Use of radio occultation for long-term tropopause studies: uncertainties, biases, and instabilities

Paul W. Staten and Thomas Reichler

Department of Meteorology, University of Utah

Corresponding author: Paul W. Staten

Department of Meteorology, University of Utah,
135 S 1460 E, Rm. 819 (WBB), Salt Lake City, UT 84112

e-mail: paul.staten@utah.edu phone: 801-585-1405

May 2008

Submitted to: Journal of Geophysical Research

Special section: Diagnosis and Modeling of the Tropopause: Structure, Dynamics, and Variability

Keywords: radio occultation, tropopause, climate
Abstract

Research suggests that changes in tropopause structure can both indicate and impact changes in the global climate system. The Global Positioning System radio occultation (RO) technique shows tremendous potential for monitoring the global tropopause due to its precision, temporal consistency, and global measurement density. This study examines the capability of RO to monitor the global tropopause by addressing three specific objectives: (1) quantify sources of uncertainty in individual RO tropopause measurements, (2) examine mean bias and long-term stability of RO tropopause parameters with respect to those obtained from radiosondes, and (3) distinguish between differences due to processing and RO instrument differences by comparing tropopause parameters from different RO products. In this study, we make use of data from four different RO missions, including the recent COSMIC (Constellation Observing System for Meteorology, Ionosphere, and Climate). RO tropopause uncertainty is shown to be largely due to the use of a highly nonlinear tropopause definition (1.6 K or 510 m), although uncertainties in the RO derived temperature profiles themselves (0.25 K or 75 m) are still significant. Global mean temperature and height biases between RO instruments and radiosondes are within 0.5 K and 75 m. One long-term RO dataset examined in this study appeared to contain spurious temperature trends, but these have since been corrected. Tropopause measurements from different RO instruments are generally within 41 m and 0.1 K for the globe. Dissimilarly processed temperature data, however, can differ by as much as 2 K in the mean. These results confirm the precision of RO data, but also demonstrate the importance of consistent processing for long-term tropopause temperature studies. Tropopause height data do not appear to be significantly affected by the differences in processing examined in this study.

1. Introduction
Changes in the structure of the tropopause have recently received increased attention both as important factors in climate processes and as sensitive indicators of human induced climate change. For example, changes in the structure of the tropopause may affect stratosphere-troposphere exchange [Holton et al., 1995; Shapiro, 1980], the stratospheric Brewer-Dobson circulation [Randel et al., 2006], and stratospheric moisture content [Mote et al., 1996; Randel et al., 2004; Zhou et al., 2001] by modulating the saturation vapor pressure at the tropical cold point tropopause. In addition, increases in the latitudinal extent of the tropical tropopause suggest a widening of the tropical Hadley cell overturning circulation [Reichler and Held, 2005; Seidel et al., 2008], and the rising of the tropopause with time is considered a sign of anthropogenic climate change [Santer et al., 2003; Sausen and Santer, 2003]. The need for accurate determination of the structure and long-term behavior of the tropopause is clear.

Measuring the global tropopause, however, has been problematic. Radiosondes are often considered the ‘gold standard’ for measuring tropopause parameters due to their direct, high resolution measurements. However, early radiosondes were often unable to reach the tropopause, and radiosonde products today, while generally of good quality, are expensive and distributed unevenly across the globe. Reanalyses make up somewhat for what radiosondes lack, by providing excellent global spatial and temporal coverage. However, reanalyses are model driven and subject to sudden changes as new data are assimilated [Sturaro, 2003; Sterl, 2004]. In addition, reanalyses are sometimes of questionable quality [Trenberth and Caron, 2001; Trenberth et al., 2001; Trenberth and Stepaniak, 2002; Greatbatch and Rong, 2006; Zhao and Li, 2006], and have low vertical resolution. This low vertical resolution causes reanalyses to miss important features, such as the tropopause inversion layer [Birner et al., 2006].
Global Positioning System radio occultation (RO) is an innovative new technology for monitoring the global atmosphere. RO instruments measure the time delay in occulted signals from one satellite to another. Processing these time delays yields atmospheric bending angle profiles, information which can then be used to derive profiles of atmospheric temperature and moisture, although additional external data is needed to distinguish between the effects of the two. The near absence of moisture in the tropopause region makes temperature profiles in that region particularly accurate. In addition, RO data are not susceptible to instrument drift, and are self-calibrating [Anthes et al., 2000].

In 1995 the first Earth-observing RO mission, the Global Positioning System Meteorology (GPS/MET) experiment, was launched, and until early 1997 produced 100–150 measurements per day (see Figure 1 for a mission timeline). In 2001, the CHAllenging Minisatellite Payload (CHAMP) and the Argentinian Satelititede Aplicaciones Cientificas-C (SAC-C) became operational, and produced roughly 150 and 100 measurements per day, respectively. SAC-C profiles are available through most of 2002, while CHAMP is still operational. The Gravity Recovery and Climate Experiment (GRACE) mission was launched in 2002, and GRACE provides roughly 100–150 occultations per day. In 2006 the Constellation Observing System for Meteorology, Ionosphere and Climate (COSMIC), a constellation of six satellites was launched, and currently produces about 1500 occultations daily across the globe (although it is expected to perform even better in the future, taking as many as 2500 occultations per day).

Ware et al. [1996] and Kursinski et al. [1997] examined profiles from GPS/MET and found RO to be quite accurate (within 1 K) for measuring atmospheric temperatures between heights of 5–15 km. Hajj et al. [2004] followed up on their work, presenting a detailed characterization of the precision of RO data using CHAMP and SAC-C. They found the accuracy of RO to be within
0.5 K between 5 and 20 km. More recently, Anthes et al. [2008] estimated the temperature precision of COSMIC to be roughly 0.25 K between 10 and 20 km.

