

# Quality Control, Error Analysis, and Impact Assessment of FORMOSAT-3/COSMIC in Numerical Weather Prediction

Paul Poli<sup>1,\*</sup>, Patrick Moll<sup>1</sup>, Dominique Puech<sup>1</sup>, Florence Rabier<sup>1</sup>, and Sean B. Healy<sup>2</sup>

<sup>1</sup> *Météo-France, National Centre for Meteorological Research (CNRM-GAME-CNRS), 42 avenue Coriolis, 31057 Toulouse Cedex 1, France*

<sup>2</sup> *European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading RG2 9AX, United Kingdom*

Received 27 June 2007, accepted 21 January 2008

---

## ABSTRACT

Following several years of experimentation with the GPS radio occultation technique, the 6-satellite FORMOSAT-3/COSMIC (F3C) mission was launched mid-2006 and has been collecting data since then. In this paper we present early findings of research performed at Météo-France regarding the use of these data for assimilation in numerical weather prediction. Benefiting from the dense global coverage allowed by F3C refraction-induced observations, we first assess the quality of these data at four levels: bending angle, refractivity, refractivity lapse rate, and temperature. We compare them with calculations from Météo-France numerical weather forecasts. Learning from these various levels of data we devise quality control procedures that rely on the refractivity lapse rate. Applying a recent methodology developed in data assimilation we calculate observation bending angle error variances for our assimilation system. Using these new quality control procedures and observation error estimates we run an assimilation and forecast experiment with Météo-France's operational global 4DVAR data assimilation system used as a reference. Our results indicate a very clear positive impact of the assimilation of F3C bending angle data in the Southern hemisphere for the prediction of geopotential heights and winds. We also observe an improvement in wind forecast skill in the Northern hemisphere, albeit such an improvement is smaller than in the Southern hemisphere.

Key words: Data assimilation, GPS radio occultation, Observation errors, Refractivity lapse rate, Quality control, Monitoring

Citation: Poli, P., P. Moll, D. Puech, F. Rabier, and S. B. Healy, 2009: Quality control, error analysis, and impact assessment of FORMOSAT-3/COSMIC in numerical weather prediction. *Terr. Atmos. Ocean. Sci.*, 20, 101-113, doi: 10.3319/TAO.2008.01.21.02(F3C)

---

## 1. INTRODUCTION

In 1995, The GPS radio occultation technique demonstrated its ability to collect low cost observations of the Earth's atmosphere with a passive instrument involving neither optics nor moving parts (Ware et al. 1996). The technique requires placing an advanced radio equipment (i.e., a GPS receiver) onboard a satellite whose orbit determination needs to be precise. With this setup the tracking data of the highly stable signals transmitted by the constellation of GPS satellites can be inverted to yield ionospheric and atmospheric information.

The outline of the paper is as follows. In section 2 we present the FORMOSAT-3/COSMIC (F3C) observations. We discuss the characteristics of one week of F3C data for

temperature, refractivity, refractivity lapse rate, and bending angle in section 3. Section 4 introduces Météo-France's operational assimilation and forecast system ARPEGE (the French acronym for Action de Recherche Petite Echelle et Grande Echelle, i.e., Research Project on Small and Large Scales). That section also details the observation operators used to simulate F3C data. Section 5 presents a comparison of F3C data with ARPEGE forecasts. We devise quality control procedures in section 6 and evaluate observation error covariances for bending angles in section 7. Section 8 presents the results of an assimilation and forecast experiment. Section 9 contains conclusions and perspectives for future work.

---

\* Corresponding author  
E-mail: paul.poli@ecmwf.int

## 2. THE FORMOSAT-3/COSMIC (F3C) OBSERVATIONS

### 2.1 The F3C Mission

The F3C mission comes after four successful experiments with the GPS radio occultation. The first experiment (GPS/MET) carried a so-called first generation GPS receiver instrument and collected the first-ever observations of the refraction properties of the atmosphere on GPS signals from space (Ware et al. 1996). Second-generation instruments have been experimented with onboard the CHAMP, SAC-C, and GRACE experiments (Hajj et al. 2002a). Advances in data processing have further enabled a fair level of improvement in data quality (Beyerle et al. 2004). The F3C constellation carries third-generation receivers.

