridional moisture convergence using a large meridional second derivative of ~4×10^{-7} g/kg/km^2 representative of conditions near 15S and 500mb, yield a diffusivity of 2.6 km^2/s. With v' ~ 5 m/sec, based on the P&O day-to-day variations in meridional velocity in this region (P&O figure 7.20), yields a diffusion length scale of ~500 km, such that several eddies would span the ITCZ to downwelling zone. Satellite images do show large scale swirls in the water vapor through this region.

Sharp meridional gradient

Further, the focused downwelling region of P&O between 10S and 15S and associated divergence of downward moisture flux implies a sharp latitudinal moisture structure should exist in this vicinity if moisture is supplied meridionally.

The Hadley downwelling of P&O is very large between 10S and 15S causing a meridionally focused vertical moisture flux divergence which must be balanced by an equally focused meridional moisture flux convergence if eq. (7.6.4) holds. Therefore, the structure of $\bar{q}$ should exhibit a large meridional change near the zone of maximum downwelling. The close spacing of $\bar{q}$ contours in figure 7.31 in the 10S to 15S zone hint of high meridional gradients. Figure 7.31 shows the logarithmic meridional gradient of $\bar{q}$. Despite the noise, a systematically high meridional gradient of $\bar{q}_{\,GEW}$ is apparent in the 10S to 20S region (figure 7.32). The CEC data exhibits the gradient more clearly due to the better averaging with the region of high $d\bar{q}_{\,GEW}/dy$ lying slightly poleward of the maximum $d\bar{q}_{\,CEC}/dy$ (figure 7.32). The poleward bias of the GEW data may result from the limited

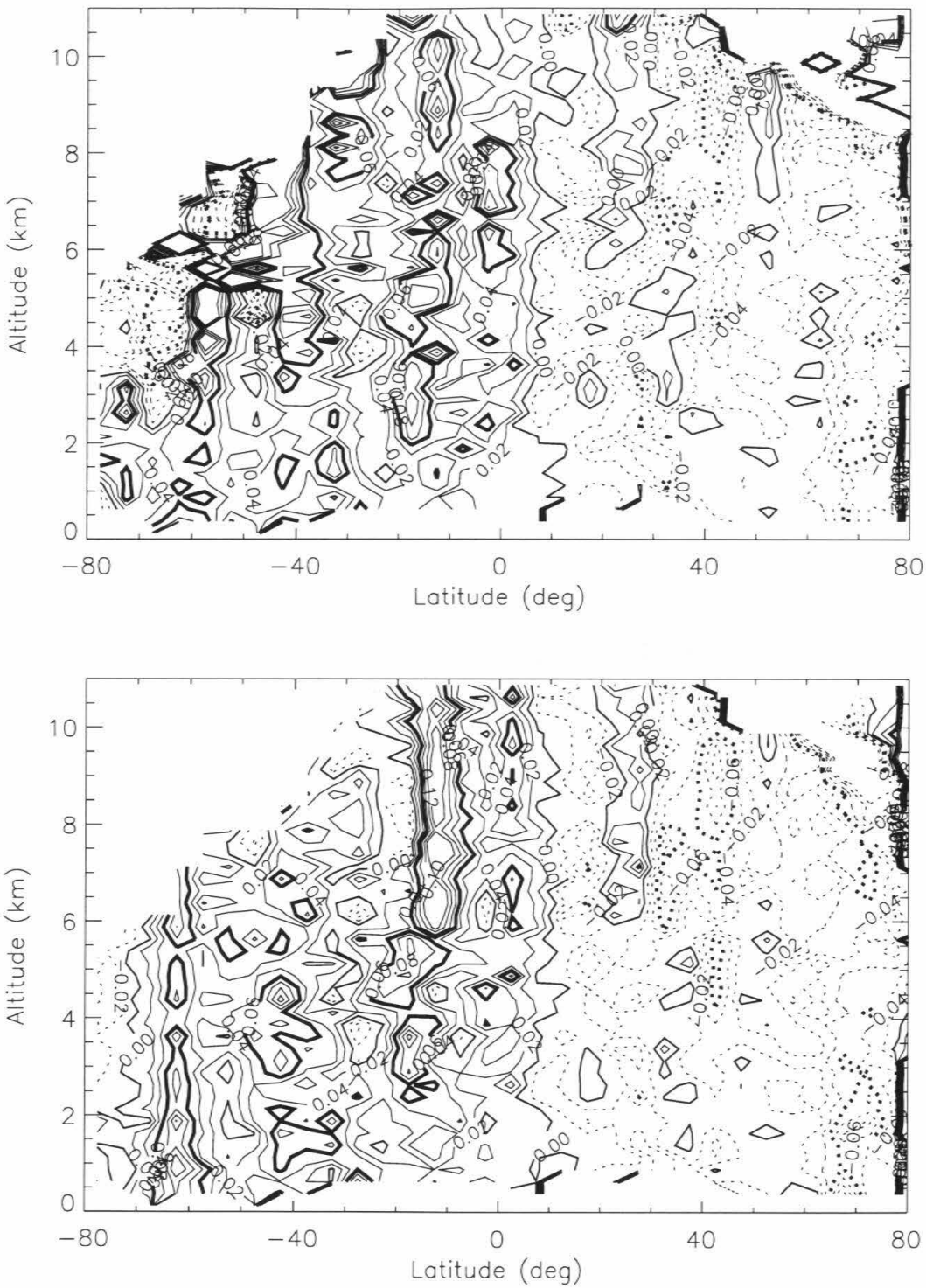

Figure 7.31 Normalized meridional gradient of specific humidity for a. GEW and b. IEC (degree^{-1}).

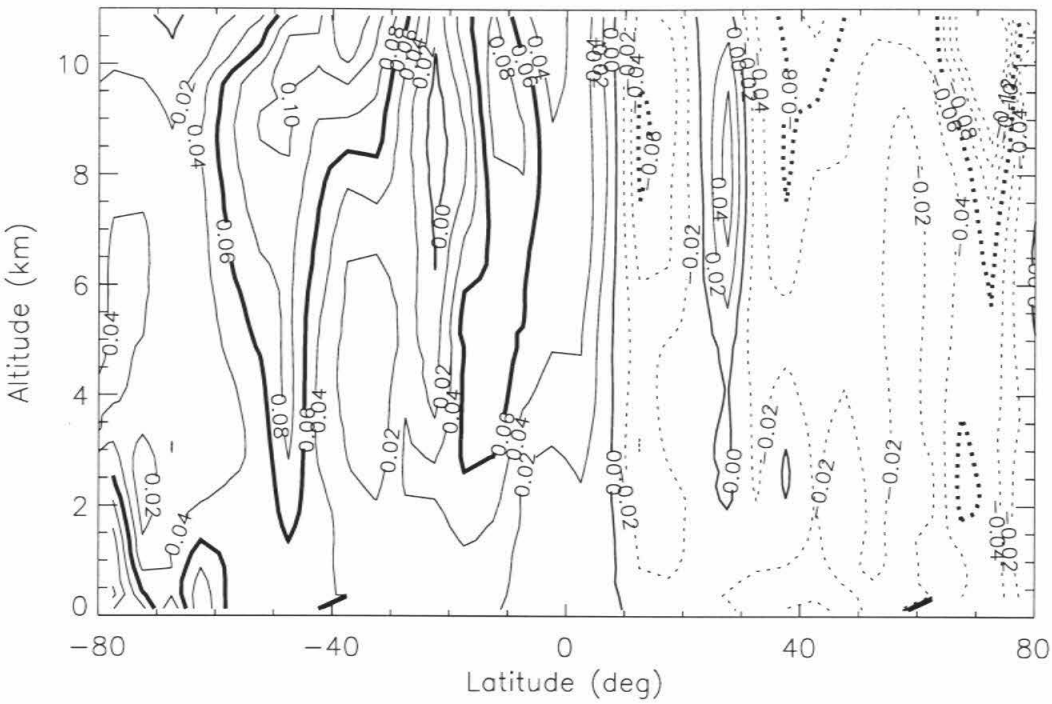

Figure 7.31 c. Meridional gradient of specific humidity for CEC (degree^{-1}).

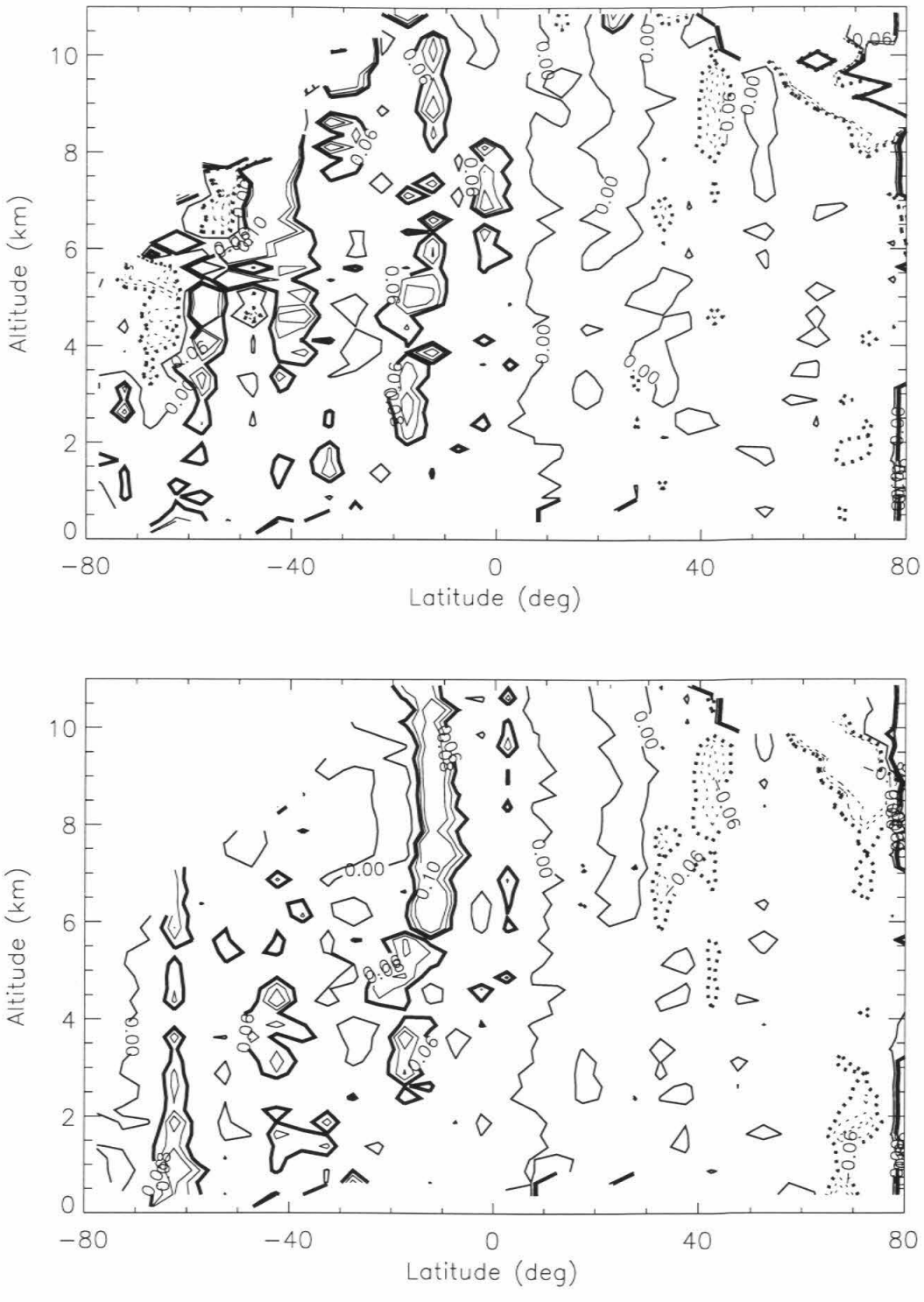

Figure 7.32 Meridional gradient of specific humidity greater than 0.06 per degree latitude for a. GEW and b. IEC (degree^{-1}).

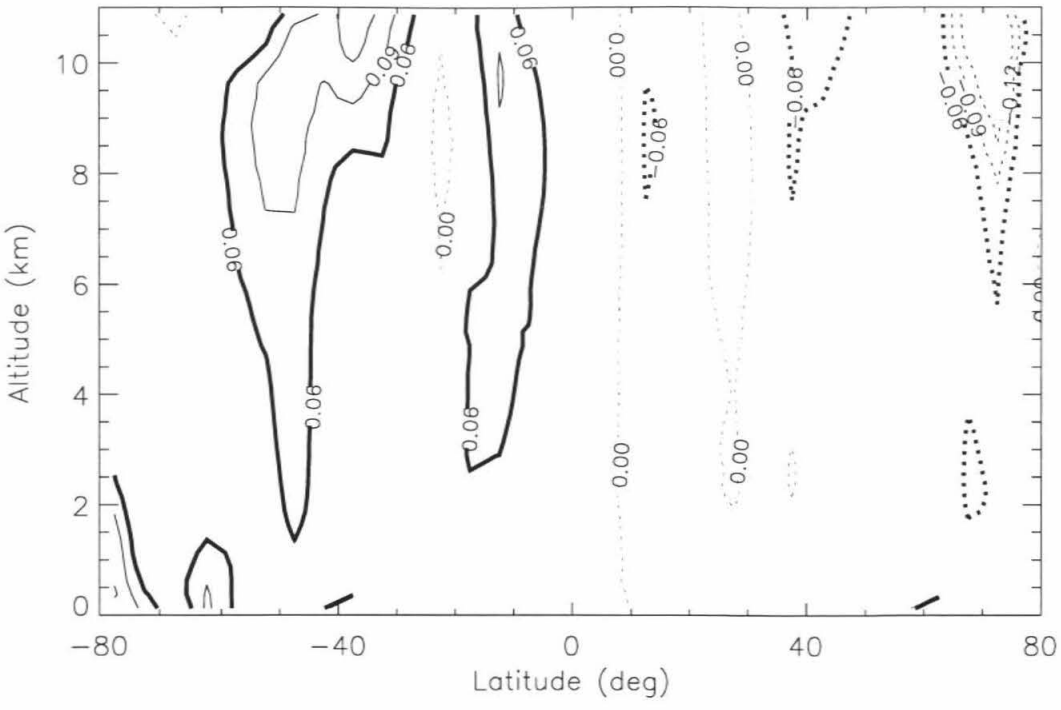

Figure 7.32 c. Meridional gradient of specific humidity greater than 0.06 per degree latitude for CEC (degree^{-1}).

sampling of the GEW since the region does move around over the 11 day period. The high $d\bar{q}/dy$ regions of GEW and CEC exhibit similar shifts in latitude versus height. High meridional gradients of $\bar{q}_{GEW}$ in regions south of 45S reflect the pressure and refractivity bias pattern there (Sections 7.5.4 and 7.5.5) and cannot be trusted. The P&O JJA $\bar{q}$ climatology reflects only a hint of the downward deflection of the $\bar{q}$ contours in the descending branch which may also reflect inadequate sampling.

Taking the diffusive model to its logical conclusion, we set the vertical moisture flux divergence equal to the meridional moisture flux convergence and solve for the vertical velocity assuming a constant diffusivity. We choose a diffusivity of 2.6 km^2/s which yields a peak inferred vertical velocity similar in magnitude to that of P&O. Since first and second derivatives must be taken, we use the CEC rather than GEW $\bar{q}$ field. The resulting inferred vertical velocity is shown in figure 7.33. The peak inferred vertical velocity lies distinctly to the south of the maximum vertical velocity of P&O. In fact, the maximum w of P&O aligns much better with the maximum $d\bar{q}/dy$ rather than the high q second derivative. The misalignment of vertical velocity fields either means the P&O velocity field is not representative of this period or the diffusive model is too simple. The ambiguity can be resolved in the near future by comparing the inferred vertical velocity with the ECMWF analysis vertical velocity which we do not presently have. If the ECMWF w agrees with that of the simple diffusion model, it gives strong weight to the diffusion concept. If the ECMWF vertical velocity structure more closely resembles the P&O w field, then a somewhat more complex model can be developed which has separates advection and diffusion.