Several studies have also been performed which use RO data to study the tropopause region. Randel et al. [2003] used GPS/MET and Schmidt et al. [2004] used CHAMP data to study the mean structure and variability of the tropical tropopause. In addition, Randel et al. [2006] used RO data to document the occurrence, mean height, and mean temperature of multiple tropopauses. However, none of these studies examine in detail the sources and magnitude of RO tropopause uncertainty.

An additional issue in RO tropopause measurements is the effect of data processing. RO temperature data are not direct measurements, but derived from the time delay of a signal. Thus the choice of processing method plays a significant role in the final temperature product. Von Engeln [2006] highlighted this “structural uncertainty,” and found troubling differences in lower tropospheric and upper stratospheric temperature processed with different algorithms from the same data.

In the present study, we seek to address the suitability of RO data for the study of long-term tropopause trends by addressing three specific objectives.

- **Objective 1.** Determine the magnitude, structure, and sources of uncertainty in individual RO tropopause measurements.
- **Objective 2.** Characterize the mean bias and long-term stability of RO tropopause parameters with respect to those measured by radiosondes.
- **Objective 3.** Distinguish between the contributions of processing and instrument differences to biases between different RO products.
Our paper is organized as follows. In Section 2 we describe the data used in our study, and in Section 3 we outline our methods. In Section 4 we present our results, followed by a discussion of these results in Section 5. We summarize and offer our conclusions in section 6.

2. Data

In our study, we make use of moisture-corrected atmospheric temperature profile datasets derived from several RO instruments and provided by the UCAR COSMIC Data Analysis and Archive Center (CDAAC) and one ‘dry’ temperature profile dataset provided by the GFZ Potsdam Information Systems and Data Center (ISDC). These profiles have a 100 m or 200 m vertical resolution depending on the product, and use a one-dimensional variational method [Kuo et al., 1998] and low-resolution European Centre for Medium-Range Weather Forecasts (ECMWF) analysis to estimate the relative contributions of moisture and temperature to the measured bending angle. While the difference between the moisture-corrected and the ‘dry’ profiles is very small in the tropopause region, we focus on the moisture-corrected profiles in order to provide the closest possible comparison with radiosonde data.

The RO data we use in this study are processed in a variety of different ways, which we group into three categories: near-real-time, post-processed, and ISDC. The near-real-time processing methods change with time and are intended mainly for weather forecasting and producing atmospheric analyses. CDAAC provides near-real-time CHAMP and COSMIC data. In addition, CDAAC also frequently reprocesses their RO datasets with the intent to provide consistent, high quality, post-processed data for use in climate studies. ISDC also reprocesses their data. See Table 1 for a list of RO datasets used in this study.
GPS/MET data used in this study was post-processed in 2006 (see Table 1). Since that time, errors in CDAAC processing have been found and the CHAMP, SAC-C, and COSMIC datasets have all been reprocessed to correct those errors. The near-real-time data are not reprocessed, however, so corrections to the processing methods are incorporated midway through the datasets. Thus, GPS/MET and some near-real-time data will be subject to additional errors in processing.

In addition, the GPS/MET instrument was unable to obtain accurate retrievals when anti-spoofing (A/S) encryption was enabled [Hajj et al., 2004], which led to much noisier refractivity profiles during time periods when A/S remained on. A/S was turned off for four three-week periods, however. We used data from the entire GPS/MET dataset, and from these “prime times” alone, and found no significant differences.

Because of numerous software uploads and testing [Hajj et al., 2004], the number of successful daily occultations from SAC-C is highly variable. Data from JJA 2002 are often examined, as this represents a relatively long period of high quality and quantity data retrievals. The JJA SAC-C data in this study is computed from 2002 alone.

In order to examine the biases of the above RO datasets, we need a reference dataset. For this purpose, we use radiosonde data from the Integrated Global Radiosonde Archive [Durre et al., 2006]. These data (hereafter denoted radiosondes) are characterized by their consistency, as well as their reasonable spatial and temporal coverage.

3. Methodology

To obtain the tropopause parameters for use in this study, we apply the WMO lapse rate tropopause criterion [WMO, 1957] to the RO temperature profiles from CDAAC and ISDC, considering only the lowest tropopause in each profile. We use a tropopause detection algorithm...
based on that proposed by Reichler et al. [2003], interpolating between vertical levels to determine the pressure and temperature of the tropopause. On the other hand, radiosonde tropopauses are already reported from high vertical resolution data using the WMO criterion. While the reported lapse-rate tropopauses are at times missing and may contain errors [J. C. Antuña et al., 2006], comparing tropopauses from high resolution RO data with tropopauses calculated from the available low resolution radiosonde data may necessitate degrading the RO dataset. To take advantage of the high vertical resolution of RO data, we decided to use the reported tropopause heights and temperatures as our best estimates of the real tropopause. For both RO and radiosonde tropopause data, we consider tropopauses above 75 hPa or below 450 hPa as outliers and omit them.

The results for our study are all concerned with differences between nearby tropopause measurements from the above datasets. We thus need to define what classifies measurements as “nearby” for our purposes. Rocken et al. [1997] use a 4° latitude and longitude radius, and ±6 hr time window, while Hajj et al. [2004] compare measurements within 200 km and 30 min of one another. However, neither of these studies focuses on the tropopause, and neither had available the high resolution data from COSMIC.

In this study we use a collocation requirement of 150 km and ±6 hr. This requirement is based on the observed variability in the tropopause parameters themselves as measured by COSMIC (see section 4.2.1). Our choice represents a compromise between precision and quantity; tightening the collocation requirements naturally yields smaller collocation error, but it also reduces the number of nearby measurements and thus the robustness of our statistical analysis. At this stage, we apply one additional filter to our data; we ignore tropopause matches differing by more than 175 hPa, as these almost certainly are not measurements of the same physical tropopause (e.g., a
first tropopause in one profile may correspond closely to the second tropopause in a neighboring profile).