The sensing principle of the GPS radio occultation technique was presented in detail by e.g., Kursinski et al. (1997) and Hajj et al. (2002b). Briefly, the GPS signals radiating from the GPS constellation of transmitters (more than 24 satellites in mid-Earth orbit) undergo bending and slowing down when propagating inside the Earth's atmosphere, by comparison with the propagation that occurs in a vacuum (i.e., outside the atmosphere). A GPS receiver on a low-Earth orbiting satellite can measure the Doppler effect induced by its relative motion with respect to the transmitter, assuming the receiver is able to track the phase of the received GPS signals. If the positions and velocities of the two platforms are known accurately, and assuming straight-line propagation between the GPS transmitter and receiver, it is possible to calculate (or predict) the theoretical Doppler effect that one should observe at the receiver. However, if an atmosphere is present along the propagation path, the angle of incidence of the tracked signal differs from what is assumed by straight-line propagation. The measured Doppler-shifted frequency is then different from the prediction. From that difference one can yield a bending angle observation. A series of such observations can be collected whenever a GPS satellite rises or sets behind the Earth's atmospheric limb (so-called occultation event).

Due to the presence of the ionosphere above the Earth's neutral atmosphere (0 - 100 km altitude), the bending angles observed by tracking of the two GPS frequencies are different. A linear combination of the two (and possibly the use of constraints via a model) can be used to remove the ionospheric contribution and retain only the neutral atmosphere contribution to bending angle (e.g., Kursinski et al. 1997). Hereafter we simply call that neutral contribution a bending angle. Assuming spherical symmetry, the resulting time-series of bending angle observations can be assigned a series of impact parameters (or asymptotic ray-miss distances) and further converted, via an Abel transform, into a vertical profile of refractivities as a function of height (e.g., Hajj et al. 2002b). These can in turn be inverted into temperature and water vapour profiles. The inversion process

requires one to assume hydrostatic equilibrium and a variational approach or equivalent scheme in order to provide constraints and make use of a priori estimates of the vertical distributions of these two meteorological parameters.

### 2.2 F3C Data Products Used in This Study

The F3C observations are processed in near-real time by the COSMIC Data Analysis and Archive Center (CDAAC) at the University Corporation for Atmospheric Research (UCAR). We use in the present study observations released in October and November 2006, and available via the internet (<http://cosmic-io.cosmic.ucar.edu/cdaac/index.html>). The geographical coverage achieved by the six F3C satellites around that time period is illustrated in Fig. 1. Note that the coverage has been increasing since, with software and firmware upgrades and with the growing distance between the six micro-satellites, allowing for more data downloads. We use the data rendered in the World Meteorological Organization (WMO)-approved Binary Universal Form for the Representation of meteorological data (BUFR). We extract for each occultation event the profile of bending angles versus impact parameters, refractivities versus altitudes, and temperatures versus geopotential heights. Note that bending angles and refractivities share the same vertical grid while temperatures are given at different levels. Each profile contains 200 levels and extends up to an altitude of 40 km except for temperatures which only extend up to around 30 km altitude. We also extract for each occultation event geometrical parameters such as the latitude and longitude profiles available on the same levels as the impact parameters, the geoid undulation ( $\Delta$ : difference between the geoid height and the ellipsoid height), and the local radius of curvature of the Earth with respect to the ellipsoid, noted  $R_c$ . The starting date and time of each occultation is also used as time tag. In order to evaluate the physical conditions under which bending occurred, we also compute for each individual refractivity observation the refractivity lapse rate ( $dN/dz$ ) using the next adjacent refractivity observation located above.

## 3. ONE WEEK OF F3C DATA

We consider here 7 days of F3C data obtained between 0300 UTC 4 October 2006 and 0900 UTC 11 October 2006 (total of 8432 occultation events). We bin the observations in 10 degree latitude and 1 km thick vertical bins. The vertical coordinate used in this operation for bending angle, refractivity, and refractivity lapse rate is the altitude found in the observations. We use geopotential heights for the temperatures.

We compute the mean for temperatures, refractivities, refractivity lapse rates, and bending angles within each bin (no outlier removal). The result is shown in Figs. 2a - d. The zonal temperature structure is consistent with climatology:

e.g., tropopause around 15 km altitude in the tropics, double structure around 60°N latitude. Note that the mean refractivity appears fairly constant with respect to latitude as a result of the log-scale. The refractivity lapse rate closely follows the atmospheric density lapse rate in the stratosphere and in the troposphere with temperatures below 250 K (typically down to 8 km altitude in the tropics, and near-surface at the poles). But more importantly, the vertical distribution of the refractivity lapse controls primarily the refraction-induced measurement process. The refractivity lapse rate is usually negative, and increases with altitude to less negative values, except around the tropical tropopause where that decay is slower.