The real process is a combination of advection and diffusion which have been mapped into a single diffusivity parameter in eq. (7.6.5). As a slightly more complex model, I have in mind a "leaky bucket" scenario where air parcels diffusively exchange

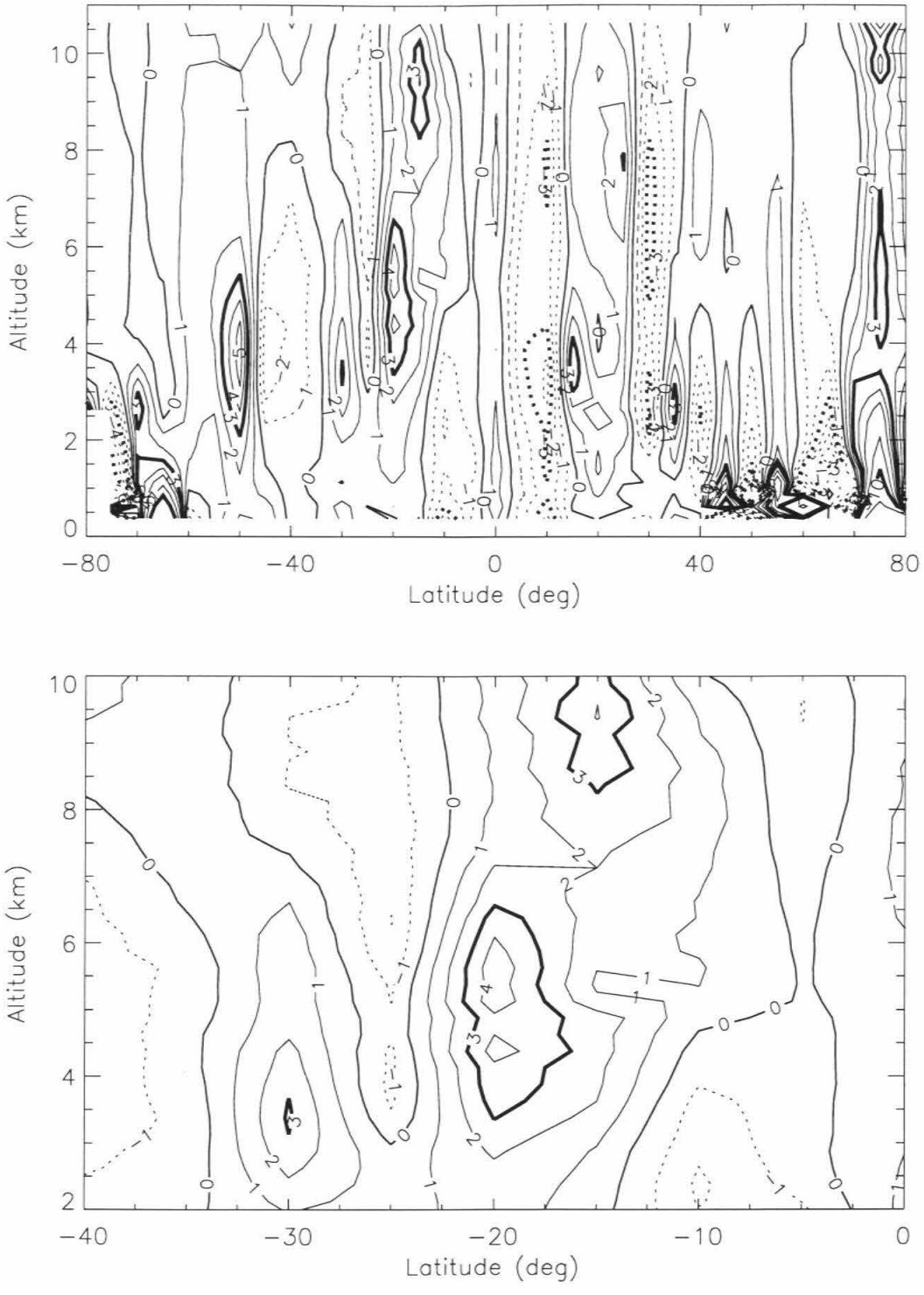

Figure 7.33 Vertical velocity inferred from a balance between vertical moisture divergence and diffusive meridional moisture convergence.
a. global and b. downwelling region.

moisture with background air through which they are advected. In this case diffusion is represented by following the simple differential equation

$$\dot{q}_p = -\frac{q_p - q_b}{\tau}$$

where q_p is q in an advected air parcel, $\dot{q}_p$ is its time rate of change, q_b is the q of background air through which the parcel is moving and τ is the diffusive exchange time constant. As a parcel from the equatorial region traverses the region of high meridional gradient, the parcel rapidly gives up moisture into the surrounding dry air at a rate which depends on τ and the meridional gradient of the background air times the meridional advection velocity. The location of the maximum flux can be adjusted by adjusting τ. This model will be further explored in the future.

In any case, the fact that a sharp meridional feature in $\bar{q}$ does exist at approximately the correct latitude indicates there is a significant change in meridional flux associated with the vertical moisture flux divergence indicative that meridional flow is balancing a large portion of the downward divergence drying.

Feedback concept

If the simple diffusive model of moisture is correct where moisture rises in the ITCZ and diffuses outward to higher latitudes where it is sinked by downward mean flow in the subsidence region with a small amount of moisture diffusing further to the south, we may be able to use the concept to gain some understanding of what happens when climate changes. According to the heat engine concept espoused by Renno and Ingersol (1996), an increase in surface temperatures resulting from increased greenhouse gases will increase the temperature difference between the absorption and emission altitudes of the atmosphere which will increase atmospheric circulation. A moisture feedback will exist if the eddy

velocities are tied to the convection circulation. I suspect that the eddy velocities are related to rate of subsidence via geostrophic balance and the subtropical high in surface pressure required to drive near-surface air into the tropics. If convection intensifies, the subsidence rate will increase causing the pressure contrast between the ITCZ and subsidence zones to increase, which increases eddy magnitudes increasing eddy diffusivity of moisture from the tropics to the subtropics. One interesting scenario is that the diffusivity increase might offset the increase in subsidence drying.

The feedback model must include coupling with radiative transfer. From a radiative perspective, all other things being equal, increased circulation gives descending air less time to cool and offset the rate of adiabatic warming during descent. It is not entirely clear what the temperatures will do because of meridional mixing between the rising and descending branches which is the basis of the diffusive moisture transport model. The mass of the column of air at the subtropics must be slightly higher than the column mass at the ITCZ to drive surface subtropical air toward the tropics. However, decreased cooling would decrease air density reducing the subtropical surface pressure high required to drive air back to the ITCZ. The dryness of the downwelling increases the mean molecular mass relative to the ITCZ upwelling which helps produce a surface high. Clearly a coupled model based on the moisture diffusion model will require some work to understand the implications of diffusive meridional moisture transport.

7.6.4 Peak in Mean q near 30N

$\bar{q}_{GEW}$, $\bar{q}_{IEC}$ and $\bar{q}_{CEC}$ all exhibit relative maxima near 30N above the 400 mb level. $\bar{q}_{GEW}$ at 300 mb at 30N is somewhat larger than at the ITCZ while the $\bar{q}_{IEC}$ and $\bar{q}_{IEC}$ are somewhat less. On the northern side of the ITCZ, a relative maximum in q variability exists between 20N and 30N extending from the surface to 250 mb (11 km) overlapping with the northern hemisphere region of minimum $\bar{U}$ above the 600 mb level. The standard deviation of temperature variations, σ_T, exhibit an interesting vertical structure with maxima

in the same region at high and low altitudes. These indicators of variability imply a perturbation transport mechanism may be operating here.

Monsoons and topography.

An examination of mid-troposphere moisture fields in the ECMWF analyses for the June 1995 period reveal a number of important features in this zonal band. Of interest here is the fact that the highest moisture levels in the mid-troposphere exist just south of the Himalayas reflecting the impact of the summer Indian Monsoon. Another region of substantial moisture exists near the west coast of the central America roughly delineating mountain ranges from Mexico to northern South America. In between these two moist regions is a distinctly dry region over the Pacific, presumably the result of large scale downwelling associated with the upwelling zones on either side. This moist-dry longitudinal structure defines the Walker circulation which is a basic feature of moisture climatological studies [e.g., Soden and Bretherton, 1996]. The observed relative maximum in $\bar{q}$ and the high variability in q near 30N (figures 7.19 and 7.229) apparently reflect the strong longitudinal dependence of the monsoons and Walker circulation. A future study examining the zonal structure of low latitude GEW data should be carried out. However, sampling over the June-July 1995 period is relatively sparse.

How robust are the $\bar{q}$ peaks at 30N? Sampling here is slightly more sparse than in latitudes to either side. The maxima of the 0.5 and 1.0 g/kg contours represent an increase of ~50% above a simple linearly decreasing trend of contour height with latitude. At 300 mb near 30N, B, based on the P&O JJA climatology, is ~100 so the estimated accuracy of $\bar{q}$ is ~20%. Using the extrapolation from the 100% error contour near 230 K, the error is 35%. At 400 mb near 30N, B, based on the P&O JJA climatology, is ~40 so the estimated accuracy of $\bar{q}$ is ~8%. Using the extrapolation from the 100% error contour near 230 K, the error is 9%. Therefore, the maxima, particularly the one at 400 mb, appear to be real.

The region of highest σ_q in the northern hemisphere is between 15N and 25N, roughly the center of the U minimum on the north side of the equator. This implies that on the north side of the equator the horizontal transport term may be transferring moisture right through the relative minimum in RH to points farther north (because the flux divergence is where material in the flux is deposited). Perhaps this transfer process is involved in filling the northern hemisphere with moisture as spring turns to summer there.

7.6.5 Peaks in q near 30S and baroclinic structure to the south

The position of relative maxima in $\bar{q}$ near 30S may be related to the transition between subtropical and winter baroclinic air masses. A significant difference exists between the IEC and GEW $\bar{q}$ contours near the region of high meridional gradient near 30S at 8 km and above. The IEC exhibits a relative maximum near 25S and minimum near 15S in the 0.1 and 0.2 g/kg contours. The CEC exhibits a much weaker maximum and minima in this area. While the GEW $\bar{q}$ contours exhibit somewhat similar bumps, they do not achieve the status of relative maxima and minima. The difference is significant because maxima imply an upper level moisture source near 25S diffusively supplying moisture to surrounding latitudes. According to P&O, the mean meridional flow is almost 0 at 30S so the source term is most likely a vertical one. The mean vertical flow of P&O at 30S is downward at 10^{-4} mb/s apparently drying this region of the atmosphere (figure 7.6.2) which leaves the vertical perturbation correlation term, $\overline{q'w'}$, as the implied moisture source. So, if the relative maxima are real, a convective source may be implicated. In marine boundary layers under regions of subsidence, conditional instability is usually large with θ_E generally significantly larger than in the overlying lower free troposphere, a situation which is made stable by the strong capping inversion. It may be that poleward of the strongest subsidence, the weaker capping inversion along with weak low level convergence due to subtropical high pressure air from the Hadley boundary layer encroaching on baroclinic boundary layer air allows the Convective Available Potential

Energy (CAPE) to be tapped. The altitude to which the relative maxima between 25S and 40S extend in figure 7.29c may be correlated with the height where θ_E equals either the mean surface θ_E or the positively perturbed values.

Alternatively, the ECMWF 0.1 and 0.2 g/kg contours near 25S may be incorrect. Although absolute GEW data accuracy is limited at 8 km altitudes, the difference between the ECMWF and GEW q structures may be due to vertical resolution between GPS and TOVS such that the vertical structure of water vapor derived from TOVS radiances has been misplaced vertically. If true, this suggests that GEW data can produce climatologically representative structure with fewer samples than TOVS. In either case, the southward decrease in q away from the ITCZ implies that only a small fraction of moisture lifted at the ITCZ reaches high latitudes via the free troposphere. Additional GPS data is required complete with an understanding of the pressure and refractivity biases discussed in Sections 7.5.2 and 7.5.4.

7.7 Summary and conclusions concerning tropospheric results

With a relatively simple receiver modification and processing system adjustments, the occultation profiles extend generally to 1 to 3 km altitude during the June 21 to July 4, 1995, period when the GPS AS encryption was turned off. Beginning with the launch of the next generation receivers at the end of 1997, still better performance is expected even when the AS encryption is enabled.

Given the global but sparse sampling of GPS-MET, a latitude versus height refractivity crosssection of the atmosphere was formed. This crosssection clearly demonstrated the contribution of water vapor to refractivity at lower latitudes. A strong summer/winter contrast is readily apparent in terms of q. Distribution of water vapor is

distinctly peaked at the ITCZ in terms of specific and relative humidity and the descending branches of the Hadley circulation are apparent as relative minima in relative humidity. The monsoonal Walker circulation appears as a relative maxima in the mean and the standard deviation of specific humidity near 30N. Weak maxima near 30S may imply a convective moisture source at that latitude.

Distinct differences exist relative to the P&O JJA climatology which are much greater than differences between GEW and ECMWF. Differences relative to the JJA P&O climatology may simply be natural variations over the times scales of weeks to months. On the other hand, the P&O data is very sparse in the southern hemisphere where humidities are significantly drier than the P&O JJA climatology. Specific humidities are more peaked near the ITCZ than in the P&O. GEW variability is similar to although somewhat larger than that of P&O.

Comparisons of individual water vapor profiles derived from occultation-derived refractivity, using ECMWF temperatures, has both confirmed ECMWF water vapor profiles as in the case of the Falkland Islands comparison as well as identified flaws in the analyses associated with insufficient vertical resolution such as the Madagascar comparison. Statistical comparison of the occultation refractivity with that derived from the ECMWF 6 hour global analyses indicates agreement between means generally consistent to within 1.5%. There are small systematic biases. Occultation refractivities are lower over a crescent shaped region centered about 800 mb and 10N below 8 km altitude. Although the source of the crescent bias is unclear, it appears to be due largely to the occultations because the combination of surface pressure, temperature and lapse rate biases which reproduce the observed bias structure requires a surface pressure bias large compared with present accuracy estimates. This bias carries through to the GEW humidities which are somewhat lower than their ECMWF counterparts in general.

A distinct tripole bias structure between the GEW and IEC refractivity and specific humidity exists near 800 mb. In the subtropics, the occultation data is drier than the IEC results whereas at the tropics the occultation data is decidedly moister. This latitudinal structure is consistent with that observed in columnar water comparisons between TOVS and SSM/I. The bias structure seems to be related to the Hadley circulation and, in particular, a systematic error in the height of the marine boundary layer presumably associated with the limited vertical resolution of the nadir viewing TOVS data.

Discrepancies exist at high southern latitudes which indicate ECMWF temperatures are noisy in this zone. Negative $\bar{q}$ near 50S and 350 mb reflect a cold bias in ECMWF temperatures while supersaturated $\bar{q}$ over Antarctica reflect a warm bias in ECMWF temperatures. Relative humidity in the region near 50S and 350mb is negative over a greater spatial extent than the region of negative $\bar{q}$ indicating the variation in ECMWF temperature errors is quite large in the region. Clearly the occultation data will substantially improve the dry and moist atmospheric structure in this zone when it is assimilated. For the immediate future the possibility of correcting the noisy ECMWF temperatures using the refractivity, pressure and temperature derived from the occultation profiles in the upper troposphere needs to be explored.

In the statistical comparison with ECMWF, a high altitude specific humidity bias of order 0.1 g/kg was revealed in the occultation results which is associated with a pressure bias of ~0.5% or 30 m in geopotential height between the occultation and ECMWF results north of 30S. More work is needed to separate the various error contributions by their signatures. For instance, occultation accuracy generally varies little with latitude to first order but more so with increasing water vapor. Therefore, sharp changes in agreement with latitude are more likely associated with ECMWF.

We have investigated control of moisture in the downwelling branch of the Hadley circulation, in particular a hypothesized balance between downward drying and meridional

eddy moistening. Given the focused climatological downwelling structure of the analysis of P&O, the existence of a sharp latitudinal gradient in q was predicted and, in fact, a large meridional gradient was found near 10S to 15S. A simple diffusional model was used to further predict the mean downward flow. It does not match the location of the P&O downwelling which either means the P&O results are not representative of this period or the model is too simple which is likely because the model in maps advection and diffusional processes into a single diffusivity constant. In fact, it would be remarkable if the control of moisture over such a large fraction of the Earth could be described by one such variable. Nonetheless, I suspect that a first order description of the water vapor in this region can be achieved with a few more degrees of freedom. If this can be achieved, then the climate change problem should attacked with this model. As mentioned, there is probably coupling between the eddy advection/diffusion and the strength of the Hadley circulation which must be understood. If this understanding is reached, then I believe a simple model can be implemented with some predictive power.

Development of simple first order model representations of moisture control mechanisms are necessary to understand the physical coupling. A low order meridional advection/diffusion model properly coupled to the strength of the Hadley circulation may go a long way toward making useful predictions for the distribution of moisture as climate changes. The one caveat in my prediction is condensed moisture because it has such an enormously nonlinear impact on radiative transfer. The amount of condensed moisture in the atmosphere is generally much less that the vapor and yet it can completely change the radiative properties of the atmosphere. The asymmetry between the meridional length scales of the northern and southern hemispheres clearly tied to higher meridional temperature gradients in the winter hemisphere. However, beyond that simple statement it gets tricky because the mean humidity is not at saturation. The atmosphere has a range of humidities which can locally reach saturation and the controlling mechanism in a simple model must be sufficiently correct to properly represent the ensemble of atmospheric

moisture states such that the model reaches saturation the right percentage of the time and further that it will continue to do so when the high latitudes warm and baroclinicity diminishes. This is not trivial, but I think it is doable. Unfortunately, one won't know if the prediction was correct until it actually happens (details....).

Concerning the near future, several specific intentions are suggested. First the ECMWF vertical wind field must be compared with the diffusion model prediction in the descending branch of the Hadley circulation. If they match, the water vapor control can be described via a remarkably simple mechanism. If, in the more likely scenario, it does not, the diffusion model needs more complexity. Predictions may be tested using short lived chemical tracers in the downwelling branch defining when the air was last in contact with the surface. More generally, the mean and eddy-type moisture fluxes and divergences need to be calculated from the ECMWF analyses to examine how well they balance and what the implied missing terms are.

Another task is to bin the GEW and ECMWF data into longitude versus height at low latitudes to examine the monsoonal-Walker circulation pattern and understand how the temperature and moisture variability observed near 30N relate to the Walker circulation and moisture transport. Unfortunately, it is not clear that the GEW sampling density is adequate to do much.

8. Summary and Conclusions

In summary, the work and results presented here represent a detailed first order examination of both the expected performance and capabilities in terms of resolution and accuracy of atmospheric structure derived from GPS occultation observations as well as a demonstration of those capabilities using initial results derived from the GPS-MET prototype experiment. The work presented here represents a marriage between JPL and Caltech objectives and goals with a JPL emphasis more on technology and Caltech on interpretation. (The JPL side is winning so far).