We next bin the differences between measurements for specific regions and seasons, count the number of differences $N$ in each bin, and calculate root mean square (RMS) difference and mean difference $\mu$ for each bin. We can derive the standard deviation $s$ for our differences using the identity

$$s^2 = RMS^2 - \mu^2,$$

and then compute a two-sided 95% confidence interval width $w$ for the differences from:

$$w = t \sqrt{\frac{s^2}{N}},$$

where $t$ is the t-value from t-statistics, assuming our differences are independent. While this assumption may be called into question, we justify it for the sake of generating useful statistics, and since the actual amount of dependence is difficult to quantify.

4. Results

4.1 A theoretical RO uncertainty estimate

To assist us in determining the sources of uncertainty in individual tropopause measurements (see Objective 1), we begin by discussing a conceptual tropopause model used by Shepherd [2002]. In this model, the actual temperature profile about a tropopause is approximated above and below by constant lapse rates (see Figure 2), and the tropopause is depicted geometrically as just the vertex of temperature profile line segments in a height-temperature plane. Although a highly simplified view of the tropopause, this is useful for describing the magnitude and cause of changes in tropopause height and temperature [Gettelman et al., 2008]. Here, we use this model
in order to estimate the amount of uncertainty in tropopause height measurements due strictly to
RO temperature uncertainties in the tropopause region.

When comparing two model tropopauses in the same plane (see Figure 2), it is evident that a
change in tropopause height \( \Delta Z \) and temperature \( \Delta T \) depends only on the changes in
temperatures \( \Delta T_s \) above and \( \Delta T_t \) below the tropopause and on the lapse rates \( \gamma_s \) above and
\( \gamma_t \) below the tropopause, where the lapse rate \( \gamma \) is defined as
\( \gamma = -\frac{\partial T}{\partial z} \). Specifically, the
temperature \( T \) can be expressed as \( T = T_s - \Delta Z \gamma_s \) or as \( T = T_t - \Delta Z \gamma_t \). Setting these two
expressions equal to each other and rearranging yields the following expressions:

\[
\Delta Z = \frac{\Delta T_t - \Delta T_s}{\gamma_t - \gamma_s}, \quad (1)
\]

and

\[
\Delta T = \frac{\Delta T_s \gamma_t - \Delta T_t \gamma_s}{\gamma_t - \gamma_s}. \quad (2)
\]

These expressions can be further simplified to analyze the sensitivities of tropopause parameters
to perturbations above and below the tropopause [Austin and Reichler, 2008]. For example, (2)
suggests that the extratropical tropopause temperature is more sensitive to temperature changes
in the stratosphere than in the troposphere, as lower stratospheric lapse rates are generally
smaller than upper tropospheric lapse rates. We can also see from (1) that a change in height is
maximized by opposing temperature changes above and below the tropopause. If we consider the
average tropopause height uncertainty \( \delta Z \) to be half the distance between the highest and lowest
possible tropopauses given our temperature uncertainty \( \delta T \), then from (1), we have
From Anthes et al. [2008], individual profiles from COSMIC are precise to about 0.25 K between 10 and 20 km. If we let $\delta T = 0.25$ K, and let $\gamma_t = 6.5$ Kkm$^{-1}$ and $\gamma_s = 0$ in the high latitudes and $\gamma_t = 4$ Kkm$^{-1}$ and $\gamma_s = -4$ Kkm$^{-1}$ in the Tropics, then from (3) we can estimate RO tropopause height uncertainty in high latitudes to be within 75 m and in the Tropics within 65 m.

This estimate utilizes a highly simplified tropopause model, and does not account for uncertainty due to the highly nonlinear, threshold definition of the tropopause (i.e. two very similar temperature profiles can have very different tropopauses) or collocation error resulting from the natural variability of the atmosphere in space and time. Our RO uncertainties measured in the following sections will include these additional errors, and by comparing these errors to those described above, we will be able to quantify the different sources of error.

4.2 RO tropopause self-comparisons

In this section, we continue to address Objective 1 by examining measurement differences from individual RO instruments. We focus mainly on data from COSMIC because of its high global measurement density. We will first examine how uncertainty in RO temperature and height vary with distance in space and time. This will allow us to estimate the actual uncertainty in individual tropopause measurements and the effect of natural atmospheric variability. We will then briefly examine the geographical structure of tropopause height and temperature uncertainty.

4.2.1 RO tropopause uncertainty by distance in space and time
To discuss the temporal and spatial structure of RO tropopause uncertainty, we use COSMIC data and focus on the Northern Hemisphere (NH) high latitudes (60°N - 90°N) and on the Tropics (24°S-24°N) during DJF (see Figure 3). We calculate RMS differences in COSMIC tropopause parameters and present the results binned by distance in space and time. We do not see any systematic biases here, because this is a self-comparison.

In DJF, the number of matching measurements (Figure 3a, d) increases with distance, but not strictly with separation in time. In the NH high latitudes during DJF, the RMS differences in height (Figure 3b) and temperature (Figure 3c) increase with separation in space and time, as one would expect. In the Tropics (Figure 3e, f) the increase in RMS differences with distance is obvious, but the trend with separation in time is not as clear. This may be related to the diurnal convective cycle. In the high latitudes, RMS contours have a fairly well defined slope, indicating that tropopause fluctuations propagate at characteristic phase speeds. In all cases, differences tend to some non-zero value as the separation in space and time approaches zero. Applying a planar fit to this data and extrapolating to zero yield RMS temperature (geopotential height) uncertainty near 1.7 K (480 m) in the NH high latitudes and 1.6 K (510 m) in the Tropics. These values are much larger than the estimates using the conceptual tropopause model described in the preceding section. RMS differences in the NH Subtropics (24°N - 60°N, not shown) are even higher (about 2.3 K or 800 m) due to the occurrence of double tropopauses and the resulting strongly discontinuous nature of the first tropopause there [Schmidt et al., 2006; Randel et al., 2007].