We also evaluate within each bin the variability of each retrieved parameter by computing the standard deviation in Figs. 2e - h. The result encompasses the natural variability of the observed quantities as well as the measurement and processing errors. In terms of temperature, the mid-latitudes are areas of larger variability than the Tropics. The Southern polar vortex (around 20 km in altitude) exhibits the largest temperature variations (up to more than 10 K). In terms of refractivity, refractivity lapse rate, and bending angle, an area of relatively large variability can be identified in the same region but starting at an altitude of 20 km and extending higher up. The moist troposphere is also a region where natural variability and possibly measurement and processing error lead to larger standard deviations in observed refractivity lapse rate. Overall,

the various levels of F3C data seem consistent between one another. One exception is the temperature retrieval in the Southern hemisphere near the 17 km altitude level where the variability seems larger than in the other products. In fact, plotting the temperature variability in percents yields a result that is more consistent with refractivity standard deviations: the observed temperature standard deviations are less than 2% throughout most of the domain except for the Southern high latitudes. Furthermore, this area corresponds to the natural variability associated with the break-up of the polar vortex in the Southern hemisphere spring.

## 4. MÉTÉO-FRANCE DATA ASSIMILATION SYSTEM

### 4.1 Reference System

We used the ARPEGE assimilation and forecast system which became operational at the end of 2006 (version CY30T2\_op2) as a baseline for the present work (Janiskova et al. 1999; Gauthier and Thépaut 2001). The forecast model and the assimilation scheme share the same 46 vertical levels between the surface and the 0.1 hPa pressure level (or about 65 km altitude). The horizontal resolution of the forecasts is T358 on a stretched grid (about 23 km over France and 133 km over France's antipodes). The assimilation horizontal resolution is T149 on a uniform grid (or about 133 km). The 4DVAR assimilation scheme bins all observa-