We began with a description of the derivation of atmospheric refractivity from GPS occultation observations and how bulk dry gas and water vapor densities contribute to microwave refractivity. Further, we quantified the effects of particle scattering and found them to be small. By examining sources of error, the altitude range of useful refractivities was shown to extend roughly to an upper limit near 60 km limited by the ionosphere, initialization conditions and signal to noise ratio. Near the surface, the most significant problems in the deriving atmospheric structure result from a combination of non-spherical structure, defocusing, atmospheric multipath and diffraction.

The resolution of the geometric optics approximation was described both in terms of the contribution of vertical atmospheric layers to bending angle and diffraction yielding ~1.4 km vertical resolution in the middle atmosphere improving to ~500 m or better near the surface. The unique combination of long wavelengths and 1 km or better vertical resolution offers the first opportunity to routinely probe the troposphere in a limb-viewing geometry. Working with Roger Linfield at JPL, we developed a general method to describe the effects of diffraction in a thick atmosphere combining geometric optics within the Huygens-Fresnel representation of electromagnetic wave propagation and used the

method to assess the accuracy of the geometric optics approximation and to predict the effects of diffraction near the tropopause which will be discussed in a future paper.

A detailed examination of the impact of error on derived results was described. A methodology was developed for characterizing error propagation through the retrieval process culminating in comprehensive error budgets for derived refractivity, geopotential of pressure surfaces and temperature as functions of height, latitude and solar cycle. Observations will yield the most accurate results near the tropopause with predicted individual profiles accuracies of ~0.2% in refractivity, ~10 m in geopotential height and 0.4 K in temperature. When averaged, the accuracies should improve by at least another order of magnitude.

We then developed a first order assessment of expected accuracy of mid and lower tropospheric water vapor. Specific humidities in the upper troposphere are too small for accurate derivation via refractivity at GPS wavelengths. Water vapor accuracy is limited by errors and uncertainties in pressure, temperature and retrieved refractivity and is a strong function of water vapor abundance. The best measurements will be obtained in the warmer regions of the troposphere at low latitudes in the middle and lower troposphere where individual profile accuracies of better than 5% in the boundary layer and 20% up to 6 or 7 km altitude are predicted. Averaging for climatological purposes may improve upon these individual profile accuracies by a factor of several.

Several classes of applications have been described including weather forecasting, climate research, the hydrological cycle, boundary layer processes and atmospheric waves. Long term GPS occultation observations will provide data to test several predicted climate change signatures including stratospheric cooling, increased specific humidity at low latitudes and expansion of the troposphere. Further refractivity will undoubtedly become an important climate diagnostic because of the GPS measurement accuracy and the ease with which it can be calculated in models. Given the unique combination of good vertical

resolution and insensitivity to clouds, GPS observations will improve our knowledge of the vertical humidity structure of the warmer troposphere significantly, particularly over regions such as the oceans where radiosonde coverage is sparse or nonexistent.

With the launch of the GPS-MET mission in April 1995, first order demonstration of several occultation features and capabilities was enabled. We described initial results using system whose development I oversaw and participated in at JPL. Comparison with radiosondes and ECMWF analyses revealed accurate reconstruction of tropopause structure with ~1 km level resolution and ~1 K level accuracy as predicted. Further, the reduced accuracy of weather model structure in the southern hemisphere is evident in the comparisons indicating the impact of limited vertical information of the analyses and suggesting where occultation data will significantly improve on present knowledge. We also summarized the range of climatological behavior in the stratosphere and troposphere, demonstrating the extremely cold tropical tropopause, the large seasonal differences between the winter and summer hemispheres and extremely cold winter polar stratosphere demonstrating consistency with present knowledge of temperature climatology.

Examination of many temperature profiles revealed wave activity at low latitudes significantly greater than that at higher latitudes. The recovery of equatorial wave structure provides another demonstration of vertical resolution and suggests these are the first such observations from space and may provide a method of monitoring tropical convective activity and intensity. A preliminary study was made attempting to identify the wave types which indicated the most likely candidate appears to be inertio-gravity waves. However, the results are not terribly satisfactory and will require a more in-depth examination and additional observations. The existence and magnitude of these waves suggest they play a role linking convective activity with tropical exchange between the stratosphere and troposphere both in terms of modifying minimum temperatures near the tropopause and the freeze-drying effect as well as vertical motion at the tropopause. Minimum temperatures

appear sufficiently cold to explain lower stratospheric specific humidities of ~ 3 ppmv. Examination of the ability of RGW and IGWs to force flow across the tropopause also made the importance of radiation apparent. Combining this understanding with the global observations of the tropopause, the energy equation and some intuition concerning the vertical signature of radiative transfer led to the hypothesis that air is systematically rising across the tropopause and returning across tropopause discontinuities. Further, this behavior is general and must occur in any large, low dissipation atmosphere. This hypothesis mat be tested using radiative transfer calculations and observations of short lived atmospheric tracers.

An initial set of water vapor profiles were derived from June-July 1995 using the ECMWF global analyses temperature to isolate and remove the dry contribution to refractivity. The validity of this approach is particularly apparent at low latitudes where temperature variations are small and but refractivity variations due to water vapor are large. The derived water vapor results were notably drier in southern hemisphere than P&O and in general agreement with the ECMWF humidities and revealed some systematic differences. This work included examination of refractivity, pressure, temperature and water vapor in sorting out errors limiting water vapor accuracy. Results indicate consistency at the 0.5% level near 8 km altitude and above. It appears that occultation refractivities may be slightly underestimated. We suggested a possible cause associated with diffraction which will be investigated in the near future using the diffraction method just mentioned. A tri-pole bias pattern was revealed near 2 km altitude with ECMWF humidities being lower in the tropics and higher in the subtropics consistent with previously observed bias between columnar water vapor derived from SSM/I and TOVS data. The vertical resolution of the GPS data has made the vertical structure of the bias apparent. Signatures of noisy ECMWF temperatures in southern hemisphere are apparent both in terms of standard deviation and biases. A future endeavor is to correct the ECMWF

temperature-pressure structure using the occultation refractivity and pressure structure in the upper troposphere.

We identified the large scale features of the meridional Hadley circulation in the GEW results such as the ITCZ which stands out as a maximum in relative humidity and the downwelling zones which are apparent as regions of minimum relative humidity. The southern hemisphere downwelling zone is very dry with relative humidities as low as 12%. The effects of monsoons and Walker circulation are apparent as well as relative maxima in mean and standard deviation of specific humidity near 30N. Based upon the focused downwelling in the descending branch of the Hadley circulation of the P&O climatological analysis, a large change in the meridional gradient of mean specific humidity was predicted and one was, in fact, observed. This feature was used in conjunction with the hypothesis that moisture in the downwelling region is controlled by a balance between downward moisture divergence and diffusive meridional convergence to predict the location and structure of the Hadley downwelling. The resulting zone of intense downwelling sits roughly 5^{o} south of its location in the P&O climatology. Vertical winds of the ECMWF analysis will be used in the near future to determine whether the single diffusion parameter model can describe the flow of moisture. If true, it will provide insight into a basic mechanism and may be used in climate models to extrapolate behavior to future climates. This represents one example of many where the ECMWF wind fields need to be combined with the existing analysis to better constrain moisture fluxes and divergences.

In conclusion, the initial GPS results have been very accurate as expected indicating a very promising future and increasing roles in many directions in the future. Several missions are scheduled to launch over the next three years and developing plans for receiver constellations should yield sampling densities sufficient to have a significant impact in many areas of atmospheric research.

A. Appendices

A.1 Scattering by Atmospheric Particles

Scattering by particles suspended in the atmosphere will affect the propagation of electromagnetic signals. Scattering was exploited in Voyager radio occultation observations of the rings of Saturn to infer particle size distributions within the rings [Tyler, 1987]. Here we are concerned primarily about the effects of liquid water droplets in the lower atmosphere on the phase and amplitude of the occulted GPS signals. Water droplets are the primary source of scattering both because of their potentially large amount and because of the large relative dielectric constant of liquid water at microwave wavelengths.

Because the GPS wavelengths are of order 20 cm, scattering by these particles falls into the Rayleigh regime. To get an approximate answer spherical particles will be assumed. Following the approach of Gresh (1990), the scattering extinction efficiency, Q_S (extinction cross-section divided by the geometric cross-section), under conditions of no absorption is

$$Q_S \cong \frac{8(kr)^4}{3}\left(\frac{m^2-1}{m^2+2}\right)^2 \qquad \text{(A.1.1)}$$

where $k = 2\pi/\lambda$, λ is the wavelength, r is the particle radius and m is the index of refraction. The finite conductivity of water leads to an analogous absorption efficiency, Q_A, which at GPS wavelengths dominates what little effect there is on signal amplitude.

$$Q_A \cong 4kr\ Im\left(\frac{m^2-1}{m^2+2}\right) \qquad \text{(A.1.2)}$$

The corresponding but less familiar phase term is the phase-shift efficiency, P, which is defined analogously to be

$$P \cong 2\,kr\ Re\left(\frac{m^2-1}{m^2+2}\right) \tag{A.1.3}$$

The index of refraction of water at microwave wavelengths is

$$m = \sqrt{\frac{\varepsilon'}{\varepsilon_o}\left(1+i\frac{\sigma}{\omega\varepsilon'}\right)} = \sqrt{\frac{\varepsilon'}{\varepsilon_o}\left(1+i\frac{\varepsilon''}{\varepsilon'}\right)} \tag{A.1.4}$$

where ε' is the real part of the dielectric constant of water, ε_o is the dielectric constant of free space, σ is the electrical conductivity of water and ω is the frequency of the signal in radians/sec and ε'' is the imaginary part of the dielectric constant of water.

These terms represent the amplitude and phase shift of the scattered signal relative to the incident signal caused by a single scattering particle. A many particle thick model of scatterers results in

$$\frac{E}{E_i} = e^{-\tau/2}\, e^{i\phi_c} \tag{A.1.5}$$

where $\tau = \pi\, r^2\, n_p\, L\, Q$, $\phi_c = \pi\, r^2\, n_p\, L\, P$, E_i is the incident electric field, n_p is the particle number density and L is the thickness of the layer. The liquid water content, W, of the atmosphere is the particle number density times the volume per particle times the density of liquid water, ρ_L, integrated across all particle radii. For simplicity we assume a monochromatic particle size distribution with all droplets having the same radius, a reasonable first order assumption because number densities fall off exponentially with drop radius and the distributions tend to be narrow (Atlas, 1964). With a monochromatic assumption, particle number density is proportional to liquid water content divided by r^3. Since ϕ_c is proportional to $n_p\, r^3$, ϕ_c is independent of particle size for a given W. From this one can show the refractivity due to scattering is

$$N_S = 1.5x10^6 \frac{W}{\rho_L} Re\left(\frac{m^2-1}{m^2+2}\right) \tag{A.1.6}$$

This expression evaluated with $m^2 = 80$, roughly the real part of the dielectric constant at GPS wavelengths (Bohren and Huffman, 1983), is the source of the scattering refractivity term in eq. (2.3.1). Since m^2 (ice) $\cong 3$, N_{Sice} is roughly 40% that of liquid water, N_{Sliq}. A liquid water density of 1 g/m^3, a large value for the atmosphere, produces a refractivity of ~1.4, which would contribute less than 1% of the total refractivity in the lower troposphere. Producing a perturbation of this magnitude requires that condensed water extend continuously along the horizontal signal path through the lower troposphere. Therefore stratiform clouds and their precipitation would be the most significant contributor to the effect. The impact of scattering will be most significant where absolute humidities are high; conditions under which the refractivity profile will most likely be used to recover water vapor abundances (Section 4). Under these conditions, refractivity caused by scattering will be interpreted as being caused by water vapor resulting in a slight overestimate of derived water vapor. The fractional overestimate of water vapor will be N_{Sliq}/N_{WV} which from eq. (2.3.1) is

$$\frac{N_{Sliq}}{N_{wv}} = \frac{T\,W}{1.2x10^3\,\rho_{wv}} \cong \frac{W}{4\,\rho_{wv}} \qquad \text{(A.1.7)}$$

approximately one quarter of the ratio of atmospheric liquid water to water vapor averaged over the horizontal sampling volume of the signal path. The fractional overestimate of water vapor when the condensed phase of water is ice will be N_{Sice}/N_{WV} which, based on the 40% scaling of N_{Sice} to N_{Sliq} mentioned above, will be

$$\frac{N_{Sice}}{N_{WV}} \cong \frac{W}{10\,\rho_{WV}} \qquad \text{(A.1.8)}$$

Representative values of W/ρ_{wv} in the atmosphere are given as a function of rain rate and temperature in figure A-1. Typical values for rain rate are of order 4 mm/hr. Under warm temperature conditions where water vapor can make a significant contribution to total refractivity, W/ρ_{wv} is generally a few % or less. Consider the contours labeled

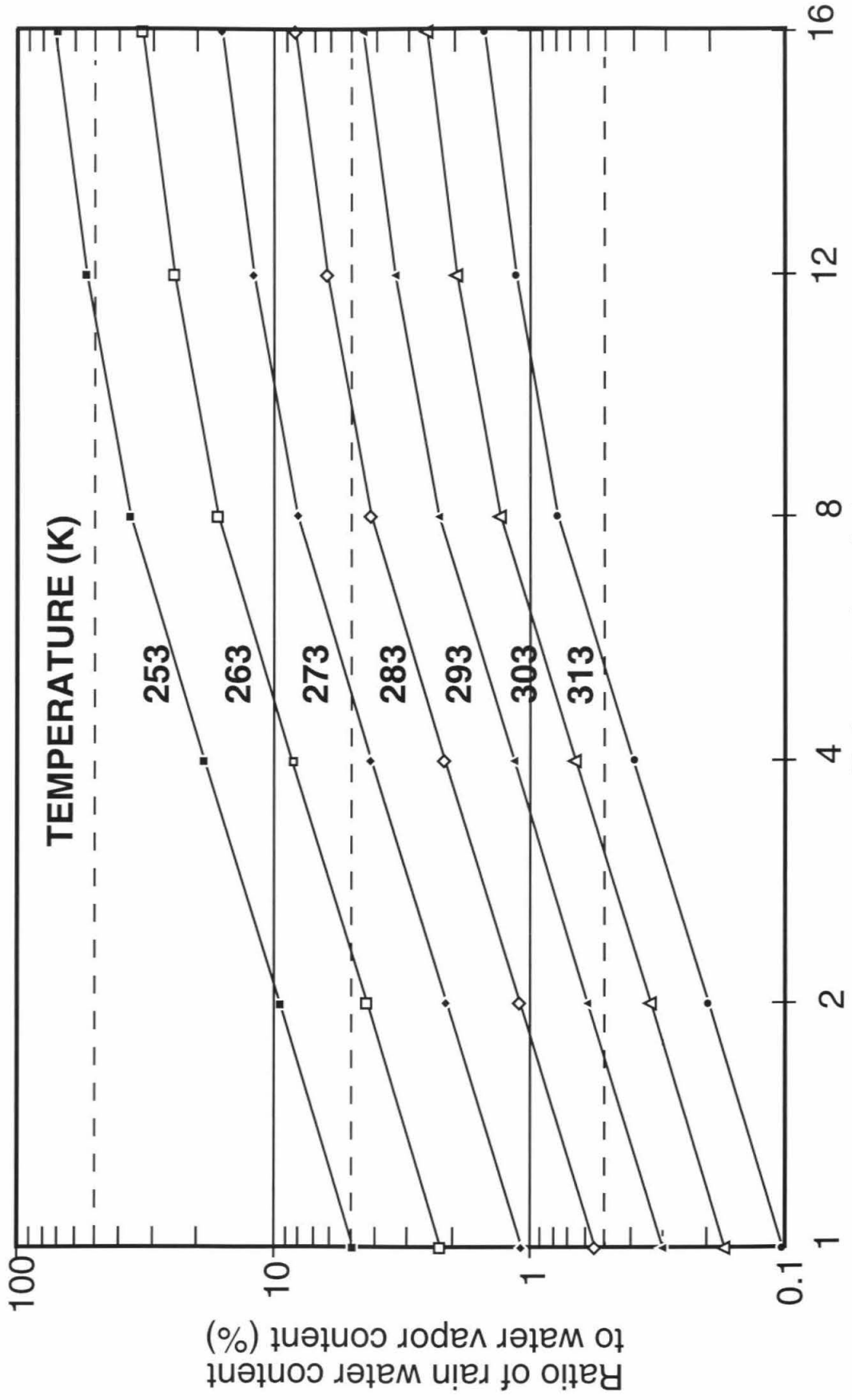

Figure A-1 Ratio of liquid water to water vapor densities as a function of rain rate and temperature

"5%" in figure 4.1 which represent the approximate height at which water vapor can be retrieved with a 5% accuracy under the assumption that temperature is known to 1.5 K. The 5% contours correspond crudely to a temperature of 285 K. Assuming liquid water is present along the ray path in the lower troposphere at a ratio $W/\rho_{wv} \sim 2\%$ (corresponding to a rain rate of 4 mm/hr and temperature of 285 K), then eq. (A.1.7) indicates water vapor abundance will be overestimated by ~0.5%, an insignificant amount. The overestimate corresponding to the 10% water vapor contours in figure 4.1 (for a rain rate of 4 mm/hr and temperature of 273 K) is ~1%, which is again insignificant. For a given rain rate, as temperature decreases, W/ρ_{wv} increases, but water vapor density is decreasing very rapidly so the scattering contribution to *total* refractivity is decreasing. At 253 K, a temperature where water vapor contributes little to total refractivity, W/ρ_{wv} would be roughly 20% for a 4 mm/hr rain rate which, via eq. A.1.8, leads to scattering refractivity of order 2% of water vapor refractivity. Therefore, the contribution of scattering to refractivity will be a minor error source in general.