The planar approximation also gives us a convenient estimate of the slope of the RMS contours in Figure 3, and thus of the characteristic phase speed of tropopause disturbances. Phase speeds for NH disturbances are estimated in this way to be about 8 ms\(^{-1}\), and for the Tropics to be much
smaller (2 ms⁻¹) and less well defined. Phase speeds are largest in the NH high latitudes during DJF, consistent with a strengthening of the westerlies.

4.2.2 Geographical RO tropopause uncertainties

We now analyze the global distribution of collocated tropopause measurements and their RMS differences, using the 150 km and ±6 hr window we described in our methods. As before, we focus on COSMIC data from DJF, but we examine measurements binned in 30° latitude, longitude bins (Figure 4). The number of measurements (Figure 4a) follows a clear meridional pattern, with the most measurements in the Subtropics and the least in the Tropics. The RMS differences in tropopause height (Figure 4b) are generally on the order of 500 m, and RMS differences in tropopause temperature (Figure 4c) range from 1 K to greater than 2.5 K, with the bulk around 1.5 K. Height and temperature differences also show a meridional structure, with higher values in the Subtropics. Again, these higher RMS differences in the Subtropics are likely related to the occurrence of double tropopauses.

4.3 Meridional bias structure of RO when compared to radiosondes

In order to understand the biases of different RO datasets (see Objective 2), we now examine the meridional bias structure in RO-derived tropopause heights and temperatures by comparing the different RO products to radiosonde measurements. We examine the zonal, seasonal, 3° latitude mean temperature and height biases between RO and radiosonde tropopauses, as well as the 95% confidence intervals for those mean biases. We also examine the RMS differences between RO and radiosonde tropopause parameters by latitude and season. The seasonal data is calculated from all instances of each season during the time span of each RO product, although we use the average number of measurements for each season in our confidence interval calculation. This is
intended to estimate the actual amount of confidence one can place in the measurements from an
individual season. When displaying our zonal mean biases, RMS differences, and confidence
intervals, we apply a ±30° latitude Gaussian smoothing.

We begin our comparison by considering the tropopause height biases between the four RO
instruments from CDAAC – GPS/MET, SAC-C, CHAMP, and COSMIC – and radiosondes (see
Figure 5). No obvious pattern is evident in height biases between GPS/MET (red curve) and
SAC-C (red curve) and radiosondes, although there appears to be generally negative height bias
in SAC-C tropopause data. Global mean COSMIC tropopause height biases (blue curve) are
slightly negative (ca. -44 m) when averaged over all four seasons, and are in no seasons larger
than 200 m. While confidence intervals for GPS/MET – radiosonde biases are often as large as 2
km, confidence intervals for COSMIC are between 20 – 150 m.

We now examine temperature biases between the same four RO instruments and radiosondes
(Figure 6). Zonal mean tropopause temperature biases between GPS/MET and radiosondes
(yellow curve) tend to be cool in the high latitudes (nearly –2 K) and are often warm in the
Tropics (by as much as 4 K). SAC-C biases (red curve) vary substantially from season to season,
but demonstrate no clear pattern. In contrast, CHAMP and COSMIC (green and blue curves)
show generally small warm biases (between 0.25 K and 0.5 K in the global mean), and are never
larger than 1 K. As with height biases, confidence intervals for temperature biases are small for
COSMIC and large for GPS/MET.

To determine the factors that determine the confidence interval sizes described above, we now
investigate the meridional distribution of the collocated RO and radiosonde measurements used
above, as well as the meridional structure of RMS differences in height and temperature. The
meridional measurement distribution (Figure 7) clearly reflects the high density of radiosonde stations over land surfaces in the NH, as well as higher RO measurement density in the subtropics. The confidence intervals over the Southern Hemisphere and the small confidence intervals in the NH are due to the much smaller and larger numbers of measurements at the respective latitudes. However, in the NH Subtropics, measurement density is reasonably high, so the noticeable width of the height and temperature confidence intervals is likely due to increased variability at those latitudes. This is confirmed upon examination of the RMS differences in tropopause height and temperature of the different RO products with respect to radiosondes (Figure 8, 9). The Subtropics are regions of consistently high RMS differences between RO datasets and radiosondes. The RMS differences for the individual comparisons are fairly similar, although we do see some variability in the RMS values from season to season in GPS/MET and SAC-C data, as well as in CHAMP and COSMIC data at far latitudes (see Figure 8b, d), all related to the larger sampling uncertainty for those instruments or regions.

4.4 Meridional bias structure between RO products

Now that we have examined the absolute biases for RO instruments using radiosonde data, we compare tropopause height and temperature data between different RO instruments in order to quantify the differences between instruments as part of Objective 3. Although we consider four RO instruments in this study, only two sets of instruments have temporal overlap: CHAMP and SAC-C, and CHAMP and COSMIC. Here we do not examine seasonal biases, because of the lack of matching SAC-C and CHAMP tropopause measurements. Rather, we show the mean biases based on the entire time period in which the two products overlap: a year for CHAMP and SAC-C, and SON for CHAMP and COSMIC. We compare tropopause height and temperature biases of SAC-C and of COSMIC, using CHAMP as a reference (Figure 10). Global mean height
biases between SAC-C and CHAMP (teal curve) are ~41 m, while those between COSMIC and CHAMP (purple curve) are ~8 m. Similarly, although temperature biases (Figure 10b) between CHAMP and SAC-C are not large, they are larger than those between CHAMP and COSMIC.

To examine the effect of processing on measured tropopause parameters, we examine the height and temperature biases between near-real-time COSMIC and post-processed COSMIC data (Figure 11), with the post-processed data as our reference. We do not show confidence intervals for these biases, but note that their widths are all within ±0.2 K due to the high density of COSMIC measurements. Biases in tropopause height in the near-real-time data (dashed lines) are high by roughly 20 m, except in the high latitudes during SON. Biases in near-real-time temperature (solid lines) for MAM (SON) are about 0.1 K (1 K) for the Tropics and –0.4 K (–2 K) for high latitudes. A similar latitudinal pattern in temperature biases (warm in the Tropics, cool at high latitudes) is evident in the temperature comparison between GPS/MET and radiosondes (Figure 6), suggesting that some common flaws – perhaps in the processing – could be responsible for the larger biases. The smaller biases in DJF and SON are likely due to a correction of the near-real-time processing scheme.