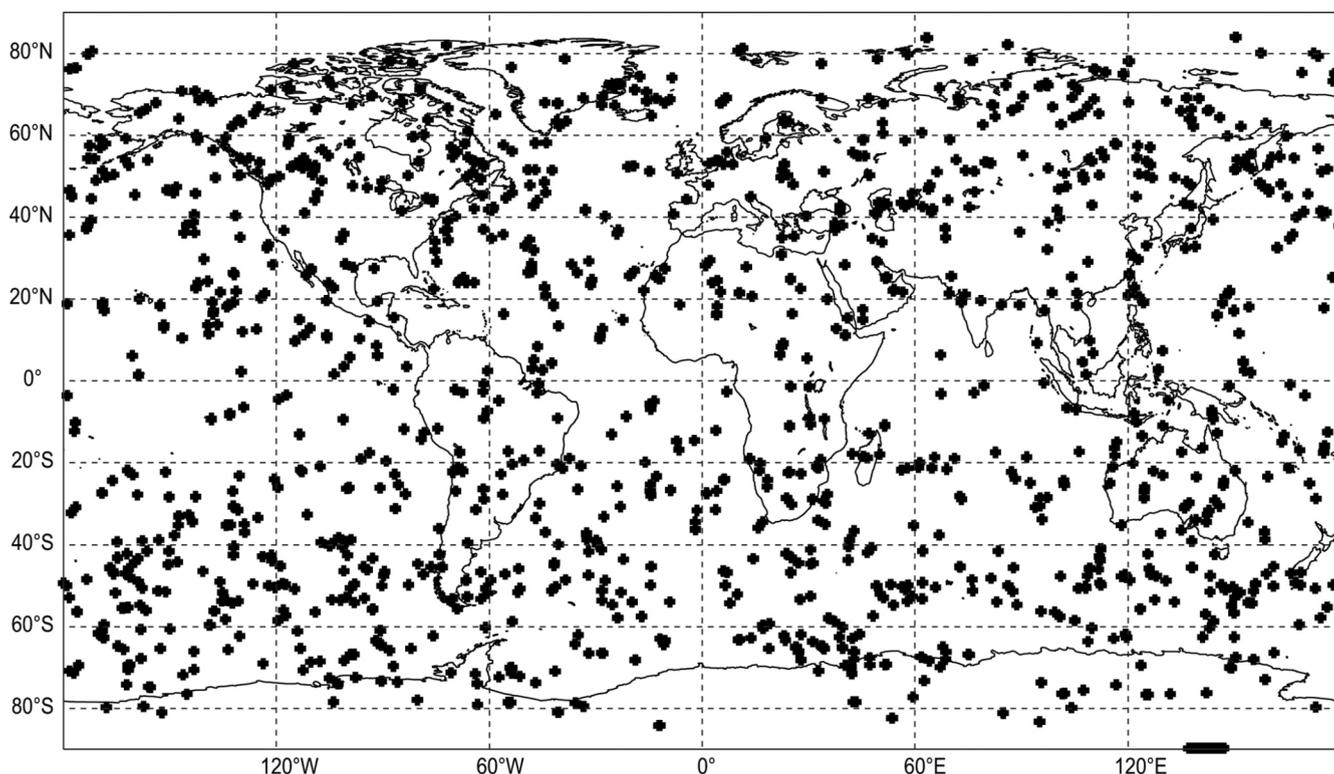


Fig. 1. A one-day coverage illustration of the GPS radio occultation events collected by F3C on 6 October 2006 (total: 1323 events).

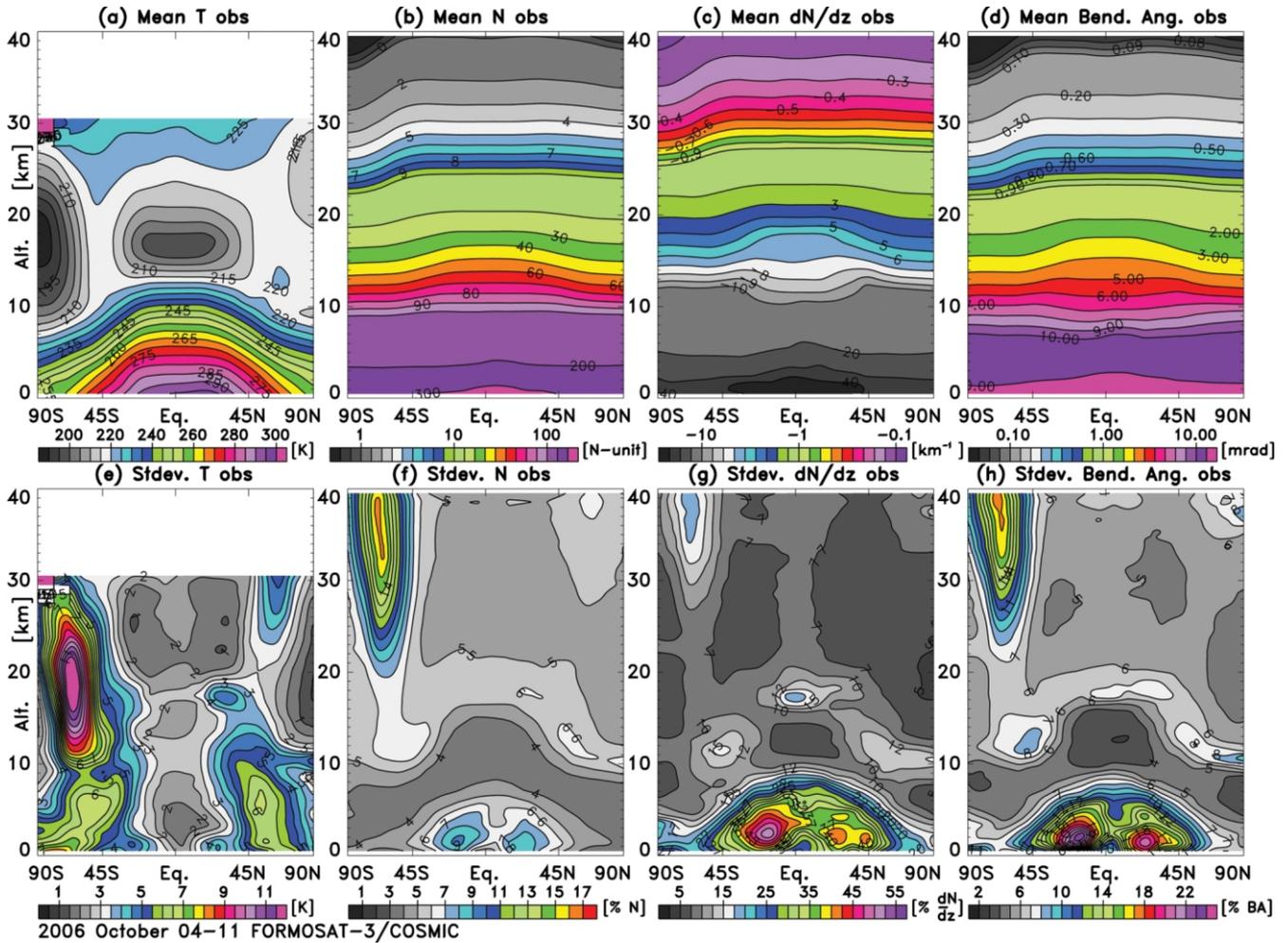


Fig. 2. The zonal average (1 km vertical resolution, 10° latitude) of the F3C observations for (a) temperature in K, (b) refractivity, (c) refractivity lapse rate in  $\text{km}^{-1}$ , and (d) bending angles in milli-radians. Lower panel plots [(e), (f), (g), and (h)] show the standard deviations calculated within each bin respectively in K, in percents of refractivity, in percents of refractivity lapse rate, and in percent of bending angle. Note the log-scale for [(b), (c), and (d)].