In evaluating signal extinction, data from Atlas (1964) has been used to relate liquid water content to droplet particle diameter. The resulting changes in signal amplitude are shown in figure A-2. The path length through the scattering particles is taken to be $L = 200$ km, the approximate horizontal path length through a layer having a vertical dimension of the first Fresnel zone. Therefore, these results should be taken as upper bounds on the effect of atmospheric water droplets. The amplitude attenuation is negligible for particles less than 1 mm in radius which corresponds to the largest particle sizes typically observed over extended horizontal regions.

A radio occultation system operating at multiple wavelengths shorter than those of the present GPS system could sense and discriminate between water droplet sizes and distributions. With sufficient rain along the path of an occultation, attenuation due to scattering will occur along with signal defocusing due to bending. The refractivity

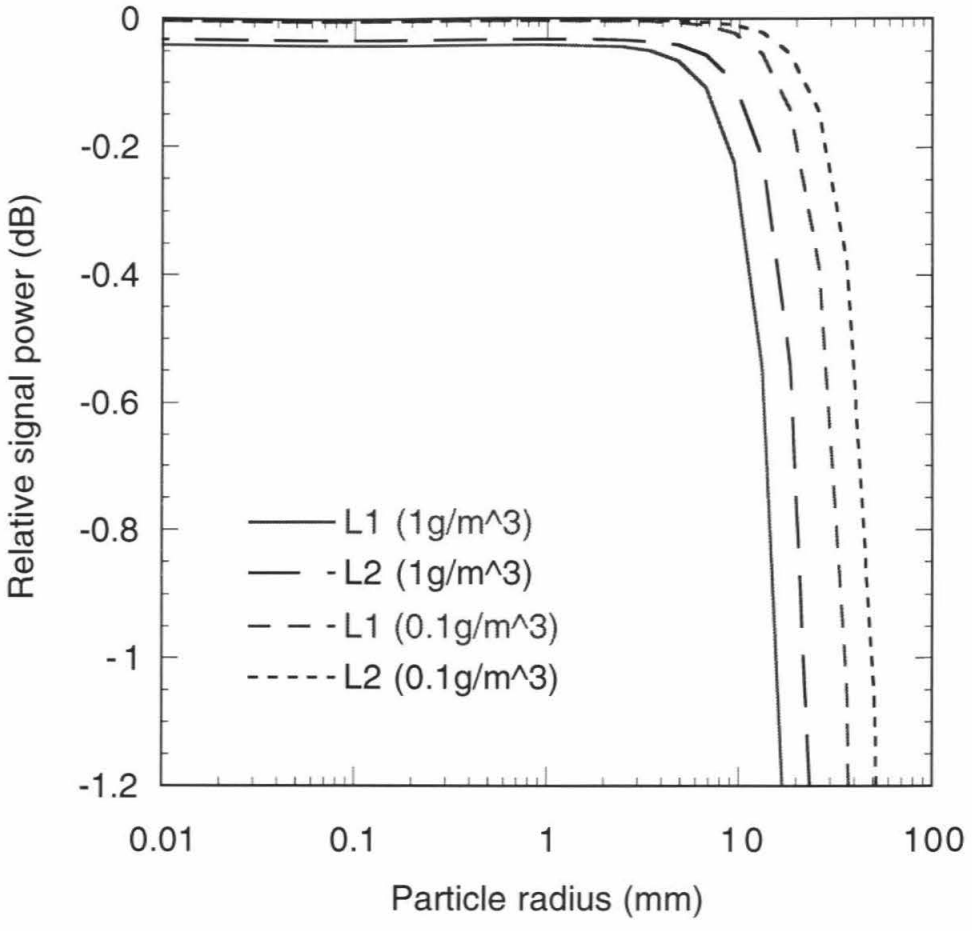

Figure A-2 Extinction of GPS signals due to scattering by water droplets.

structure retrieved from the phase data allows the attenuation due to defocusing to be estimated and subtracted from the observed attenuation. Any residual signal decrease would be extinction due to scattering. The same approach has been applied to occultation data taken at the outer planets to estimate the ammonia vapor density. The amount of attenuation would be a function of wavelength and the attenuation difference between wavelengths could be used to determine the size of the raindrops much in the manner in which particle distributions were determined for Saturn's rings.

A.2 Description of Errors in the Fourier Domain

Certain noise processes such as clock noise and local multipath are best characterized in the fourier domain in terms of power spectra, $S_x(f)$, defined as the fourier transform of the autocovariance, $\langle x(t)\, x(t+\tau)\rangle$ where $\langle x\rangle$ means the expected value of x. The purpose of this section is to relate the mean square error (MSE) ($\equiv \langle x(t)^2\rangle$) of the refractivity, density, pressure and temperature errors to the power spectrum of the phase measurement noise. In order to do so, extensive use will be made of the following relation: if A can be related to $\delta\phi(t)$ as follows,

$$A = \int_{-\infty}^{\infty} dt'\, \delta\phi(t')\, w(t') \qquad \text{(A.2.1)}$$

then the MSE of A is defined as

$$\langle A^2 \rangle = \int_{-\infty}^{\infty} df\, S_{\delta\phi}(f) \left| W(f) \right|^2 \qquad \text{(A.2.2)}$$

where $S_{\delta\phi}(f)$ is the power spectrum of $\delta\phi(t)$, W(f) is the fourier transform of w(t) and |W(f)| denotes the magnitude of W(f). Eq. (A.2.2) requires that the quantity, $\delta\phi(t)\, w(t)$, be stationary, a requirement which can be met with an appropriate choice of w(t). Eq.

(A.2.2) relates the MSE of A to the power spectrum of the phase noise, $\delta\phi(t)$, and its use requires that the relationship between A and $\delta\phi$ can be expressed in the form of eq. (A.2.1). Eq. (A.2.1) is in fact very similar to the convolution integral used in Section 3.2 to relate phase measurement error to refractivity error. Therefore, the purpose of this section will be to first write each of the four convolution integrals relating phase measurement error to errors in refractivity, density, pressure and temperature in the form of eq. (A.2.1) thereby defining w(t) for each of these four variables. The second step will be to derive the fourier transform of each w(t) such that the MSE of refractivity, density, pressure and temperature can be estimated via eq. (A.2.2). The proportionality between time and height above the tropopause has been described in Section 3.2.

Refractivity and density error descriptions in the fourier domain

As described in Section 3.2, refractivity error can be described as the convolution of phase error and the impulse response of the combined differentiator-Abel transform processing steps, $h_{DA}(t)$. Refractivity error can therefore be written as

$$\delta N(t_a) = \int_0^{t_a} d\beta\ \delta\phi(\beta)\ h_{DA}(t_a-\beta) = \int_{-\infty}^{\infty} \delta\phi(\beta)\ h_{DA}(t_a-\beta)\ rect_{0,t_a}(\beta)\ d\beta \qquad (A.2.3)$$

where $rect_{0,ta}(t)$ is the rectangle function equal to 1 for $0 \le t < t_a$ and 0 otherwise and t_a is time measured since the time of the first measurement used in the Abel transform. The presence of the rectangle function means the length of the impulse response increases with t_a, an important dependence to be maintained when estimating error via the fourier domain. Eq. (A.2.3) is of the form of eq. (A.2.1) with

$$w_N(t_a;t) \equiv h_{DA}(t_a-t)\ rect_{0,t_a}(t) \qquad (A.2.4)$$

so the mean square refractivity error is given as

$$\left\langle \left[\delta N(z)\right]^2 \right\rangle = \int_{-\infty}^{\infty} df\ S_{\delta\phi} \left| W_N(t_a;f) \right|^2$$

where $W_N(t_a;f)$ is the fourier transform of $w_N(t_a;t)$. $W_N(t_a;f)$ is found by taking the fourier transform

$$W_N(t_a;f) = \int_0^{t_a} d\beta\, h_{DA}(t_a-\beta)\, e^{-i2\pi f\beta}$$

Changing variables results in

$$W_N(t_a;f) = e^{-i2\pi f t_a} \int_0^{t_a} d\sigma\, h_{DA}(\sigma)\, e^{i2\pi f\sigma} \equiv e^{-i2\pi f t_a} H_{DA}(t_a;-f) \qquad \text{(A.2.5)}$$

where $H_{DA}(t_a;f)$ is the fourier transform of the truncated version of $h_{DA}(t)$ which has been set to 0 for $t > t_a$. The mean square refractivity error is given as

$$\left\langle \left[\delta N(z)\right]^2 \right\rangle = \int_{-\infty}^{\infty} df\, S_{\delta\phi} \left| H_{DA}(t_a;-f) \right|^2 \qquad \text{(A.2.6)}$$

The density error is

$$\left\langle \left[\delta\rho(z)\right]^2 \right\rangle = b_d^2 \left\langle \left[\delta N(z)\right]^2 \right\rangle \qquad \text{(A.2.7)}$$

Pressure error in the fourier domain:

Pressure is derived as the hydrostatic integral of density and the error in pressure is therefore the hydrostatic integral of the density error. Because phase error is a function of time, the previously discussed proportionality between height and time discussed in Section 3.2 is used such that the temporal and vertical pressure gradients are related as

$$\frac{dP}{dt} = V_\perp \frac{dP}{dz} = -\frac{V_\perp\, g\, m\, N}{R\, b_1} = b_P\, N$$

where $V_\perp$ is the ray path ascent or descent velocity, R is the gas constant, m is mean molecular mass and b_1 is defined in eq. (2.3.1). b_P is of order 100, ranging from 88 to 132 (Pa s^{-1} N-unit^{-1}) for ray path descent or ascent velocities of 2 to 3 km/sec respectively. The pressure error is the integral of the refractivity error:

$$\delta P(z) = b_P \int_0^{t_a} \delta N(t')\, dt' = b_P \int_0^{t_a} dt' \int_0^{t'} d\beta\, \delta\phi(\beta)\, h_{DA}(t'-\beta)$$

Reversing the order of integration,

$$\delta P(z) = b_P \int_0^{t_a} d\beta\, \delta\phi(\beta) \int_\beta^{t_a} dt'\, h_{DA}(t'-\beta)$$

$$\delta P(z) = b_P \int_{-\infty}^{\infty} d\beta\, \delta\phi(\beta)\, rect_{0,t_a}(\beta) \int_\beta^{t_a} dt'\, h_{DA}(t'-\beta) \equiv b_P \int_{-\infty}^{\infty} d\beta\, \delta\phi(\beta)\, w_P(t_a;\beta) \quad \text{(A.2.8)}$$

Now, since δP has been related to $\delta\phi$ in the form of eq. (A.2.1), the next step is to calculate $W_P(t_a;f)$, the fourier transform of $w_P(t_a;t)$, which is defined as

$$W_P(t_a;f) = \int_0^{t_a} d\beta\, e^{-i2\pi f\beta} \int_\beta^{t_a} dt'\, h_{DA}(t'-\beta)$$

Reversing the order of integration,

$$W_P(t_a;f) = \int_0^{t_a} dt' \int_0^{t'} d\beta\, h_{DA}(t'-\beta)\, e^{-i2\pi f\beta}$$

and changing variables with $\sigma = t'-\beta$ results in

$$W_P(t_a;f) = \int_0^{t_a} dt'\, e^{-i2\pi ft'} \int_0^{t'} d\sigma\, h_{DA}(\sigma)\, e^{i2\pi f\sigma} = \int_0^{t_a} dt'\, e^{-i2\pi ft'} H_{DA}(t';-f) \quad \text{(A.2.9)}$$

where $H_{DA}(t';f)$ is defined as in eq. (A.2.5). Therefore, $W_P(t_a;f)$ is the truncated fourier transform of the truncated fourier transform of $h_{DA}(t)$. The mean square error of pressure is then defined by inserting eq. (A.2.9) into eq. (A.2.2) such that

$$\left\langle \left[\delta P(z)\right]^2 \right\rangle = b_P^2 \int_{-\infty}^{\infty} df\, S_{\delta\phi} \left| W_P(t_a;f) \right|^2 \quad \text{(A.2.10)}$$

Temperature error in the fourier domain:

Temperature error is related to density and pressure errors in a fractional form as

$$\frac{\delta T}{T} = \frac{\delta P}{P} - \frac{\delta \rho}{\rho} = \frac{1}{\rho}\left(\frac{\delta P}{g\,H_P} - \delta\rho\right) \qquad (3.2.6)$$

where g is gravitational acceleration and H_P is the pressure scale height. Combining eq. (A.2.7) and (A.2.3) into eq. (3.2.6), the fractional temperature error is given as

$$\frac{\delta T}{T}(t_h) = \frac{1}{\rho(z)}\int_{-\infty}^{\infty} d\beta\, \delta\phi(\beta)\left(\frac{b_P}{g\,H_P}\, w_P(t_a;\beta) - b_d\, w_N(t_a;\beta)\right) \qquad (A.2.11)$$

The fractional mean square temperature error is therefore

$$\left\langle\left(\frac{\delta T}{T}(t_a)\right)^2\right\rangle = \frac{1}{\rho^2(z)}\int_{-\infty}^{\infty} df\, S_{\delta\phi}(f)\left|\frac{b_P}{g\,H_P}\, W_P(t_a;f) - b_d\, W_P(t_a;f)\right|^2 \qquad (A.2.12)$$

where $W_P(t_a;f)$ and $W_N(t_a;f)$ are defined in eqs. (A.2.9) and (A.2.5).

Fourier representations of impulse responses, $h_D(t)$ and $h_A(t)$

The refractivity, density, pressure and temperature error estimates have been derived in this appendix in terms of $h_{DA}(t)$ which is the combined impulse response of differentiation and the Abel transform steps. Since the combined response is the convolution of the two individual impulse responses, the fourier transforms of each impulse response will be multiplied together to represent the fourier representation of the combined impulse response.

$h_D(t)$ represents differentiation and its fourier transform is $i2\pi f$ when a continuous derivative is used. However, because of the discrete nature of the measurements, a finite

differencing form of differentiation will be used. The fourier transform of the simple, discrete form of differentiation defined in eq. (3.2.2) is

$$H_D(f) = \frac{\left(e^{-i\pi f\tau} - e^{i\pi f\tau}\right)}{\tau} \qquad \text{(A.2.13)}$$

where τ is taken as the time for the ray path tangent point to descend a Fresnel zone. The magnitude squared from the definition in eq. (3.2.2) is

$$\left|H_D(f)\right|^2 = \frac{4\sin^2(\pi f\tau)}{\tau^2} \qquad \text{(A.2.14)}$$

Note that in using the discrete form of eq. (3.2.2), the frequency estimates are shifted in time by $\tau/2$ relative to the phase estimates. The remaining portion of $H_{DA}(f)$ is $H_A(f)$, the fourier transform of $h_A(t)$ representing the impulse response of the conversion of Doppler frequency versus time into refractivity versus height. As described in Section 3.2, $h_A(t)$ can be estimated numerically by passing a set of perfect measurements plus one bad one through the retrieval system and using the resulting refractivity error to represent the impulse response of the system. This approach raises two important issues: the finite thickness of the delta function error approximation and the discrete sampling of the impulse response.

The first issue lies in the fact that the single erroneous measurement is not a true delta function because of the finite time interval between the measurements. The interpolator in the numerical implementation of the Abel transform interpolates across the finite intervals between consecutive frequency estimates transforming the single frequency error into a continuous error somewhat triangular in form centered on the original error. The exact shape of the continuous error is the impulse response of the interpolator, $h_I(\tau;t)$ where τ is the time interval between samples. The estimated impulse response, $\tilde{h}_A(t)$, is therefore the true impulse response, $h_A(t)$, convolved with $h_I(\tau;t)$. For example, $h_I(\tau;t)$ is a

triangle function when a linear interpolator is used. The fourier transform of $\tilde{h}_A(t)$ is therefore

$$\tilde{H}_A(f) \cong H_A(f)\ H_I(\tau;f)$$

where $H_A(f)$ is the exact fourier transform of the true impulse response, $h_A(t)$, and $H_I(\tau;f)$ is the fourier transform of the $h_I(\tau;t)$. The triangular impulse response of the linear interpolator is defined as

$$h_I(\tau;t) = \begin{cases} -\frac{|t|}{\tau} + 1 & for\ |t| < \tau \\ 0 & for\ |t| > \tau \end{cases}$$

and its fourier transform is

$$H_I(\tau;f) = \tau \frac{\sin^2(\pi f \tau)}{(\pi f \tau)^2}$$

Generalizing from this response, the interpolator will scale $H_A(f)$ by a factor of τ and attenuate the high frequency response to some degree.

The second issue raised about the estimation of the impulse response is its representation as a set of discrete samples separated by the Fresnel interval to be consistent with the geometric optics approximation as described in Section 3.2. According to the Nyquist theorem, the spectrum of $h_A(t)$ whose maximum frequency, f_{max}, is $\leq (2\tau)^{-1}$, is equal to τ times the spectrum of the discretely sampled $h_A(t)$ where τ is the time interval between samples. Because the geometric optics approach is valid only for structures greater than or equal to the Fresnel diameter, Z_D, the maximum fourier frequency, f_{max}, of the spectrum of $h_A(t)$ is $(2Z_D/V_\perp)^{-1}$. This discrete representation of the spectrum therefore assumes attenuation by smoothing in the phase and frequency estimation process is sufficient to make noise power at frequencies greater than f_{max} negligible.