4.5 Long-term CHAMP temperature stability

Now that we have shown the effect of processing on RO measurements, we are ready to address the instrument stability concern in Objective 2. Here, we examine the mean tropopause temperature bias between CHAMP – our longest-running RO dataset – and radiosondes by latitude and season (Figure 12). We examine data from CHAMP data downloaded in June 2007, CHAMP data processed at ISDC, and CHAMP data processed in December 2007.
Temperature biases between the early CDAAC CHAMP dataset and radiosondes (Figure 12a) show strong shifts. In particular, we note the abrupt shifts in JJA 2005 and MAM 2006, where biases change by as much as 3 K. Height biases (not shown) show no such sudden change; i.e. they appear to be much more stable in time. The bias pattern in the ISDC CHAMP data (Figure 12b) is qualitatively and quantitatively similar to the temperature bias pattern described above in the GPS/MET data and near-real-time data. Although the data from ISDC are somewhat more variable in time than the post-processed data, they do not exhibit unusual shifts in time. Biases between the newly processed CHAMP data from CDAAC and radiosondes (Figure 12c) are quite a bit smaller, and show no such strong shifts.

4.6 Quality of radiosonde data

Digressing from our main objectives for a moment, we point out one additional result from our study. In addition to their use for examining RO tropopause data, our methods are also useful for examining the regional differences in tropopause parameters as measured by radiosondes. To do this, we use COSMIC as our reference and examine the regional pattern of mean biases between radiosondes and COSMIC. We show here only results for DJF 2006 (Figure 13) since the other seasons are qualitatively very similar. While height biases for the individual regions are generally of the same order as those for the zonal mean biases (Figure 5), India’s radiosonde network exhibits exceptionally strong and consistently negative biases; several between −1 km and −2 km. Temperature biases (not shown) are consistently too warm over much of India (about +4 to +7 K) as well. Our results substantiate those mentioned in Kuo et al. [2005] and support using RO data to characterize differences between radiosonde products.

5. Discussion
In this section, we discuss the results shown in our previous section, and how they address each of our objectives.

5.1 Magnitude, structure, and sources of uncertainty in RO tropopause measurements

In section 4.1, we use a conceptual tropopause model to estimate the uncertainty in RO tropopause height precision to be about 75 m, based on an RO temperature precision of 0.25 K [Anthes et al., 2008]. In section 4.2.1, we examine matching COSMIC data, and determine that the actual uncertainty in determining tropopause temperature (height) for the same location and time are closer to 1.7 K (480 m) in the NH high latitudes and 1.6 K (510 m) in the Tropics. Previous studies found tropopause parameter differences between RO and radiosondes to be in the neighborhood of 2–4 K (700 m) [Randel et al., 2003; Schmidt et al., 2004], and tropopause temperature uncertainty in radiosonde data to be roughly the same, at ~2–3 K [Tsuda et al., 1994].

We explain this large discrepancy between the theoretically expected and actual differences with three sources of uncertainty: (1) Instrument uncertainty, (2) tropopause definition uncertainty, and (3) collocation error. By instrument uncertainty, we mean the uncertainty in RO temperature measurements in the vicinity of the tropopause, and the resulting uncertainty in height based on the conceptual tropopause model in Section 3. By tropopause definition uncertainty, we mean the additional uncertainty brought on by the nonlinear definition of the tropopause that cannot be accounted for by instrument uncertainty alone. Collocation error simply results from comparing measurements that differ in time and space. See Table 2 for a summary of the different sources of uncertainty and their respective magnitudes.
The 0.25 K figure from Anthes et al. [2008] and the corresponding tropopause height uncertainty of 75 m from Section 3.1 are achieved by filtering out natural variability and by applying our conceptual tropopause model (see Section 3.1). Anthes et al. [2008] filtered out atmospheric variability by extrapolating measured differences to a collocation distance of zero. On the other hand, the 1.6 K figure obtained in this study is an estimate of the actual uncertainty in RO tropopause measurements for a collocation distance of zero, and thus includes tropopause definition uncertainty as well. If we assume that the planar extrapolation used in Section 3.2 effectively filters out collocation error, and that the sources of uncertainty (as estimated by RMS differences) add linearly in the square, then we can estimate the amount of uncertainty in tropopause measurements due to the tropopause uncertainty itself using

\[
\left( \frac{RO}{\text{uncertainty}} \right)^2 + \left( \frac{\text{tropopause}}{\text{uncertainty}} \right)^2 + \left( \frac{\text{collocation error}}{\text{uncertainty}} \right)^2 = \left( \frac{\text{total uncertainty}}{\text{uncertainty}} \right)^2. \tag{4}
\]

From (4), setting the collocation error to zero, and recalling a combined RO and tropopause uncertainty of roughly 1.6 K (510 m), we estimate that the uncertainty due to the nonlinear definition dominates the total uncertainty with a contribution of nearly 1.6 K (510 m). This estimate is itself, however, a function of the RO uncertainty; tropopause definition uncertainty is a function of instrument precision.

The large (~3 K) RO tropopause uncertainties reported by Randel et al. [2003] and Schmidt et al. (2004) also include collocation error, which will vary depending on the collocation requirement used. Using our collocation criteria (within 150 km and ±6 hrs), we have an RMS RO tropopause temperature uncertainty of roughly 2 K (700 m; see Figure 4). Recalling the different sources of uncertainties discussed above, from (4) we can estimate the portion of this uncertainty due to collocation error to be 1.2 K (480 m). We caution that these are rough estimates and that actual
uncertainties vary strongly with latitude (e.g., tropopause uncertainty and collocation error are both higher in the Subtropics). It should be clear, however, that estimates of instrument uncertainty obtained by comparing collocated tropopause measurements will overestimate instrument uncertainty if the collocation error is not removed and, in this case, if tropopause uncertainty is not accounted for.