tions within 3 hours of the analysis time in 7 different time-slots. The first and last time-slots are 30 minutes long; the remaining time-slots are 60 minutes long.

The following data-types were assimilated in the operational ARPEGE assimilation system as of October 2006: in-situ measurements from the surface (by land, ship, and buoys) and upper-air (via radiosondes and aircraft), atmospheric motion vector winds from geostationary and polar-orbiting imagery, wind profiles from European and American wind profilers, sea winds from the QuikSCAT scatterometer, zenith total delays from ground-based GPS stations over Europe, and brightness temperatures from the following passive radiometers: the Advanced Microwave Sounding Unit (AMSU)-A, AMSU-B, High Resolution Infrared Radiation Sounder (HIRS), Atmospheric InfraRed Sounder (AIRS), and Special Sensor Microwave/Imager (SSM/I) on-board (resp.) the NOAA-15, -16, -18, and NASA Aqua satellites, the NOAA-16, -17, and -18 satellites, the NOAA-17 satellite, the NASA Aqua satellite, and the DMSP-13 satellite.

## 4.2 GPS Radio Occultation Observation Operators

We map the information from forecast fields (or background) into GPS radio occultation observation space with the help of observation operators.

The bending angle observation operator used in this study is the same as that used by Healy and Thépaut (2006). The assumption of spherical symmetry enables to calculate the bending angle from a one-dimensional integration of the background refractivity lapse rate profile located at the vertical of the observation point. The computation of this integral from the background temperature, pressure, humidity and geopotential height information requires the product  $n$  times  $r$ . The refractive index  $n$  equals the unity plus  $10^{-6}$  times the radio refractivity  $N$  (calculated using the formula of Smith and Weintraub 1953). The radial distance  $r$  equals the geometric altitude (calculated from the geopotential height assuming an altitude and latitude-dependent gravity acceleration) plus  $R_c$  and

We added observation operators to compute refractivity and refractivity lapse rate as a function of impact parameter (see Appendix A), and temperature as a function of geopotential height from the background fields. The observation operator also returns for each bending angle observation the background layer index containing the observed impact parameter.

Our observation operators assume that the products ( $n.r$ ) calculated at background level interfaces decay with altitude. Consequently, when  $n.r$  increases locally with altitude in a background layer (i.e., ducting conditions, see von Engel and Teixeira 2004) we cannot compute background equivalent quantities for refractivity, refractivity lapse rate, and bending angle at any of the levels below.

## 5. COMPARISON OF THE F3C DATA WITH ARPEGE FORECASTS

The observation operators presented above are used to project the ARPEGE operational 6-hour forecasts onto the various observation spaces: temperature, refractivity, refractivity lapse rate, and bending angle. The forecasts represent the background onto which F3C data are to be assimilated later on. Before applying variational data assimilation algorithms we must evaluate the validity of an important underlying assimilation hypothesis regarding the F3C observations and the forecasts projected in the same space; namely, the two datasets must reproduce the same phenomena, within a known error range forming a Gaussian distribution whose moments are understood. The purpose of this section is to answer first the question of what are the mean and the standard deviation of the distribution of the differences between F3C data and their equivalent as calculated from the forecasts.

### 5.1 Temperature

Figure 3a shows the zonal cross-section of the mean difference between F3C temperature and background temperature. The background presents a cold bias in the tropics reaching more than 2 K at the tropical tropopause (this bias is known and was also observed with radiosondes) and a warm bias at high latitudes as compared to F3C temperature retrievals. We also note a vertically sign-oscillating bias over Antarctica (similar to that reported by Healy and Thépaut (2006) but in a different assimilation system). The F3C temperature information is invaluable in the sense that it provides us with a global estimate of the background temperature biases.