The fourier transform of the discretely sampled $\tilde{h}_A(t)$ ($\equiv \tilde{h}_A^d(t)$) is

$$\tilde{H}_A^d(f) = \int_{-\infty}^{\infty} \sum_{j=0}^{j\,max} \left[\tilde{h}_j\, \delta(\alpha - j\tau)\right] e^{-i2\pi ft}\, dt = \sum_{j=0}^{j\,max} \left[\tilde{h}_j\, e^{-i2\pi fj\tau}\right]$$

which, from the finite delta function and discrete sampling discussions, must equal

$$\tilde{H}_A^d(f) = H_A^d(f)\, H_I(f) = \frac{1}{\tau}\, H_A(f)\, H_I(f)$$

For linear interpolation,

$$\tilde{H}_A^d(f) = H_A(f)\, \frac{\sin^2(\pi\, \tau f)}{(\pi\, \tau f)^2}$$

With a linear interpolator, the fourier transform of the discretely sampled estimate equals the transform of the true impulse response at low frequencies but is somewhat attenuated at higher frequencies where the estimated spectrum is somewhat attenuated. Because the actual interpolator used is a cubic interpolator, the high frequency attenuation will be less than that of the linear interpolator and the approximation will be made that

$$\tilde{H}_A^d(f) \cong H_A(f)$$

A.3 Local Multipath Description

The broad-beam GPS receiver antennas are sensitive to "mirages" or multiple images of the same signal arriving at the antenna having scattered off structures in the immediate vicinity of the receiving antenna. This effect is referred to as *local* multipath (as opposed to *atmospheric* multipath which is produced by sharp variations in refractivity structure in the atmosphere). In the presence of local multipath, the receiver tracks the sum of the multiple images of the signal. Each signal image has, in general, some non-zero propagation delay relative to the direct signal such that the phase of the sum of the images differs in general from that of the direct signal alone. The sum tracked by the receiver is

$$S(t) = A\, e^{i(\omega t - \omega\tau - \phi)} + \sum_{j=1}^{j\,max} A_j\, e^{i(\omega t - \omega\tau_j - \phi - \phi_j)} = A\, e^{i(\omega t - \omega\tau - \phi)} \left(1 + \sum_{j=1}^{j\,max} \frac{A_j}{A}\, e^{i(-\omega\Delta\tau_j - \phi_j)} \right)$$

where $\Delta\tau_j$ is the extra travel time of the multipath signal reflecting off point j. Therefore, an extra phase shift (in units of radians) of

$$\delta = \arctan\left(\frac{\sum_{j=1}^{j\,max} \frac{A_j}{A} \sin\left(\omega\Delta\tau_j + \phi_j\right)}{1 + \sum_{j=1}^{j\,max} \frac{A_j}{A} \cos\left(\omega\Delta\tau_j + \phi_j\right)} \right) \qquad \text{(A.3.1)}$$

exists relative to the direct signal. For $A_j << A$ for all j's, eq. (A.3.1) becomes

$$\delta = \sum_{j=1}^{j\,max} \frac{A_j}{A} \sin\left(\omega\Delta\tau_j + \phi_j\right) \qquad \text{(A.3.2)}$$

Two important points can be made about eq. (A.3.2). First, the magnitude of the multipath-induced phase delay, δ, depends directly on the amplitude of the multipath signals relative to the direct signal and it is therefore desirable that multipath signals be attenuated as much as possible. Secondly, δ is a function of time because the geometry between the transmitter, receiver and each scattering location, j, is continually changing due to orbital motion.

Because bending angle is derived from phase rate, the time variation of $\Delta\tau_j$ relative to the duration of an occultation determines the impact local multipath has on derived atmospheric quantities. The time variation depends in turn on both the angular velocity at which the transmitter orbits the receiver and the distance of the scattering point from the receiving antenna. To quantify the time dependence, consider a reflecting surface near the antenna (figure A-3). The extra travel time between the direct signal and the signal specularly reflected off the surface, j, is

$$\Delta\tau_j = \frac{2L_j \sin\theta_j}{c} \qquad \text{(A.3.3)}$$

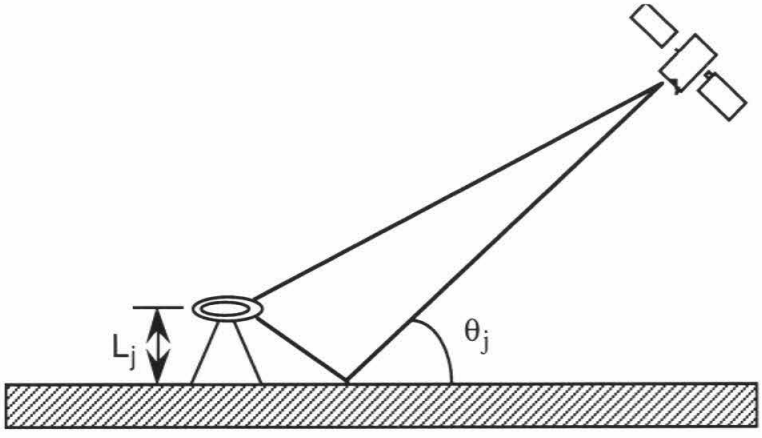

Figure A-3 Geometry of local multipath.

where L_j is the height of the antenna above the reflecting surface and θ_j is the angle between the direction to the GPS transmitter and the reflecting surface. Time variation enters through θ_j. Let $\theta_j = \omega_s t$ where ω_s is the angular revolution frequency of the transmitting satellite as seen by the receiving satellite and expand $\sin \omega_s t$ about the value of $\omega_s t$ at the midpoint of the occultation ($\equiv \theta_{rj}$) to linearize the time dependence:

$$\sin(\omega_s t) = \sin(\theta_{rj}) + (\omega_s t - \theta_{rj}) \cos(\theta_{rj}) \qquad \text{(A.3.4)}$$

Plugging eq. (A.3.4) into eq. (A.3.2) gives the period of the jth component of multipath phase, defined such that the change in $\Delta\tau_j$ equals $2\pi/\omega$, as

$$\tau_{mpj} = \frac{\lambda \tau_s}{4\pi L_j \cos(\theta_{rj})} \qquad \text{(A.3.5)}$$

or equivalently a frequency

$$\omega_{mpj} = \frac{8\pi^2 L_j \cos(\theta_{rj})}{\lambda \tau_s} \qquad \text{(A.3.6)}$$

where τ_{mpj} and ω_{mpj} are the period and angular frequency of the beat frequency between the direct signal and the jth component of multipath oscillation, τ_s is the orbital period and λ is the GPS signal wavelength. It is apparent that the frequency of the multipath-induced phase oscillation is proportional to the distance between the reflecting surface and the antenna, proportional to $\cos(\theta_{rj})$, and inversely proportional to the revolution period of the transmitter as viewed from the receiver. The GPS satellites, LEO receivers and Earth-based receivers have 12 hour, ~100 minute and 24 hour orbital periods respectively. Therefore, the LEO receiver will see phase oscillations which are 5 to 14 times more rapid than the ground receiver for the same local scattering geometry. The maximum frequency of multipath for a 1 m spacecraft is ~0.01 Hz. While spacecraft have finite size which limits L_j, antennas on the Earth can sit farther from scattering locations and may therefore experience higher frequency multipath oscillations. However, most multipath effects tend

to be dominated by structures within a few meters of the antenna. The CA signal is more sensitive to large multipath delays of up to a CA pseudo-range "chip" length of 300 m.

The minimum frequency of 0 occurs when θ_{rj} is 90° or 270°. However, since θ_{rj} depends both on geometry and electrical properties of numerous scatterers, it will be difficult, in general, to control, limiting the utility of the $\cos(\theta_{rj})$ term. A rough average magnitude of the $\cos(\theta_{rj})$ term can be taken as 0.5.

Magnitude of the multipath phase error

Now the relative amplitude of local multipath signals must be considered. Attenuation depends on the relative antenna gain between the direction of the direct signal and the direction to each scattering point, the reflectivity at each scattering location and the distance from the antenna to the scattering point. The multipath signals can be suppressed further in post-processing if the physics of the multipath process can be adequately modeled. These issues, when taken together, demonstrate that local multipath is very dependent on the particular configuration of and around the GPS receiver antenna and therefore the impact of local multipath on retrieved refractivity, etc., can only be generalized to some degree.

From eq. (A.3.2), the peak phase delay for the jth multipath signal is A_j/A in radians or $\lambda A_j/(2\pi A)$ in units of length for $A_j << A$. Using a reasonable antenna rejection in the aft direction such that $A_j/A = 0.1$, the peak and rms multipath phase delays are 3 and 2 mm. The formation of the ionosphere-free linear combination to remove ionospheric effects magnifies multipath phase errors by approximately the same factor of 3 as for thermal noise increasing the rms error to ~6 mm (Section 3.7). Differencing schemes to remove clock error will also increase multipath noise because each transmitter to receiver link used is contaminated by local multipath (Section 3.8). If two links are used to remove the receiver clock instability, another factor of square root of two increase in rms phase

error can be expected bringing the rms error to ~10 mm. A double differencing scheme can introduce multipath observed at ground receiving sites, which while generally low in multipath amplitude includes scattering from structures significantly farther from the receiving antenna than is possible on a spacecraft. Therefore the maximum fourier frequencies and the erroneous atmospheric Doppler shifts can be larger. Antennas with higher gain and better multipath rejection are under consideration for future missions.

A.4 Effect of Incorrect Radius of Curvature

The error associated with horizontal refractivity structure when the Abel transform is used has been an issue since the Pioneer 10 and 11 Jovian occultation measurements (Kliore, 1974; Eshleman, 1975; Kliore et al., 1976). In this appendix we examine the impact of a simple horizontal error due to an incorrect radius of curvature and derive eq. (3.11.1).

The angles in eq. (2.2.9) can be written in terms of β, the angle between the straight line signal path and the spacecraft velocity vector, γ, the angle between the radial direction and the straight line signal path, and α, the additional angle due to the bending. Eqs. (2.2.9) and (2.2.10) can be rewritten as

$$\Delta f = \frac{f_t}{c}\left\{ V_t\left[\cos(\beta_t - \alpha_t) - \cos(\beta_t)\right] - V_r\left[\cos(\beta_r - \alpha_r) - \cos(\beta_r)\right]\right\} \qquad \text{(A.4.1)}$$

$$a = n_t R_t \sin(\gamma_t + \alpha_t) = n_r R_r \sin(\gamma_r + \alpha_r) \qquad \text{(A.4.2)}$$

where V_t and V_r are the magnitude of $\boldsymbol{V_t}$ and $\boldsymbol{V_r}$ respectively, R_t and R_r correspond to the distance from the center of curvature to the transmitter and receiver respectively, and n_t and n_r are the indices of refraction at the transmitter and the receiver respectively. In the case of the GPS satellite and a LEO that is above most of the ionosphere, these can be taken as unity. In this case, a corresponds to the asymptotic miss distance as shown in figure

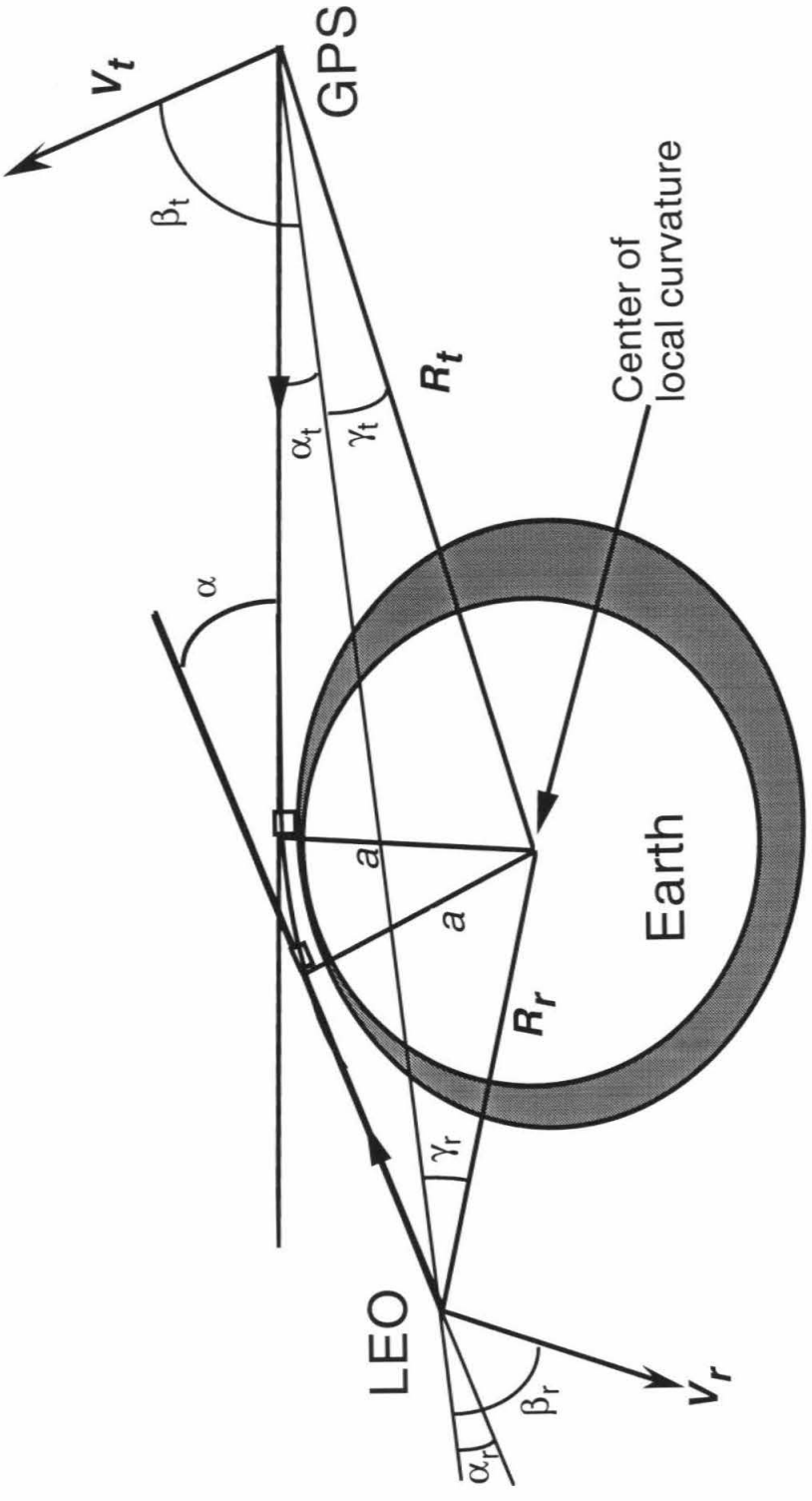

Figure A.4 Occultation geometry for deriving the impact of an incorrect radius of curvature on retrieved refractivity.

A.4 and eqs. (A.4.1)-(A.4.2) provide the two equations necessary to solve for α_t and α_r. The total bending angle $\alpha = \alpha_t + \alpha_r$ can then be expressed as a function of a. All the other variables in eqs. (A.4.1)-(A.4.2) are known from the satellite ephemerides and from the Doppler observations.

Taking eq. (A.4.1) and expanding cos $(\beta_t - \alpha_t)$ for small α_t we get (to first order)

$$\Delta f = \frac{f_t}{c}\left(V_t \, \alpha_t \sin\beta_t - V_r \, \alpha_r \sin\beta_r \right) \qquad \text{(A.4.3)}$$

Now as shorthand define $V_{t\perp} = V_t \sin \beta_t$ and $V_{r\perp} = V_r \sin \beta_r$ so

$$\Delta f = \frac{f_t}{c} \left(V_{t\perp} \, \alpha_t - V_{r\perp} \, \alpha_r \right) \qquad \text{(A.4.4)}$$

$V_{t\perp}$ and $V_{r\perp}$ are known as plane of sky velocities because they are the velocity component in the plane of the sky as viewed by the opposite end of the link. Figure A.4 defines the meaning of the signs.