We now address how future tropopause climate studies may reduce the sources of uncertainty described above. On one hand, collocation error is not an instrument error at all, but an error that occurs when comparing collocated measurements. Thus, collocation error only needs to be considered when estimating instrument uncertainties, not when analyzing long term trends. Instrument temperature uncertainties, however, are unavoidable as long as a one uses a temperature-based tropopause definition, or is interested in RO-derived tropopause temperatures. Rao et al. [2007] construct a bending angle tropopause definition, however, which may be useful for studying long-term tropopause height trends, as the RO bending angle is not subject to the same degree of uncertainty as the RO derived temperature.

Uncertainty stemming from the tropopause definition may be reduced by adding additional filters to the data, i.e., examining strictly the tropical or an extratropical tropopause, or by defining a tropopause “strength” threshold, to avoid tropopauses that are likely to be missed in nearby observations. Calculating a “climatological tropopause” from time-mean temperature profiles would also likely reduce, if not eliminate, much of the uncertainty due to the nonlinear nature tropopause; time-averaging eliminates the fine structure responsible for the tropopause uncertainty we described previously.

5.2 Biases and long-term stability of RO tropopause parameters
In order to put the following discussion in perspective, we first mention that the temperature stability requirement for climate studies in the troposphere is 0.04 K according to Ohring et al. [2005]. In other words, if the bias in a dataset changes by more than 0.04 K per decade, it may prevent the detection of a climate trend with any reasonable degree of confidence. However, this stability requirement is related to measurement of surface temperature changes associated with global climate change, and we relax the requirement here, as tropopause temperature trends are stronger; about −0.5 K per decade at the tropopause [Seidel and Randel, 2006] as opposed to +0.13 K per decade at the surface [Solomon et al., 2007]. Accordingly, we consider any change in tropopause temperature biases larger than 0.2 K to be seriously detrimental for use in climate studies. Using our conceptual tropopause model, we estimate a 0.2 K bias shift to be analogous to a height shift of about 60 m.

Keeping this in mind, we now proceed to discuss the biases of RO tropopause parameters based on our comparison with radiosonde data. Global mean height biases between COSMIC and radiosondes are roughly -44 m. Temperature biases of GPS/MET against radiosondes are as large 4K depending on altitude (1.06 K for the globe) due to processing, while those for SAC-C, CHAMP and COSMIC are smaller; 0.25 K for SAC-C and COSMIC, 0.5 K for CHAMP. Similar biases (−100 m and 0.6 K) were noted for RO derived cold point tropopause parameters in Randel et al. [2003]. Temperature biases between RO and radiosonde tropopause parameters are larger than the stability requirements we specified above, and which would suggest that one should not simply combine RO and radiosonde tropopause temperature into one long-term tropopause dataset. There may be some utility in combining the tropopause height data from the two sets, however.
While mean biases in post-processed datasets appear small, we show in Section 4.5 that tropopause temperature biases between CDAAC CHAMP data post-processed prior to November 2007 and radiosondes show sudden shifts as large as ±2.5 K. The intent of post-processed RO datasets is to provide as stable and accurate a dataset as practically possible for climate studies. Thus, the presence of such artifacts in even one post-processed dataset, and the resulting failure of the dataset to meet our temperature stability requirements for tropopause climate research, are disconcerting. It is clear that even post-processed RO datasets should undergo careful quality control and reality checks. Fortunately, height biases in all datasets (not shown) appear stable.

5.3 Sources of uncertainty between RO products

Finally, we discuss the differences in tropopause parameters between different RO instruments or processing methods. The two instrument comparisons we have available illustrate the similarity of data from different instruments. Global average biases between RO products are about 44 m (Figure 10a), while global temperature biases range from 0.3 – 0.15 K (Figure 10b), as in Hajj et al. [2004]. On the other hand, biases in temperature between dissimilarly processed datasets can be as large as 2.5 K. Thus, height biases between RO tropopause datasets, whether processed similarly or not, are smaller than the stability thresholds we specified above. Temperature biases for similarly processed datasets are also within the required limits, while temperature biases for dissimilarly processed datasets can be well outside the required range.

6. Summary and conclusion

In our introduction we state the three primary objectives of our study: (1) determine the magnitude, structure, and sources of uncertainty in RO tropopause measurements, (2)
characterize the mean bias and long-term stability of RO tropopause parameters with respect to those measured by radiosondes, and (3) distinguish between the contributions of processing and instrument differences to biases between different RO products. We now summarize our results for each objective, suggestions for future tropopause climate studies, and a few possible caveats in this study.

**Objective 1:** We find that differences in tropopause parameters between observations are due largely to collocation error. Removing this error, the remaining uncertainty in measurements is mostly due to the nonlinearity of the tropopause definition.

**Objective 2:** We examine the time-mean meridional error structure in RO height and temperature against radiosonde, and find generally good agreement in tropopause, aside from noise, and a negative global mean height bias (-44 m) in tropopause height in RO data. We see that RO processing methods can strongly affect temperature biases between RO and radiosonde data, with earlier-processed data having large warm (cool) bias in the Tropics (Extratropics). Newly processed data exhibit warm biases of 0.25–0.3 K against radiosondes. These biases agree well with those determined by Randel et al. [2003].

Upon examining the temporal structure of zonal mean tropopause parameter biases between post-processed CHAMP data and radiosondes, we find no noticeable instability in height biases, but very strong instabilities in temperature biases (2.5 K) prior to the November 2007 reprocessing. ISDC CHAMP tropopause temperatures and CDAAC CHAMP temperatures processed after December 2007, when compared to those from radiosondes, show no such instabilities. We conclude that one should be cautious in examining trends from any RO dataset, as errors due to processing can introduce spurious trends.
Objective 3: We find that tropopause heights in different RO instruments are sufficiently similar (within ~41 m) for the straightforward combination of RO datasets for use in tropopause height studies without regard to the processing methods used. RO tropopause temperature datasets, on the other hand, are only similar enough to be combined for long-term studies if the products are processed similarly.