Figure 3e shows the standard deviation of the temperature differences between F3C retrievals and background. The largest deviations are observed in the moist troposphere, possibly due to a water vapour ambiguity in the temperature retrieval as well as measurement and back-

ground errors. Between the 10 and 20 km altitude (except around the tropical tropopause) the standard deviations reach below 1.0 K indicating that the background and F3C agree fairly well in that region. In contrast with Fig. 2e, we observe small deviations above the Southern high latitudes around 17 km altitude. This suggests that the temperature variability observed earlier in the observations is correctly simulated by the background and results from natural variability (and not from processing or retrieval error).

### 5.2 Refractivity

Figure 3b shows the zonal bias of the refractivity departures between F3C and the background. We can identify a region of negative bias in the moist troposphere which is similar to the so-called refractivity bias observed by other GPS radio occultation experiments (e.g., Beyerle et al. 2004). Over Antarctica we note again vertical dipoles of bias. Excluding the Southern high latitudes and the tropical tropopause, the bias is less than 0.5% between 8 and 20 km in altitude.

The standard deviations of the refractivity departures shown in Fig. 3f are less than 1% almost everywhere below 25 km altitude except in the moist troposphere.

### 5.3 Refractivity Lapse Rate

If horizontal gradients of refractivity are neglected, the refractivity lapse rate ( $dN/dz$ ) determines locally how much bending is to occur. We further use it as an indicator of whether, over a particular region, the F3C observations intrinsically saw the same physical refraction phenomena as the background. One caveat to this approach is that the refractivity retrieved under abnormal propagation conditions differs from the true refractivity (e.g., Xie et al. 2006).

Figure 3c shows the zonal mean differences in refractivity lapse rate between that calculated from F3C refractivity observations and that calculated from the background refractivities. Between 8 and 35 km in altitude, the differences remain within 3%, suggesting that F3C observations and background present similar vertical structures. We note an exception above Antarctica where we observe oscillations similar to those seen in the refractivities.

Figure 3g shows the standard deviation of the differences between the refractivity lapse rate from F3C and that calculated from the background. The agreement is within 9% between 8 and 35 km altitude, exceeding 15% around the tropical tropopause.

### 5.4 Bending Angle

Figures 3d and h show the zonal mean and standard deviation of the bending angles departures between F3C observations and background calculations. The mean agree-