We can write a similar first order expression from eq. (A.4.2).

$$n_t R_t \left(sin \, \gamma_t + \alpha_t cos \, \gamma_t \right) = n_r R_r \left(sin \, \gamma_r + \alpha_r cos \, \gamma_r \right) \qquad \text{(A.4.5)}$$

Defining the following terms to simplify notation,

$$b_1 = n_t \, R_t \, sin \, \gamma_t - n_r \, R_r \, sin \, \gamma_r$$

$$b_2 = -n_t \, R_t \, cos \, \gamma_t$$

$$b_3 = n_r \, R_r \, cos \, \gamma_r$$

$$b_4 = V_{t\perp} \, f_t / c$$

$$b_5 = -V_{r\perp} \, f_t / c$$

eq. (A.4.5) can then be written as

$$b_1 = b_2 \, \alpha_t + b_2 \, \alpha_r$$

Combining eqs. (A.4.4) and (A.4.5) yields linear expressions for α_t and α_r:

$$\alpha_t = \frac{\Delta f - \frac{b_5 b_1}{b_3}}{b_4 - \frac{b_5 b_2}{b_3}} \quad \text{and} \quad \alpha_r = \frac{\Delta f - b_4 \alpha_t}{b_5}$$

so

$$\alpha = \alpha_t + \alpha_r = \alpha_t\left(1 - \frac{b_4}{b_5}\right) + \frac{\Delta f}{b_5}$$

$$\alpha = \frac{\frac{\Delta f}{b_5} - \frac{b_1}{b_3}}{-\frac{V_{t\perp}}{V_{r\perp}} - \frac{b_2}{b_3}}\left(1 - \frac{V_{t\perp}}{V_{r\perp}}\right) - \frac{c\Delta f}{V_{r\perp} f_t}$$

$$\alpha = -\Delta f\left[\frac{\left(1 + \frac{V_{t\perp}}{V_{r\perp}}\right)}{b_5\left(\frac{V_{t\perp}}{V_{r\perp}} + \frac{b_2}{b_3}\right)} + \frac{c}{V_{r\perp} f_t}\right] + \frac{b_1\left(1 + \frac{V_{t\perp}}{V_{r\perp}}\right)}{b_3\left(\frac{V_{t\perp}}{V_{r\perp}} + \frac{b_2}{b_3}\right)}$$

In the case where $n_t = n_r$, $b_1 = 0$ and

$$\alpha = -\frac{\Delta f\, c}{f_t\, V_{r\perp}}\left[\frac{(V_{r\perp} + V_{t\perp})}{-V_{t\perp} + V_{r\perp}\frac{R_t \cos\gamma_t}{R_r \cos\gamma_r}} + 1\right] \qquad \text{(A.4.6)}$$

$$\alpha = -\frac{\Delta f\, c}{f_t\, V_{r\perp}}\left[1 + \frac{\left(V_{r\perp} + V_{t\perp}\right)}{V_{r\perp}\frac{L_t}{L_r} - V_{t\perp}}\right]$$

where $L_t = R_t \cos\gamma_t$ and $L_r = R_r \cos\gamma_r$.

A similar argument can be made for the asymptotic miss distance, a, such that

$$a \equiv n_t R_t\left[\sin\gamma_t + \frac{\left(\Delta f - {}^{b_5 b_4}/_{b_3}\right)}{b_4 - {}^{b_5 b_2}/_{b_3}}\cos\gamma_t\right]$$

The equivalent first order expression for a in the case where $n_t = n_r$, (so $b_1 = 0$) is

$$a = n_t R_t \sin\gamma_t \left[1 + \frac{\Delta f\, c}{f_t V_{t\perp}} \frac{1}{\tan\gamma_t \left[1 - V_{r\perp} L_t / V_{t\perp} L_r \right]} \right] \qquad \text{(A.4.7)}$$

When $V_{r\perp}$ and $V_{t\perp}$ have the same sign, $\left(\tan\gamma_t \left[1 - V_{r\perp} L_t / V_{t\perp} L_r \right] \right)^{-1}$ ~ -0.5 whereas, when $V_{r\perp}$ and $V_{t\perp}$ have opposite signs, the same term is ~ 0.4. Eqs. (A.4.6) and (A.4.7) can be used to map errors in frequency and geometry to errors in bending angle and asymptotic miss distance.

Error in radius of curvature only

A change in the radius of curvature in the y axis only leaves the straight-line lengths from spacecraft to limb, L_t and L_r, unchanged. The radius to the straight-line tangent height, R_{tan}, is changed by δR_{tan}. The radii of the spacecraft, R_t and R_r, measured relative to the center of curvature are changed.

$$R_t' = (L_t^2 + (R_{tan} + \delta R_{tan})^2)^{1/2} = R_t (1 + R_{tan}\, \delta R_{tan} / R_t^2)$$

$$\sin\gamma_t' = \sin\gamma_t (1 + \delta R_{tan} / R_{tan}) / ((1 + R_{tan}\, \delta R_{tan} / R_t^2)$$

$$\tan\gamma_t' = \tan\gamma_t (1 + \delta R_{tan} / R_{tan})$$

There is no change in $V_{t\perp}$ or $V_{r\perp}$ and, from eq. (A.4.6), there is no error in α associated with changing the radius of curvature along the y-axis alone. There is an error in the derived value of a, the impact parameter.

$$a' = R_{tan} \left(1 + \frac{\delta R_{tan}}{R_{tan}} \right) \left(1 + \frac{c\Delta f}{V_{t\perp} f_t} \left(\tan\gamma_t \left(1 + \frac{\delta R_{tan}}{R_{tan}} \right) \left(1 - \frac{V_{r\perp} R_t \cos\gamma_t}{V_{t\perp} R_r \cos\gamma_r} \right) \right)^{-1} \right)$$

$$\delta a = a' - a$$

$$\delta a = \delta R_{\tan}\left(1 - \frac{\delta R_{\tan}}{R_{\tan}}\frac{c\Delta f}{V_{t\perp}f_t}\left(\tan\gamma_t\left(1 - \frac{V_{r\perp}R_t\cos\gamma_t}{V_{t\perp}R_r\cos\gamma_r}\right)\right)^{-1}\right)$$

So to an error of order 0.1% or less, $\delta a = \delta R_{tan}$.

The error in a must be mapped through the inverse Abel transform defined in eq. (2.2.7)

$$\ln\left(n'(r'_o)\right) = \frac{1}{\pi}\int_{a_1'}^{\infty}\frac{\alpha(a')\,da'}{\sqrt{a'^2 - a_1'^2}}$$

where $a' = a + \delta a$. α is correct but it is assigned to an incorrect value of the impact parameter, a. Now with a change of variable, $a' = a_1' + \Delta a'$, eq. (2.2.7) becomes

$$\ln\left(n'(r'_o)\right) = \frac{1}{\pi}\int_{0}^{\infty}\frac{\alpha\,d\Delta a'}{\sqrt{\left(2a_1' + \Delta a'\right)\Delta a'}} \qquad \text{(A.4.8)}$$

Δa and Δa' are related as

$$\Delta a' = a' - a_1' = a + \delta a - (a_1 + \delta a_1) = \Delta a + \delta a - \delta a_1$$

so eq. (A.4.8) can be rewritten as

$$\ln\left[n'(r'_o)\right] = \frac{1}{\pi}\int_{0}^{\infty}\frac{\alpha\,d\Delta a'}{\sqrt{\left(2\left(a_1 + \delta a_1\right) + \Delta a + \delta a - \delta a_1\right)\left(\Delta a + \delta a - \delta a_1\right)}}$$

$$\ln\left[n'(r'_o)\right] = \frac{1}{\pi}\int_{0}^{\infty}\frac{\alpha\,d\Delta a}{\sqrt{\left(2a_1 + \Delta a\right)\Delta a}}\frac{\left(1 + \frac{d(\delta a)}{d\Delta a}\right)}{\sqrt{\left(1 + \frac{\delta a + \delta a_1}{2a_1 + \Delta a}\right)\left(1 + \frac{\delta a - \delta a_1}{\Delta a}\right)}}$$

$$\ln\left[n'(r'_o)\right] \cong \frac{1}{\pi}\int_0^{\infty} \frac{\alpha\, d\Delta a}{\sqrt{\left(2a_1+\Delta a\right)\Delta a}} \frac{\left(1+\frac{d(\delta a)}{d\Delta a}\right)\left(1-\frac{\delta a+\delta a_1}{2\left(2a_1+\Delta a\right)}\right)}{\sqrt{\left(1+\frac{\delta a-\delta a_1}{\Delta a}\right)}} \qquad \text{(A.4.9)}$$

In this form the Abel integrand is scaled by the factor on the right. Notice that the Abel transform is very heavily weighted to the region near $\Delta a = 0$. Therefore the effect of Δa is limited to the region near $a = a_1$.

If δa is constant then $d\delta a/d\Delta b = 0$ and $\delta a = \delta a_1$ such that

$$\ln\left[n'(r'_o)\right] \cong \frac{1}{\pi}\int_0^{\infty} \frac{\alpha\, d\Delta a}{\sqrt{\left(2a_1+\Delta a\right)\Delta a}}\left(1-\frac{\delta a}{2a_1+\Delta a}\right)$$

Since $(a + a_1)^{-1}$ changes much more slowly than Δa^{-1} and because the Abel transform is heavily weighted toward the tangent region, the term on the right side can be taken outside of the integral resulting in

$$\ln\left[n'(r'_o)\right] \cong \left(1-\frac{\delta a}{2a_1}\right)\frac{1}{\pi}\int_0^{\infty} \frac{\alpha\, d\Delta a}{\sqrt{\left(2a_1+\Delta a\right)\Delta a}}$$

Since $\ln(1+N) \simeq N$ to first order and $N'(r_0) = N(r_0) + \delta N_0'$, for constant δa

$$\delta N_0'/N(r_0) = -\delta a/2a_1$$

There is an additional error because the derived r is in error. r is estimated from $a = nr$ so the error in r, δr, is

$$\delta r/r = \delta a/a - \delta n/n = \delta a/a - \delta N_0'/N(r_0)\ N(r_0)/n = \delta a/a\ (1 + N(r_0)/2n)$$

$$\frac{\delta r}{r} = \frac{\delta a}{a}\left(1+\frac{N}{2}\right)$$

$$\delta r = \frac{r\,\delta a}{a}\left(1+\frac{N}{2}\right) = \frac{r\,\delta a}{r\left(1+N\right)}\left(1+\frac{N}{2}\right) \cong \delta a\left(1-\frac{N}{2}\right)$$

The "1" term compensates for the shift in the coordinate system in the y direction by δR_{tan} = δa. The additional adjustment is $-\delta a\, N/2 \equiv \delta \Delta r$ is the error in the height assigned to the derived N. To find the total error, we must compare N' at its derived radius, r_0', with the true N at r_0': $N_{true}(r_0+\delta\Delta r) = N(r_0) \exp(-\delta\Delta r/H_N)$ where r_0 is the true radius at the true tangent point associated with the true a_1 and H_N is the local refractivity scale height. The resulting error in refractivity is

$$\delta N = N'(r_0') - N_{true}(r_0') = N(r_0)\ (1 + \delta N_0'/N(r_0) - \exp(-\delta\Delta r/H_N))$$

which to first order is

$$\delta N = N(r_o)\ (1 + \delta N_o'/N(r_0) - 1 + \delta\Delta r/H_N) = -N(r_o)\ (\delta a/2a_1 + \delta a\ N/2H_N)$$

The total fractional error in N is expressed as

$$\frac{\delta N}{N_{true}(r_0')} = \frac{N'(r_0') - N_{true}(r_0')}{N_{true}(r_0')} \cong -\frac{\delta a\, N(r_0)}{N(r_0)\left(1-\frac{\delta\Delta r}{H_N}\right)}\left(\frac{1}{2a}+\frac{N(r_0)}{2H_N}\right)$$

$$\frac{\delta N}{N_{true}} \cong -\frac{\delta a}{2a}\left(1+\frac{aN}{H_N}\right)\left(1+\frac{\delta\Delta r}{H_N}\right)$$

$$\frac{\delta N}{N_{true}} \cong -\frac{\delta a}{2a}\left(1+\frac{aN}{H_N}-\frac{\delta aN}{2H_N}\right) \cong -\frac{\delta a}{2a}\left(1+\frac{aN}{H_N}\right) \qquad (3.11.1)$$

Note that N here is unitless. If N is in units of N-units, the N in eq. (3.11.1) must be divided by 10^6. The magnitude of aN/H_N can be estimated using some near-surface values with a = 6372 km, H_N = 6 km and $N = 3\times10^{-4}$, which results in $aN/H_N \sim 0.3$. Near the surface, the refractivity error increased by ~30% due to the incorrect radius so the error in radius can be significant near the surface. Notice also that to get an error of 1% in refractivity requires an error in radius of curvature of order 2% or ~120 km.

A.5 Equatorial Wave Equations

Relationships between k, m and ω and the phase and group velocities

Variable	Kelvin	RGW	IGW	Rossby
m	$-\frac{Nk}{\omega}$	$-\mathrm{sgn}\,(\omega)\,\frac{N}{\omega^2}\left(\beta+\omega\,k\right)$	$\frac{-\mathrm{sgn}(\omega)N\beta}{\omega^2}\left[\left(n+\frac{1}{2}\right)+\left\{\left(n+\frac{1}{2}\right)^2+\frac{\omega k}{\beta}\left[1+\frac{\omega k}{\beta}\right]\right\}^{1/2}\right]$	$\frac{N\beta}{\omega^2}\left[\left(n+\frac{1}{2}\right)-\left\{\left(n+\frac{1}{2}\right)^2+\frac{\omega k}{\beta}\left[1+\frac{\omega k}{\beta}\right]\right\}^{1/2}\right]$
k	$-\frac{\omega m}{N}$	$-\frac{m\lvert\omega\rvert}{N}-\frac{\beta}{\omega}$	$-\frac{\beta}{2\omega}+\sqrt{\left(\frac{\beta}{2\omega}\right)^2+\left(\frac{m\,\omega}{N}\right)^2-(2n+1)\frac{\beta\lvert m\rvert}{N}}$	$-\frac{\beta}{2\omega}\pm\sqrt{\left(\frac{\beta}{2\omega}\right)^2+\left(\frac{m\,\omega}{N}\right)^2-(2n+1)\frac{\beta\lvert m\rvert}{N}}$
$\mathbf{c_x=\frac{\omega}{k}}$	$-\frac{N}{m}$	$-\left[\frac{m}{\mathrm{sgn}\,(\omega)\,N}+\frac{\beta}{\omega^2}\right]^{-1}$	$\frac{\frac{2\,\omega^2}{\beta}}{-1+\sqrt{1+\left(\frac{2\omega}{N\beta}\right)^2\left[m^2\omega^2-(2n+1)N\beta\lvert m\rvert\right]}}$	$\frac{\frac{2\,\omega^2}{\beta}}{-1\pm\sqrt{1+\left(\frac{2\omega}{N\beta}\right)^2\left[m^2\omega^2-(2n+1)N\beta\lvert m\rvert\right]}}$
$\mathbf{c_{gx}=\frac{\partial\omega}{\partial k}}$	$-\frac{N}{m}$	$\left[\frac{m}{-\mathrm{sgn}\,(\omega)\,N}+\frac{\beta}{\omega^2}\right]^{-1}$	$\frac{2\,k+\frac{\beta}{\omega}}{2\,\omega\left(\frac{m}{N}\right)^2+\frac{\beta\,k}{\omega^2}}$	$\frac{2\,k+\frac{\beta}{\omega}}{2\,\omega\left(\frac{m}{N}\right)^2+\frac{\beta\,k}{\omega^2}}$
$\mathbf{c_z=\frac{\omega}{m}}$	$-\frac{Nk}{m^2}$	$\frac{-\mathrm{sgn}\,(\omega)\,\omega^3}{N\left(\beta+\omega\,k\right)}$	$\frac{\frac{-\omega^3}{\mathrm{sgn}(\omega)\,N\,\beta}}{\left(n+\frac{1}{2}\right)+\left[\left(n+\frac{1}{2}\right)^2+\frac{\omega k}{\beta}\left[1+\frac{\omega k}{\beta}\right]\right]^{1/2}}$	$\frac{\frac{\omega^3}{N\,\beta}}{\left(n+\frac{1}{2}\right)-\left[\left(n+\frac{1}{2}\right)^2+\frac{\omega k}{\beta}\left[1+\frac{\omega k}{\beta}\right]\right]^{1/2}}$
$\mathbf{c_{gz}=\frac{\partial\omega}{\partial m}}$	$\frac{Nk}{m^2}$	$\frac{\mathrm{sgn}\,(\omega)\,\omega^3}{N\left[2\,\beta+k\,\omega\right]}$	$\frac{(2n+1)\,N\beta\,\mathrm{sgn}(m)-2\,m\,\omega^2}{\beta k\left(\frac{N}{\omega}\right)^2+2\,\omega\,m^2}$	$\frac{(2n+1)\,N\beta-2\,m\,\omega^2}{\beta k\left(\frac{N}{\omega}\right)^2+2\,\omega\,m^2}$

Relationship between geopotential, temperature and wind perturbations

	Kelvin	RGW	IGW and Rossby Wave
$\hat{\Phi}$	$\hat{\Phi}_o\, e^{-(1/2)\eta^2}$	$i\hat{v}_o\, \omega y\, e^{-(1/2)\eta^2}$	$i\hat{v}_o \left(\frac{\beta N^3}{\lvert m\rvert}\right)^{1/2} \left[\frac{\frac{1}{2} H_{n+1}(\eta)}{\lvert m\rvert\omega - Nk} - \frac{n\, H_{n-1}(\eta)}{\lvert m\rvert\omega + Nk}\right] e^{-(1/2)\eta^2}$
$\frac{\hat{T}}{T}$	$\frac{\hat{\Phi}_o}{g}\left[\frac{1}{2H} + im\right] e^{-(1/2)\eta^2}$	$\hat{v}_o \frac{\omega y}{g}\left[\frac{i}{2H} - m\right] e^{-(1/2)\eta^2}$	$\frac{\hat{v}_o}{g} \left(\frac{\beta N^3}{\lvert m\rvert}\right)^{1/2} \left[\frac{\frac{1}{2} H_{n+1}(\eta)}{\lvert m\rvert\omega - Nk} - \frac{n\, H_{n-1}(\eta)}{\lvert m\rvert\omega + Nk}\right] \left(\frac{i}{2H} - m\right) e^{-(1/2)\eta^2}$
$\hat{v}$	-	$\hat{v}_o\, e^{-(1/2)\eta^2}$	$\hat{v}_o\, H_n(\eta)\, e^{-(1/2)\eta^2}$
$\hat{\mathbf{u}}$	$\hat{\Phi}_o \frac{k}{\omega} e^{-(1/2)\eta^2}$	$\hat{v}_o \frac{i\lvert m\rvert \omega y}{N} e^{-(1/2)\eta^2}$	$i\hat{v}_o \left(\beta\lvert m\rvert N\right)^{1/2} \left[\frac{\frac{1}{2} H_{n+1}(\eta)}{\lvert m\rvert\omega - Nk} + \frac{n\, H_{n-1}(\eta)}{\lvert m\rvert\omega + Nk}\right] e^{-(1/2)\eta^2}$
$\hat{w}$	$\frac{\hat{\Phi}_o\, \omega}{N^2}\left[-m + \frac{i}{2H}\right] e^{-(1/2)\eta^2}$	$-\hat{v}_o \frac{\omega^2 y}{N^2}\left[i\, m + \frac{1}{2H}\right] e^{-(1/2)\eta^2}$	$-\hat{v}_o \left[i\, m + \frac{1}{2H}\right] \omega \left(\frac{\beta}{\lvert m\rvert N}\right)^{1/2} \left[\frac{\frac{1}{2} H_{n+1}(\eta)}{\lvert m\rvert\omega - Nk} - \frac{n\, H_{n-1}(\eta)}{\lvert m\rvert\omega + Nk}\right] e^{-(1/2)\eta^2}$