Future climate studies may reduce uncertainty in measured tropopause parameters by filtering data as described in Section 5.1, and by considering tropopause heights computed from time-mean temperature profiles, rather than time averages of tropopause heights. It should be remembered, however, that the tropopause temperature of the time-mean temperature profile is necessarily warmer than the time mean of tropopause temperatures, as tropopauses do not occur at the same level each time. It appears that RO tropopause height datasets can safely be combined regardless of their respective levels of processing. To study RO tropopause temperature trends, however, one must combine similarly processed datasets, and check for spurious trends in each dataset. A more in-depth study of the temperature biases of different non-post-processed datasets would also be necessary to produce a tropopause temperature record going back past 2001, when CHAMP became operational.

We note the following possible caveats in our study. We make extensive use of reported tropopause data from IGRA as a reference dataset. Naturally, IGRA data does not necessarily represent truth; IGRA has documented inconsistencies, such as instrument differences between countries. In addition, the methodology for determining the lapse rate tropopause may differ between our method and those used at the radiosonde stations themselves. However, we use the reported IGRA tropopauses as a best available reference dataset. Finally, using our collocation criteria and by only comparing data from profiles which have tropopauses reduces the amount of
data substantially, which increases the sampling uncertainty in our study, particularly in smaller datasets such as GPS/MET and SAC-C.

Acknowledgments

The author would like to thank NCAR, GFZ Potsdam, and NOAA for use of their RO data, and Doug Hunt and the CDAAC team as well as Jonathan Jiang, Chi O. Ao at JPL for their discussion of RO data issues. We also thank Chuntao Liu for his assistance, and Dian Seidel, John Horel, and two anonymous reviewers for their helpful comments. This work was supported by NASA Headquarters under NASA Earth and Space Science Fellowship Program – Grant 07-Earth07F-0108 and by NSF Grant ATM 0532280.

References


Greatbatch, R. J., and P.P. Rong (2006), Discrepancies between different Northern Hemisphere summer atmospheric data products, *J. Climate*, **19**(7), 1261


Table 1. RO data products used in this study, along the general processing category (see text), the processing date or version number, starting and ending dates and the number of seasons covered. RT denotes near-real-time data, PP denotes post-processed data, and ISDC denotes data processed at the GFZ Potsdam Information Systems and Data Center.

Table 2. Sources of error in RO tropopause measurements, along with their approximate magnitudes. Also listed are the studies from which these estimates come.

Figure 1. Timeline of different products considered in this study. Arrows indicate that the dataset continues beyond the time shown.

Figure 2. Conceptual tropopause model after Shephard [2002]. The blue lines represent a temperature profile in the vicinity of the tropopause, and the red solid lines represent the profile after hypothetical temperature changes of $\Delta T_s$ above and $\Delta T_l$ below the tropopause. The black lines represent the resulting changes $\Delta Z$ and $\Delta T$ in tropopause height and pressure, respectively. The gray lines relate $\Delta T$ to $\Delta T_s$, $\Delta T_l$, $\gamma_s$, and $\gamma_l$ as described in the text. The dashed lines are for clarity.

Figure 3. Error statistics for collocated DJF COSMIC tropopause data binned by distance in space and separation in time. The number of measurements are shown in (a) and (d), RMS height errors are shown in (b) and (e), and RMS temperature errors in (c) and (f). (a) through (c) are for Northern Hemisphere high latitudes, while (d) through (f) are for the Tropics.
**Figure 4.** Geographical distribution of measurements and RMS errors for collocated DJF COSMIC tropopause data. Shown are the number of measurements in each bin (a), RMS height errors (b), and RMS temperature errors (c).

**Figure 5.** Meridional profile of mean geopotential height differences between CDAAC RO and radiosonde datasets, and the 95% confidence interval for the mean. Profiles are shown for DJF (a), (b), MAM (c), (d), JJA (e), (f), and SON (g), (h). All values are Gaussian smoothed over ±30° latitude. Gaps represent insufficient data. Average global instrument biases are -230 m for GPS/MET, -33 m for SAC-C, -42 m for CHAMP, and -44 m for COSMIC.

**Figure 6.** Meridional profile of mean temperature differences between GPS/MET and SAC-C and radiosonde datasets, and the 95% confidence interval for the mean. Profiles are shown for DJF (a), MAM (b), JJA (c), and SON (d). Gaps represent insufficient data. All values are Gaussian smoothed over ±30° latitude. Average global instrument biases are 1.06 K for GPS/MET, 0.25 K for SAC-C, 0.3 K for CHAMP, and 0.29 K for COSMIC.

**Figure 7.** Meridional distribution of the average number of matching CDAAC RO vs. radiosonde measurements per season for each RO instrument during the course of our study.

**Figure 8.** Meridional profile of RMS height differences between CDAAC RO and radiosonde data. Profiles are shown for DJF (a), MAM (b), JJA (c), and SON (d). All values are Gaussian smoothed over ±30° latitude. Gaps represent insufficient data.

**Figure 9.** Meridional profile of RMS temperature differences between CDAAC RO and radiosonde data. Profiles are shown for DJF (a), MAM (b), JJA (c), and SON (d). All values are Gaussian smoothed over ±30° latitude. Gaps represent insufficient data.
Figure 10. Meridional profile of mean geopotential height differences (a) and temperature differences (b) between the CDAAC RO datasets specified for SON. All values are Gaussian smoothed over ±30° latitude. Gaps represent insufficient data. Average global biases are -0.03 K / 41 m for SAC-C – CHAMP and 0.15 K / 8 m for COSMIC – CHAMP.