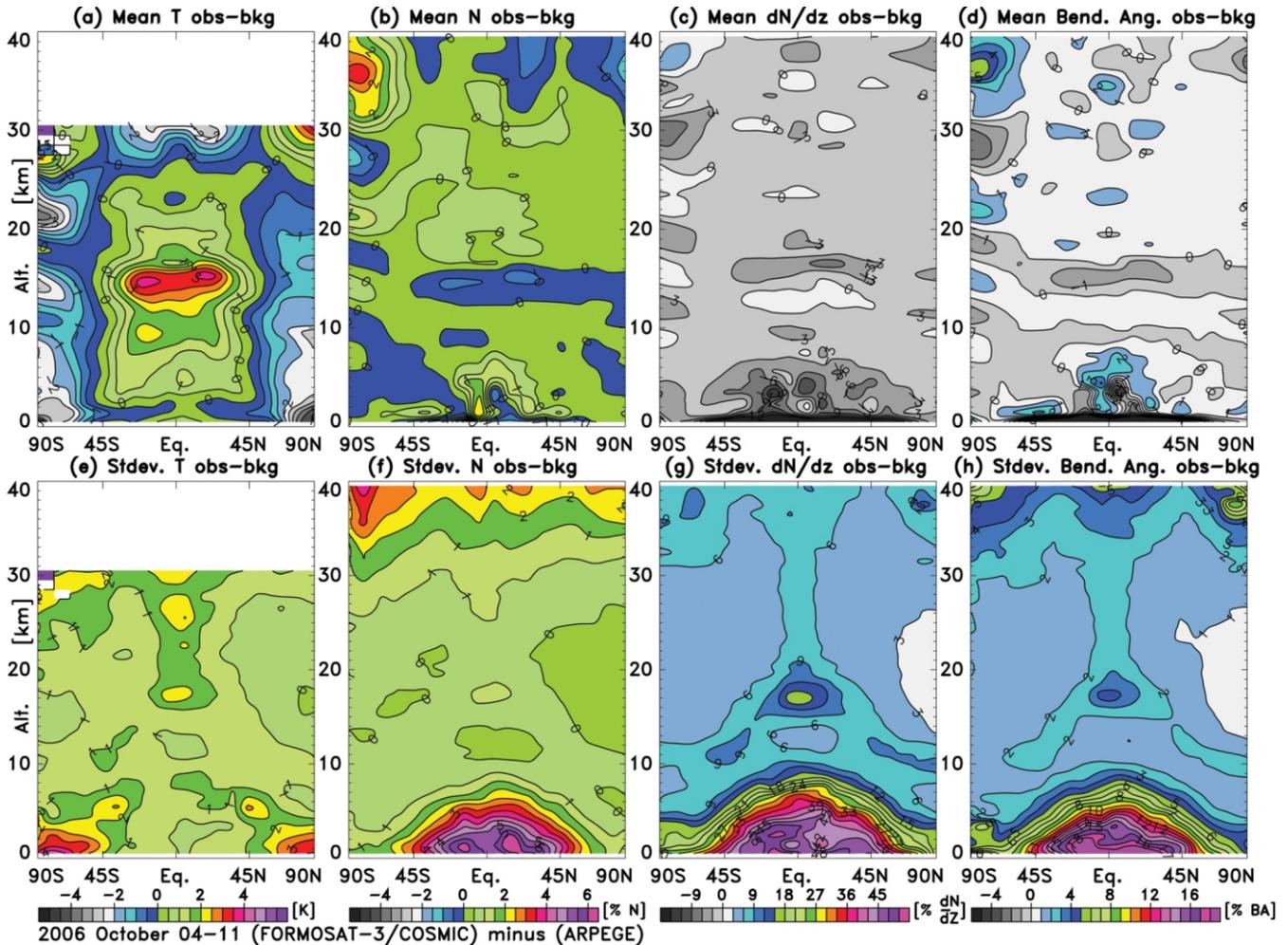


Fig. 3. The zonal averages of the differences between observations from F3C and calculations from the background for (a) temperature in K, (b) refractivity in percent difference, (c) refractivity lapse rate in percent difference, and (d) bending angles in percent difference. [(e), (f), (g), and (h)] show standard deviations. There are no F3C temperature retrievals above about 32 km in altitude. Statistics include all data except for refractivity lapse rate differences outliers in excess of 200%.

ment is best over 8 to 20 km altitude (excluding Antarctica), between -2% and +2%. The region located above 25 km altitude presents difference standard deviations increasing with height above 30 km altitude (possibly due to noise and ionospheric contamination). The moist troposphere indicates standard deviations in excess of 15%. Everywhere else the standard deviations remain below 3% (except near the tropical tropopause).

Overall, we observe that the standard deviations of departures for all quantities are larger in the Southern high latitudes than in the Northern high latitudes. This may be explained by the better accuracy of the ARPEGE forecasts in the Northern hemisphere.

## 6. QUALITY CONTROL PROCEDURES

### 6.1 Physical Considerations

The GPS radio occultation sounding technique relies on

the propagation of L-band radio waves within the Earth's atmosphere. Radar meteorologists typically consider that normal (radio) refraction occurs when the (radio) refractivity lapse rate  $dN/dz$  is between  $-79$  and  $0 \text{ km}^{-1}$  (Doviak and Zrnicek 1984). Below that range and down to  $-158 \text{ km}^{-1}$ , super-refraction is said to occur. Ducting (sub-refraction) may occur below  $-158 \text{ km}^{-1}$  (at or above  $0 \text{ km}^{-1}$ ). Based on the refractivity lapse rate calculated from F3C observations we find that about 27% of the F3C occultation events seem to include at least one of such abnormal propagation event. Our results (not shown here) indicate that most of these events occur in the lower troposphere, as expected.

### 6.2 Procedures and Validation

Attempting to show the effects of these events in the data, we show in Fig. 4a a scatter plot of the bending angle observation minus background departures versus altitudes.