For all waves: $\hat{w} = -\frac{\omega}{N^2}\left(m - \frac{i}{2H}\right)\hat{\Phi}$ and $\frac{\hat{T}}{T} = \frac{1}{g}\left[\frac{1}{2H} + im\right]\hat{\Phi}$

β is $2\Omega \cos\phi / R_{Earth}$ and Ω is the angular rotation rate of Earth, ϕ is the latitude, R_{Earth} is Earth's radius

N is the buoyancy frequency

A.6 Net Vertical Displacement by Equatorial Waves

The Lagrangian approach

A simple explanation for the systematic upward displacement near the equator is due to the phase relationship between the meridional and vertical winds. Following AHL, assuming N to be constant, a class of equatorial waves exist of the general form,

$$\left(u',v',w',\Phi'\right) = e^{z/2H}\,\mathrm{Re}\left\{\left[\hat{u}(y),\hat{v}(y),\hat{w}(y),\hat{\Phi}(y)\right]e^{i(kx+mz-\omega t)}\right\} \tag{A.6.1}$$

where u', v', w' and Φ' are the wave perturbations to zonal, meridional and vertical winds and geopotential respectively. For the RGW,

$$\hat{w} = -\frac{\omega}{N^2}\left(m-\frac{i}{2H}\right)\hat{\Phi} = -\hat{v}_0\,\frac{\omega}{N^2}\left(m-\frac{i}{2H}\right)i\omega y\exp\left[\frac{-\beta|m|y^2}{2N}\right] \tag{A.6.2}$$

so

$$w' = -e^{z/2H}\,e^{-by^2}\,\hat{v}_0\,\frac{\omega^2}{N^2}\,\mathrm{Re}\left\{i\left(m-\frac{i}{2H}\right)y\,e^{i(kx+mz-\omega t)}\right\} \tag{A.6.3}$$

where $b = \dfrac{\beta|m|}{2N}$. The key is the variation in y over the oscillation of the wave. The meridional particle velocity is

$$v' = e^{z/2H}\,e^{-by^2}\,\hat{v}_0\,\mathrm{Re}\left\{e^{i(kx+mz-\omega t)}\right\} = e^{z/2H}\,e^{-by^2}\,\hat{v}_0\cos\left(kx+mz-\omega t\right) \tag{A.6.4}$$

The change in y, the meridional component of the parcel's location, is the integral of eq. (A.6.4) over the course of the wave's propagation. Equatorial wave amplitudes are generally not large enough to produce zonal and vertical particle velocities which significantly alter the kx and mz phase terms over a wave period. Therefore integration of eq. (A.6.4) with respect to time gives the first order change in y over a wave cycle:

$$y \cong y_0 + \int_0^t v'\,dt' = y_0 - e^{z/2H}\, e^{-b\,y^2}\, \frac{\hat{v}_0}{\omega} \sin\left(kx + mz - \omega t\right) \tag{A.6.5}$$

Taking a parcel initially at the equator ($y_0 = 0$) and plugging eq. (A.6.5) into eq. (A.6.3) yields

$$w' = e^{z/H}\, e^{-2b\,y^2}\, \hat{v}_0^2\, \frac{\omega}{N^2} \sin\left(kx + mz - \omega t\right) \mathrm{Re}\left\{ i \left(m - \frac{i}{2H} \right) e^{i\left(kx + mz - \omega t\right)} \right\}$$

$$= e^{z/H}\, \frac{\hat{v}_0^2\, \omega}{N^2} \left\{ \frac{\sin\left(\phi - \omega t\right) \cos\left(\phi - \omega t\right)}{2H} - m \sin^2\left(\phi - \omega t\right) \right\} \tag{A.6.6}$$

where ϕ is kx + mz. Eq. (A.6.6) is the approximate vertical velocity within several degrees of the equator. With the dominant RGW vertical wavelength between 4 and 8 km, 1/2H << m and the amplitude of the sin*cos term will be much less than that of the $\sin^2$ term. More importantly, the contribution of the sin*cos term averages to zero over a full cycle of the wave while the $\sin^2$ term averages to 0.5. The nonzero mean of the $\sin^2$ term causes a net vertical displacement of air parcels near the equator during the passage of a RGW independent of the wave's phase.

The assumption that the vertical component of group velocity is upward (consistent with a tropospheric source) which requires ω and m to have opposite signs forces (via the second term on the RHS of eq. (A.6.6)) the average vertical velocity, $\overline{w'}$, to be **positive** and lower stratospheric air near the equator must flow **upward** on average during the passage of a RGW. The approximate magnitude of the flow is given by eq. (A.6.7).

$$\overline{w'} \cong e^{z/H}\, \frac{\hat{v}_0^2 \left| \omega\, m \right|}{N^2} \overline{\sin^2\left(\phi - \omega t\right)} = e^{z/H}\, \frac{\hat{v}_0^2 \left| \omega\, m \right|}{2\, N^2} \tag{A.6.7}$$

Eulerian residual circulation

The vertical component of residual mean meridional circulation is

$$\overline{w}^* \equiv \overline{w} + \frac{\left(\cos\phi\, \overline{v'\theta'} / \overline{\theta}_z \right)_\phi}{a \cos\phi} \tag{A.6.8}$$

where ϕ is latitude and a is the radius of the Earth. The perturbation in potential temperature, θ', is given as

$$\theta' = \frac{He^{\kappa z/H}}{R}\frac{\partial \Phi'}{\partial z}$$

$$= \frac{He^{\kappa z/H}}{R}\frac{\partial \Phi'}{\partial z}\hat{v}_0 \exp\left[-\frac{\beta|m|y^2}{2N}\right]e^{z/2H}\,\mathrm{Re}\left\{i\omega y\left(\frac{1}{2H}+im\right)e^{i(kx+mz-\omega t)}\right\} \qquad \text{(A.6.9)}$$

Combining eq. (A.6.9) with the expression for v' in Table A.5 yields

$$\overline{v'\theta'} = -\frac{\hat{v}_0^2\omega y H}{R}\exp\left[-\frac{\beta|m|y^2}{2N}\right]e^{(1+\kappa)z/H}\left(m\,\overline{\cos^2(kx+mz-\omega t)}+\frac{\overline{\sin(kx+mz-\omega t)\cos(kx+mz-\omega t)}}{2H}\right)$$

which simplifies to

$$\overline{v'\theta'} = -\frac{\hat{v}_0^2}{2}\frac{\omega y m H}{R}\exp\left[-\frac{\beta|m|y^2}{2N}\right]e^{(1+\kappa)z/H} \qquad \text{(A.6.10)}$$

because sin and cos are orthogonal. Assuming $\bar{w}$ is 0 and plugging eq. (A.6.10) into eq. (A.6.8) yields

$$\bar{w}^* = -\frac{g\sin\phi\,\overline{v'\theta'}}{\bar{\theta}N^2}-\frac{\hat{v}_0^2\,\omega\,m}{2N^2}\exp\left[-\frac{\beta|m|y^2}{2N}\right]e^{z/H}\left(1-\frac{2\beta|m|y^2}{N}\right) \qquad \text{(A.6.11)}$$

which at the equator equals

$$\bar{w}^*_{y=0} = -\frac{\hat{v}_0^2\,\omega\,m}{2N^2}e^{z/H} \qquad \text{(A.6.12)}$$

Since $\omega\ \alpha$ -m, eq. (A.6.12) is identical to eq. (A.6.7), confirming the residual mean vertical flow is the same in both Lagrangian and Eulerian representations at least at the equator.

References

Ahnert, P.R., Precision and comparability of National Weather Service upper air measurements, in *Seventh Symposium on Meteorological Observations and Instrumentation,* New Orleans, LA, 221-227, 1991.

Albrecht, B.A., V. Ramanathan and B.A. Boville, The effects of cumulus moisture transports on the simulation of climate with a general circulation model. *J. Atmos Sci., 43*, 2443-2462, 1986.

Allison, M., Planetary waves in Jupiter's equatorial atmosphere, *Icarus, 83*, 282-307, 1990.

Andersson E., A. Hollingsworth, G. Kelly, P. Lönnberg, J. Pailleux and Z. Zhang, Global Observing system experiments on operational statistical retrievals of satellite sounding data, *Mon. Weath. Rev., 119*, 1851-1864, 1991.

Andrews, D.G., J.R. Holton and C.B. Leovy, *Middle Atmosphere Dynamics*, Academic Press, Orlando, FL (1987).

Angell, J.K., Variations and trends in tropospheric and stratospheric global temperatures, 1958-1987, *J. Clim., 1*, 1296-1313, 1988.

Augstein, E., H. Schmidt and F. Ostapoff, The vertical structure of the atmospheric planetary boundary layer in undisturbed trade wind over the Atlantic Ocean, *Boundary-Layer Meteorol., 6*, 129-150, 1974.

Barnes, J. A., A.R. Chi, L.S. Cutler, D.J. Healey, D.B. Leeson, T.E. McGunigal, J.A. Mullen, W.L. Smith, R.L. Sydnor, R.F.C. Vessot and G.M.R. Winkler, Characterization of Frequency Stability, *IEEE Transactions on Instrumentation and Measurement, IM-20,* 105-120, 1971.

Barnett, T.P., and M.E. Schlesinger, Detecting changes in global climate induced by greenhouse gases, *J. Geophys. Res., 92*, 14772-14780, 1987.

Bassiri, B., and G.A. Hajj, Higher-order ionospheric effects on the global positioning system observables and means of modeling them, *manuscripta geodaetica, 18*, 280-289, 1993.

Belloul, B., and A. Hauchecorne, Horizontal homogeneities in occultation methods: the case of atmospheric gravity waves, submitted to *Radio Sci.,* 1995.

Bertiger, W.I., Y.E. Bar-Sever, E.J. Christensen, E.S. Davis, J.R. Guinn, B.J. Haines, R.W. Ibanez-Meier, J.R. Jee, S.M. Lichten, W.G. Melbourne, R.J. Muellerschoen, T.N. Munson, Y. Vigue, S.C. Wu, T.P. Yunck, B.E. Schutz, P.A.M. Abusali, H.J. Rim, M.M. Watkins, and P. Willis, GPS Precise Tracking Of Topex/Poseidon: Results and Implications, *J. Geophys. Res., 99*, 24449-24464, 1994.

Bertiger, W.I., and S. Wu, "Single Frequency GPS Orbit Determination For Low Earth Orbiters", Proceedings of the ION National Tech. Meeting, Santa Monica, CA, Jan. 1996.

Betts, A.K., and W. Ridgway, Climatic equilibrium of the atmospheric convective boundary layer over a tropical ocean, *J. Atmos. Sci., 46*, 2621-2641, 1989.

Betts, A.K., R.L. Desjardins and J.I. MacPherson, Budget analysis of the boundary layer grid flights during FIFE 1987, *J. Geophys. Res., 97*, 18533-18546, 1992.

Boering, K.A., S.C. Wolfsy, B.C. Daube, H.R. Schneider, M. Loewenstein, J.D. Podolske and T.J. Conway, Stratospheric mean ages and transport rates from observations of carbon dioxide and nitrous oxide, *Science, 274*, 1340-1343, 1996.

Bolton, D., The Computation of Equivalent Potential Temperature, *Mon. Weath. Rev., 108*, 1046-1053, 1980

Born, M., and E. Wolf, *Principles of Optics* , p.122, Sixth edition, Pergamon Press, 1980.

Cess, R.D., G.L. Potter, J.P. Blanchet, G.J. Boer, A.D. Delgenio, M. Deque', V. Dymnikov, V. Galin, W.L. Gates, S.J. Ghan, J.T. Kiehl, A.A. Lacis, H. LeTreut, Z.-X. Li, X.-Z. Liang, B.J. McAvaney, V.P. Meleshko, J.F.B. Mitchell, J.-J. Morcette, D.A. Randall, L. Rikus, E. Roeckner, J.F. Royer, U. Schlese, D.A. Sheinin, A. Slingo, A.P. Sokolov, K.E. Taylor, W.M. Washington, R.T. Wetherald, I. Yagai and M.-H. Zhang, Intercomparison and interpretation of climate feedback processes in 19 atmospheric general circulation models, *J. Geophys. Res., 95*, 16601-16615, 1990.

Chahine, M. T., D. J. McCleese, P. W. Rosenkranz and D. H. Staelin, Interaction mechanisms within the atmosphere, in *Manual of Remote Sensing*, Editor-in-chief R.N. Colwell, pp. 165-230, American Society of Photogrammetry, Falls Church, VA, 1983.

Champion, K.S.W., A.E. Cole and A.J. Kantor, Standard and Reference Atmospheres, in the *Handbook of Geophysics and the Space Environment*, Jursa A.S. (ed), Air Force Geophysics Laboratory, Bedford, MA, 1985.

Danielsen, E.F., In Situ evidence of rapid, vertical irreversible transport of lower tropospheric air into the lower stratosphere by convective cloud turrets and by larger scale upwelling in the tropical cyclones, *J. Geophys Res.*, **98**, 8665-8681, 1993.

Eluszkiewicz, J., D. Crisp, R. Zurek, L. Elson, E. Fishbein, L. Froidevaux, J. Waters, R.G. Grainger, A. Lambert, R. Harwood and G. Peckham, Residual circulation in the stratosphere and lower mesosphere as diagnosed from microwave limb sounder data, *J. Atm. Sci., 53*, 217-240, 1996.

Epstein, E.S., Detecting Climate Change, *J. App. Met., 21*, 1172-1182, 1982.

Eshleman, V.R., "Jupiter's atmosphere: problems and potential of radio occultation", Science, 189, 876-878, 1975.

Eyre, J. R., Assimilation of radio occultation measurements into a numerical weather prediction system, *European Centre for Medium-Range Weather Forecasts,* Technical Memorandum No. 199, May, 1994.

Eyre, J.R., G.A. Kelly, A.P. McNally, E. Andersson and A. Persson, *Quart. J. R. Met. Soc., 119*, 1427, 1993.

Finger, F.G., and F.J. Schmidlin, Upper-air measurements and instrumentation workshop, *Bull. Amer. Meteor. Soc*, *72*, 50-55, 1991.

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

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

Fjeldbo, G., A.J. Kliore, and V.R. Eshleman, The neutral atmosphere of Venus as studied with the Mariner V radio occultation experiments", *Astronom J., 76*, 123-140, 1971.

Flobert, J.F., E. Andersson, A. Chedin, A. Hollingsworth, G. Kelly, J. Pailleux and N.A. Scott, Global data assimilation and forecast experiments using the improved inversion method for satellite soundings, *Mon. Weath. Rev.*, ***119***, 1881-1914, 1991.

Fritts, D.C., R.C. Blanchard and L. Coy, Gravity Wave Structure between 60 and 90 km Inferred from Space Shuttle Reentry Data, *J. Atm. Sci., 46*, 423-434, 1989.

Fritts, D.C., M.A. Geller, B.B. Balsley, M.L. Chanin, I. Hirota, J.R. Holton, S. Kato, R.S. Lindzen, M.R. Schoeberl, R.A. Vincent and R.F. Woodman, Research Status and Recommendations from the Alaska Workshop on Gravity Waves and Turbulence in the Middle Atmosphere Fairbanks, Alaska 18-22 July 1983, *Bull. Amer. Meteor. Soc*, *65*, 149-159, 1984.

Gary, B. L., A novel method for monitoring global warming using unmanned aerial vehicles and GPS receivers, White Paper, Jet Propulsion Laboratory, Pasadena, CA, Dec. 8, 1992.

Gorbunov, M.E., and S.V. Sokolovskiy, Remote sensing of refractivity from space for global observations of atmospheric parameters, *Report 119, Max-Planck-Institut fur Meteorologie*, 1993.

Gurvich, A., and S.V. Sokolovskiy, Diffraction correction of GPS/MET data, *Second URSI International Workshop for Working Group AFG1: Remote Sensing and Atmospheric Research Using Radio Occultation Techniques with the Global Navigation System*, URSI, Tuscon, AZ, February 22-23, 1996.