Figure 11. Zonal mean temperature differences between post-processed (PP) and near-real-time (RT) COSMIC data for the seasons shown. Dashed lines and left axes show mean height differences, while solid lines and right axes show mean temperature differences. All values are Gaussian smoothed over ±30° latitude.

Figure 12. Zonal mean tropopause temperature bias between CHAMP and radiosonde data as a function of time and for two different processing methods. Shown are CHAMP biases from an earlier CDAAC reprocessing (a), ISDC (b), and the latest CDAAC reprocessing (c).

Figure 13. Geographical distribution of mean tropopause height biases between radiosonde and COSMIC data during the period Dec. 2006 – Feb. 2007. Gray cells represent insufficient data.
Tables

<table>
<thead>
<tr>
<th>Product</th>
<th>Processing</th>
<th>Date Processed / Software Version</th>
<th>Start</th>
<th>End</th>
<th>Seasons</th>
</tr>
</thead>
</table>

Table 1. RO data products used in this study, along the general processing category (see text), the processing date or version number, starting and ending dates and the number of seasons covered. RT denotes near-real-time data, PP denotes post-processed data, and ISDC denotes data processed at the GFZ Potsdam Information Systems and Data Center.
<table>
<thead>
<tr>
<th>Uncertainty and Error</th>
<th>ΔT (K)</th>
<th>ΔZ (m)</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>RO instrument uncertainty</td>
<td>0.25</td>
<td>75</td>
<td>Anthes et al. (2008), and conceptual model (section 3.1)</td>
</tr>
<tr>
<td>Tropopause definition uncertainty</td>
<td>1.6</td>
<td>510</td>
<td>This study, equation 4</td>
</tr>
<tr>
<td>Collocation error</td>
<td>1.2</td>
<td>480</td>
<td>This study, equation 4. For collocation criteria of 150 km and ±6 hrs</td>
</tr>
<tr>
<td>Total uncertainty and error</td>
<td>2</td>
<td>700</td>
<td>This study</td>
</tr>
</tbody>
</table>

Table 2. Sources of error in RO tropopause measurements, along with their approximate magnitudes. Also listed are the studies from which these estimates come.
Figures

**Figure 1.** Timeline of different products considered in this study. Arrows indicate that the dataset continues beyond the time shown.
Figure 2. Conceptual tropopause model after Shephard [2002]. The blue lines represent a temperature profile in the vicinity of the tropopause, and the red solid lines represent the profile after hypothetical temperature changes of $\Delta T_s$ above and $\Delta T_t$ below the tropopause. The black lines represent the resulting changes $\Delta Z$ and $\Delta T$ in tropopause height and pressure, respectively. The gray lines relate $\Delta T$ to $\Delta T_s$, $\Delta T_t$, $\gamma_s$, and $\gamma_t$ as described in the text. The dashed lines are for clarity.
Figure 3. Error statistics for collocated DJF COSMIC tropopause data binned by distance in space and separation in time. The number of measurements are shown in (a) and (d), RMS height errors are shown in (b) and (e), and RMS temperature errors in (c) and (f). (a) through (c) are for Northern Hemisphere high latitudes, while (d) through (f) are for the Tropics.
Figure 4. Geographical distribution of measurements and RMS errors for collocated DJF COSMIC tropopause data. Shown are the number of measurements in each bin (a), RMS height errors (b), and RMS temperature errors (c).
Figure 5. Meridional profile of mean geopotential height differences between CDAAC RO and radiosonde datasets, and the 95% confidence interval for the mean. Profiles are shown for DJF (a), (b), MAM (c), (d), JJA (e), (f), and SON (g), (h). All values are Gaussian smoothed over ±30° latitude. Gaps represent insufficient data. Average global instrument biases are -230 m for GPS/MET, -33 m for SAC-C, -42 m for CHAMP, and -44 m for COSMIC.
Figure 6. Meridional profile of mean temperature differences between GPS/MET and SAC-C and radiosonde datasets, and the 95% confidence interval for the mean. Profiles are shown for DJF (a), MAM (b), JJA (c), and SON (d). Gaps represent insufficient data. All values are Gaussian smoothed over ±30° latitude. Average global instrument biases are 1.06 K for GPS/MET, 0.25 K for SAC-C, 0.3 K for CHAMP, and 0.29 K for COSMIC.
Figure 7. Meridional distribution of the average number of matching CDAAC RO vs. radiosonde measurements per season for each RO instrument during the course of our study.
Figure 8. Meridional profile of RMS height differences between CDAAC RO and radiosonde data. Profiles are shown for DJF (a), MAM (b), JJA (c), and SON (d). All values are Gaussian smoothed over ±30° latitude. Gaps represent insufficient data.
Figure 9. Meridional profile of RMS temperature differences between CDAAC RO and radiosonde data. Profiles are shown for DJF (a), MAM (b), JJA (c), and SON (d). All values are Gaussian smoothed over ±30° latitude. Gaps represent insufficient data.
Figure 10. Meridional profile of mean geopotential height differences (a) and temperature differences (b) between the CDAAC RO datasets specified for SON. All values are Gaussian smoothed over ±30° latitude. Gaps represent insufficient data. Average global biases are -0.03 K / 41 m for SAC-C – CHAMP and 0.15 K / 8 m for COSMIC – CHAMP.
Figure 11. Zonal mean temperature differences between post-processed (PP) and near-real-time (RT) COSMIC data for the seasons shown. Dashed lines and left axes show mean height differences, while solid lines and right axes show mean temperature differences. All values are Gaussian smoothed over ±30° latitude.
Figure 12. Zonal mean tropopause temperature bias between CHAMP and radiosonde data as a function of time and for two different processing methods. Shown are CHAMP biases from an earlier CDAAC reprocessing (a), ISDC (b), and the latest CDAAC reprocessing (c).
Figure 13. Geographical distribution of mean tropopause height biases between radiosonde and COSMIC data during the period Dec. 2006 – Feb. 2007. Gray cells represent insufficient data.