Hajj, G.A., R. Ibanez-Meier, E.R.Kursinski and L. J. Romans, Imaging the ionosphere with the Global Positioning System, *Int. J. of Imaging Sys. and Tech.*, *5*, 174-184, 1994.

Hajj, G.A., and E.R. Kursinski, Analysis of Errors in the Vertical Temperature Profiles Recovered from GPS Occultation Observations, *EOS Trans. AGU*, *72*, 372, 1991.

Hajj, G. A., E. R. Kursinski, W. I. Bertiger, S. S. Leroy, T. Meehan, L. J. Romans, J. T. Schofield, Initial results of GPS-LEO occultation measurements of Earth's atmosphere obtained with the GPS-MET experiment, IUGG XXI General Assembly, Boulder, Co., July 2-14, 1995.

Hamilton, K. and R. A. Vincent, High-Resolution Radiosonde Data Offer New Prospects for Research, *EOS Trans. AGU*, *76*, 497, 1995.

Hansen, J., A. Lacis, D. Rind, L. Russell, P. Stone, I. Fung, R. Ruedy and J. Lerner, Climate sensitivity analysis of feedback mechanisms, in *Climate Processes and Climate Sensitivity* (ed. J. Hansen and T. Takahashi), Geophysical Monograph 29, American Geophysical Union, Washington D.C., 130-163, 1984.

Hardy, K.R., G.A.Hajj and E.R.Kursinski, Accuracies of atmospheric profiles obtained from GPS occultations, *Intern. J. Sat. Comm.*, *12*, 463-473, 1994.

Hardy, K. R., D. P. Hinson, G. L. Tyler and E. R. Kursinski, Atmospheric profiles from active space-based radio measurements, 6th Conf. on Satellite Meteorology and Oceanography, Amer. Meteorol. Soc., Atlanta, GA, January 5-10, 1992.

Haugstad, B.B., Effects of the inhomogenoeus background on radiation propagating through turbulent planetary atmospheres, *Radio Sci.*, *13*, 435-440, 1978.

Heckley, W.A., Systematic errors in the ECMWF operational forecasting model in tropical regions, *Quart. J. Roy. Meteor. Soc., 11*, 709-738, 1985.

Hinson, D. P., and J. M. Jenkins, Magellan radio occultation measurements of atmospheric waves on Venus, *Icarus, 114*, 310-327, 1995.

Hinson, D. P., and J. A. Magalhaes, Equatorial waves in the stratosphere of Uranus, *Icarus, 94*, 64-91, 1991.

Hinson, D. P., and J. A. Magalhaes, Inertio-Gravity Waves in the Atmosphere of Neptune waves in Neptune, *Icarus, 105*, 142-161, 1993.

Hinson, D. P., and G. L. Tyler, Internal gravity waves in Titan's atmosphere observed by Voyager radio occultation, *Icarus, 54*, 337-352, 1983.

Hoskins, B.J., M.E. McIntyre and A.W. Robertson, On the use and significance of isentropic potential vorticity maps, *Quart .J. Met., 111*, 877-946, 1985.

Houghton, J. T., G. J. Jenkins and J. J. Ephraums, editors of *Climate Change The IPCC Scientific Assessment*, University of Cambridge, Cambridge, UK, 1990.

Janjic, Z.I., The step-mountain coordinate physical package, *Mon. Weath. Rev.*, *118*, 1429-1443, 1990.

Jensen, E.J., O.B.Toon, L. Pfister and H.B.Slekirk, Dehydration of the upper troposphere and lower stratosphere by subvisible cirrus clouds near the tropical tropopause, *Geophys. Res. Lett., 23*, 825-828, 1996.

Karayel, E. T., and D. P. Hinson, Sub-Fresnel-scale vertical resolution in atmospheric profiles from radio occultation, submitted to *Radio Sci.*, 1996.

Karoly, D. J., Southern Hemisphere temperature trends: A possible greenhouse gas effect?, *Geophys. Res. Lett., 14,* 1139-1141, 1987.

Karoly, D.J., Northern Hemisphere temperature trends: A possible greenhouse gas effect?, *Geophys. Res. Lett., 16*, 465-468, 1989.

Kasahara, A., and A.P. Mizzi, Estimates of tropical analysis differences in daily values produced by two operational centers, *Mon. Weath. Rev.*, *120*, 279-302, 1992.

Kelley, M.C., *The Earth's Ionosphere: Plasma Physics and Electrodynamics*, 487 p., Academic Press, San Diego, 1989.

Kelly G., E. Andersson, A. Hollingsworth, P. Lönnberg, J. Pailleux and Z. Zhang, Quality control of operational physical retrievals of satellite sounding data, *Mon. Weath. Rev.*, *119*, 1866-1880, 1991.

Kerr, R.A., A military navigation system might probe lofty weather, *Science, 256,* 318-319, 1992.

Kliore, A. J., G. Fjeldbo, B. L. Seidel, T. T. Sesplaukis, D. W. Sweetnam and P. M. Woiceshyn, "Atmosphere of Jupiter from the Pioneer 11 S-band Occultation experiment: Preliminary Results", Science, 188,474-476, 1975.

Kliore, A. J., and I. R. Patel, Thermal structure of the atmosphere of Venus from Pioneer Venus radio occultations, *Icarus.*, *52*, 320-334, 1982.

Kliore, A. J., P. M. Woiceshyn and W. B. Hubbard, "Temperature of the atmosphere of Jupiter from Pioneer 10/11 radio occultations", GRL, 3, 113-116, 1976.

Kuo, Y.-H., X. Zou and W. Huang, The impact of GPS data on the prediction of an extratropical cyclone: an observing system simulation experiment, submitted to *J. Dyn. Atmos. Ocean*, February, 1996.

Kursinski, E.R., Monitoring the Earth's atmosphere with GPS, *GPS World*, *5*, 50-54, 1994.

Kursinski, E. R., W. I. Bertiger, G. A. Hajj, S. S. Leroy,T. K. Meehan, D. J. McCleese, L. J. Romans, J. T. Schofield and J. B. Thomas, Observations of tropospheric water vapor by radio occultation using the Global Positioning System, Second International Scientific Conference on the Global Energy and Water Cycle, preprint volume p. 380-381, Washington DC, June 17-21, 1996b.

Kursinski, E. R., G. A. Hajj, W. I. Bertiger, S. S. Leroy, T. K. Meehan, L. J. Romans, J. T. Schofield, D. J. McCleese, W. G. Melborne, C. L. Thornton, T. P. Yunck, J. R. Eyre and R. N. Nagatani, Initial Results of Radio Occultation Observations of Earth's Atmosphere Using the Global Positioning System, *Science*, *271,* 1107-1110, 1996a.

Kursinski, E. R., G. A. Hajj and K. R. Hardy, Moisture profile information from radio occultation measurements, *Eos Trans. AGU , 72,* 372, 1991.

Kursinski, E. R., G. A. Hajj, and K .R. Hardy, Temperature or moisture profiles from radio occultation measurements, *Proc. of the 8th Symp. on Meteorological Observations and Instrumentation*, Am. Met. Soc., pp. J153-J158, Anaheim, CA, Jan. 17-22, 1993a.

Kursinski, E. R., G. A. Hajj, and K. R. Hardy, Atmospheric profiles from radio occultation measurements of GPS Satellites, *Proc. of the SPIE Symp. on Optical Engineering and Photonics in Aerospace Science and Sensing*, paper no. 1935-13, SPIE, Orlando, FL, 1993b.

Kursinski, E. R., G. A. Hajj, and K. R. Hardy, Observing Climate Change with the Global Positioning System, *EOS Trans. AGU*, *75*, 114, 1994.

Kursinski, E. R., G. A. Hajj, K. R. Hardy, L. J. Romans and J. T. Schofield, Observing tropospheric water vapor by radio occultation using the Global Positioning System, *Geophys. Res. Lett.*, *22*, 2365-2368, 1995.

Leroy, S.S., Measurement of geopotential heights by GPS radio occultation, *J. Geophys. Res.*, *102*, 6971-6986, 1997.

Lindal, G.F., The atmosphere of Neptune: an analysis of radio occultation data acquired with Voyager 2, *Astronom. J., 103*, 967-982, 1991.

Lindal, G. F., G. E. Wood, H. B. Hotz, D. N. Sweetnam, V. R. Eshleman and G. L. Tyler, The Atmosphere of Titan: An Analysis of the Voyager 1 Radio Occultation Measurements, *Icarus*, *53*, 248-263, 1983.

Lindzen, R. S., and M. Fox-Rabinovitz, Consistent Vertical and Horizontal Resolution, *Monthly Weather Review, 117*, 2575-2583, 1989.

Linfield, R., The effect of aperture averaging upon tropospheric delay functions seen with a DSN antenna, *The Telecommunications and Data Acquisition Progress Report 42-124*, Jet Propulsion Laboratory, Pasadena, CA, pp.1-7, February 15, 1996.

Liu, Q., and C.J.E. Schuurmans, The correlation of tropospheric and stratospheric temperatures and its effect on the detection of climate changes, *Geophys. Res. Lett., 17*, 1085-1088, 1990.

Luers, J.K., and R.E. Eskridge, Temperature corrections for the VIZ and Vaisala radiosondes, *J. Appl. Met.*, *34*, 1241-1253, 1995.

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

Manabe, S., and R.T. Wetherald, Large scale changes of soil wetness induced by an increase in atmospheric carbon dioxide, *J. Atmos. Sci., 44*, 1211-1235, 1987.

Marouf, E. A., G. L. Tyler and P. A. Rosen, Profiling Saturn's Rings by Radio Occultation, *Icarus, 68*, 120-166, 1986.

McMillin, L.M., M.E. Gelman, A. Sanyal and M. Sylva, A method for the use of satellite retrievals as a transfer standard to determine systematic radiosonde errors, *Mon. Wea. Rev.*, *116*, 1091-1102, 1988.

Mesinger, F., Z. I. Janjic, S. Nickovic, D. Gavrilov, D. G. Deaven, The step-mountain coordinate-model description and performance for cases of alpine lee cyclogensis and for a case of an appalachian redevelopment, *Mon. Weath. Rev.*, *116*, 1493-1518, 1988.

Muruyama, Y., T. Tsuda, M. Yamamoto, T. Nakamura, T. Sato, S. Kato and S. Fukao, Dominant vertical scales of gravity waves in the middle atmosphere observed with the MU radar and rocketsondes, *J. Atm. Terr. Phys., 54*, 339-346, 1992.

National Research Council, *The Global Positioning System, A Shared National Asset: Recommendation for Technical Improvements and Enhancements*, p. 6, National Research Council committee on the future of the Global Positioning System, National Academy Press, Washington D.C. 1995.

Newell, R.E., and S. Gould-Stewart, A Stratospheric Fountain? *J. Atm. Sci., 38*, 2789-2796, 1981.

Oort, A.H., Global atmospheric circulation statistics 1958-1973, *NOAA Prof. Pap. 14*, 180 pp. U.S. Govt. Printing Office, Washington, D.C., 1983.

Owens, J.S., Optical refractive index of air: dependence on pressure, temperature and composition, *Appl. Opt., 6*, 51-58, 1967.

Paluch, I. R., D. H. Lenschow, J. G.Hudson and R. Pearson, Transport and mixing processes in the lower troposphere over the ocean, *JGR, 97*, 7527-7541, 1992.

Papas, C.H., Theory of electromagnetic wave propagation, Magraw-Hill, New York, New York, 1965.

Papoulis, A., *Probability, random variables, and stochastic processes*, second edition, McGraw-Hill, New York, p. 235, 1984.

Parker, D.E., On the detection of temperature changes induced by increasing atmospheric carbon dioxide" *Q.J.R. Met. Soc., 111*, 587-601, 1985.

Peixoto, J.P., and A.H.Oort, *Physics of Climate*, 520 pp., Am. Inst. Phys, New York, 1992.

Phillips, N., J. Susskind and L. McMillin, Results of a joint NOAA/NASA sounder simulation study, *J. Atm.Oceanic Tech., 5*, 44-56, 1988.

Raval, A. and V. Ramanathan, Observational determination of the greenhouse effect, *Nature, 342*, 758-761, 1989.

Remsberg, E. E., J. M. Russell III, L. L. Gordley, P. L. Bailey, W. G. Planet and J. E. Harries, The validation of NIMBUS 7 LIMS measurements of ozone, *J. Geophys Res., 89*, 5161-5178, 1984.

Rich, F.J., and S. Basu, Ionospheric Physics, in *Handbook of Geophysics and the Space Environment*, ed. Jursa A.S. Air Force Geophysics Laboratory, Bedford, Mass., 1985.

Rind, D., E.-W. Chiou, W. Chu, S. Oltmans, J. Lerner, J. Larsen, M. P. McCormick and L. McMaster, Overview of the Stratospheric Aerosol and Gas Experiment II water vapor observations: method, validation and data characteristics, *J. Geophys. Res., 98*, 4835-4856, 1993.

Santer, B. D., T. M. I Wigley, M. E. Schlesinger and P. D. Jones, Multi-variate methods for the detection of greenhouse-gas-induced climate change, in *Greenhouse-Gas-Induced Climatic Change: A Critical Appraisal of Simulatiuons and Observations*, edited by M. E. Schlesinger, Elsevier Science Publishers, Amsterdam, 1990.

Schlesinger, M. E., T. P. Barnett and X. J. Jiang, On greenhouse gas signal detection strategies, in *Greenhouse-Gas-Induced Climatic Change: A Critical Appraisal of Simulations and Observations*, edited by M. E. Schlesinger, Elsevier Science Publishers, Amsterdam, 1990.

Smith, E.K., and S. Weintraub, The constants in the equation for atmospheric refractive index at radio frequencies, *Proc. IRE, 41*, 1035-1037, 1953.

Smith, W.L, .H. M.Woolf, C. M. Hayden, D.Q. Wark, and L. M. McMillan, *Bull. Amer. Met. Soc.* , *60*, 1177, 1979.

Soden, B.J. and F.P. Bretherton, Interpretation of TOVS water vapor radiances in terms of layer-average realtive humidities: Method and climatology for the upper, middle and lower troposphere, *JGR, 101,* 9333-9343, 1996.

Solomon, S., The mystery of the Antarctic ozone "hole", *Rev. Geophys.*, *26*, 131-148, 1988.

Solomon, S., Progress towards a quantitative understanding of Antarctic ozone depletion, *Nature*, *347,* 347-354, 1990.

Spencer, R.W., and J.R. Christy, Precise monitoring of global temperature trends from satellites, *Science*, *247*, 1558-1562, 1990.

Spilker J. J., GPS signal structure and performance characteristics, *Navigation, 25*, 29-54, 1978.

Starr, D.O'C., and S.H.Melfi, editors of The Role of Water Vapor in Climate, A Strategic Research Plan for the Proposed GEWEX Water Vapor Project (GVaP), NASA Conference Publication 3120, 1991.

Sun, D.Z., and R.S. Lindsen, Water vapor feedback and the ice-age snowline record, *Ann-geoph, 11*, 204-215, 1993.

Thayer, G.D., An improved equation for the radio refractive index of air, *Radio Sci.*, *9*, 803-807, 1974.

Treuheft, R.N., and G.E. Lanyi, The effect of the dynamic wet troposphere on radio interferometric measurements, *Radio Sci.*, *22*, 251-265, 1987.

Tyler, G. L., D. N. Sweetnam, J. D. Anderson, J. K. Campbell, V. R. Eshleman, D. P. Hinson, G. S. Levy, G. F. Lindal, E. A. Marouf, and R. A. Simpson, Voyager 2 radio science observations of the Uranian system: Atmosphere, rings, and satellites", *Science*, *233*, 79-84, 1986.

Vorob'ev V.V., and T.G. Krasil'nikova, An estimation of accuracy of recovery of atmospheric refrctivity from measurements of doppler shifts at frequencies used in the Navstar system, *Ivestia AN. Fizica Atmosfery i Okeana*, *29*, 626-633, 1993.

Waliser, D.E., and C. Gautier, A satellite-derived climatology of the ITCZ, *J. Clim., 6*, 2162-2174, 1993

Ware, R., M. Exner, D. Feng, M. Gorbunov, K. Hardy, B. Herman, Y. Kuo, T. K. Meehan, W. G. Melbourne, *et al.*, GPS sounding of the atmosphere from low Earth orbit - preliminary results, *Bull. Am. Met. Soc., 77*, 19-40, 1996.

Wittmeyer, I. L. and T. H. Vonder Haar, Analysis of the global ISCCP TOVS water vapor climatology, *J. Clim., 7*, 325-333, 1994.

Wu, S.C., W.I.Bertiger and J.T.Wu, Minimizing Selective Availability Error on Topex GPS Measurements, AIAA/AAS Astrodynamics Conference, August 20-22, Portland, Oregon, Paper 90-2942, 1990.

Wu, D.L., and J.W. Waters, Satellite observations of atmospheric gravity waves, submitted to *Geophys. Res. Lett.*, 1996.

Yeh, K.C., and C.H. Liu, Radiowave scintillations in the ionosphere, *Proc. IEEE, 70*, 324-360, 1982.

Yuan, L. L., R. A. Anthes, R. H. Ware, C. Rocken, W. D. Bonner, M. G. Bevis and S. Businger, Sensing climate change using the global positioning system, *JGR*, *98*, 14925-14937, 1993.

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

Zou, X., Y.-H. Kuo and Y.-R. Guo, Assimilation of atmospheric radio refractivity using a nonhydrostatic adjoint model, *Mon. Wea. Rev., 123*, 2229-2249, 1995.