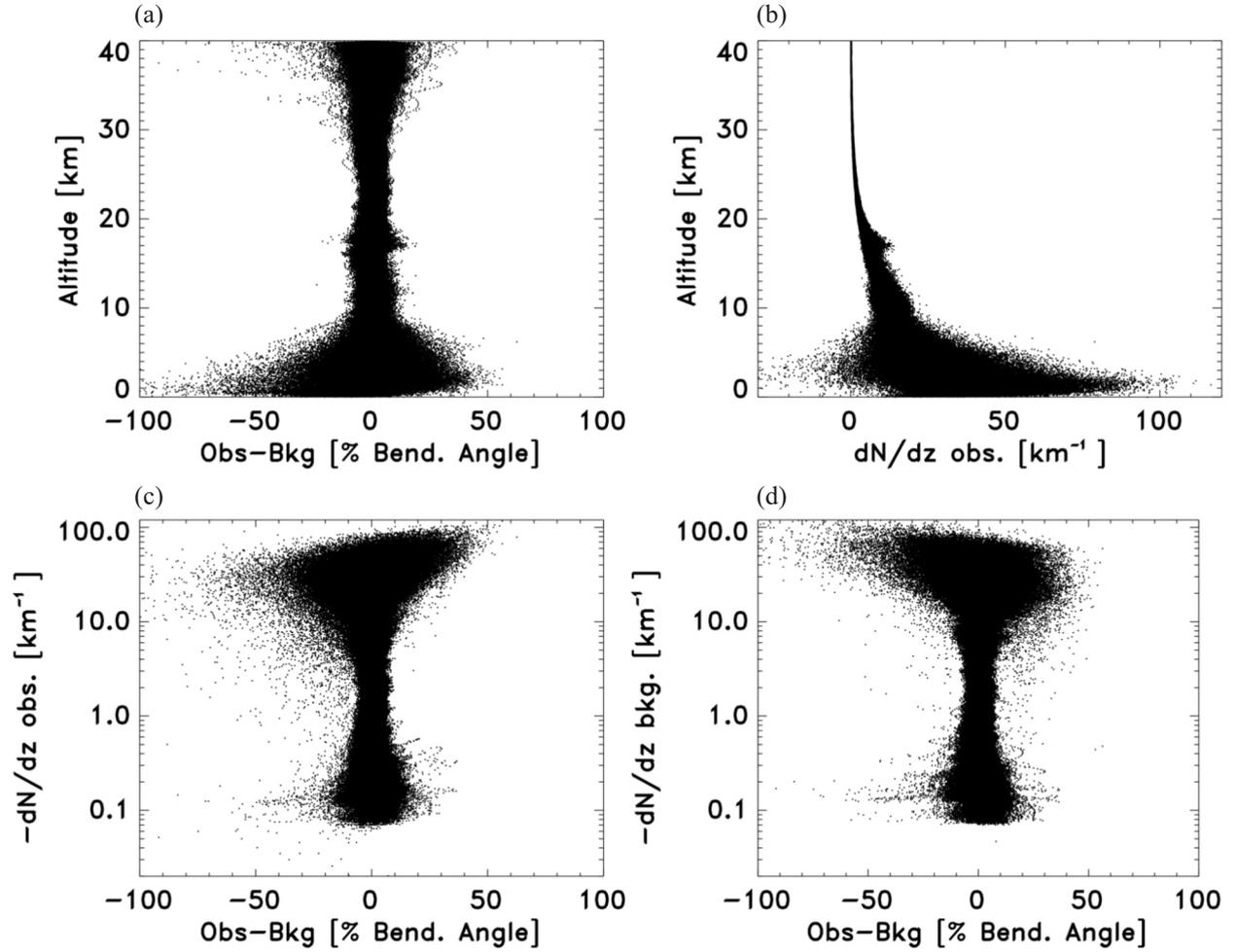


Fig. 4. (a) A scatter plot of the bending angle departure (F3C observation minus background) versus altitude; (b) scatter plot of the refractivity lapse rate measured by F3C versus altitude; (c) scatter plot of the bending angle departure versus the refractivity lapse rate measured by F3C; and (d) the same as (c) except that the refractivity lapse rate is calculated from the background. Occurrences of positive refractivity lapse rates are excluded from the plots (c) and (d).

We observe a negative tail of departures below 10 km altitude. There is no cloud contamination in the GPS radio occultation measurements unlike infrared measurements for which similar ‘cold tails’ departures are frequent without proper cloud-screening. However, this negative tail is troublesome if some data are not screened out before assimilation. One reason is that variational assimilation assumes Gaussian error distributions, as mentioned earlier in the introduction of section 5. Plotting in Fig. 4b a scatter plot of the refractivity lapse rates versus altitude, we also observe a spread with super and sub-refraction events. Attempting to find a possible relationship between abnormal refractivity lapse rates and large bending angle departures we show a scatter plot of one versus the other in Fig. 4c. For the strongest lapse rates (i.e., in the lower troposphere), we observe an anvil shape with a significant number of large negative bending angle departures. Using instead the background refractivity lapse rate in Fig. 4d we find again an anvil shape,

but reversed, indicating that the observations and the background each possess their own intrinsic lapse rate of physical limitations (leading to large positive or negative bending angle departures).

Consequently, we devise the following quality control procedure (called QC1). We flag as suspicious all data observations located below regions where  $dN/dz$  (from observation or background) falls below  $-50 \text{ km}^{-1}$ . We choose this conservative threshold to include the possibility that the F3C refractivity data may have indeed seen some abnormal propagation layers but whose effects may have been damped in the inversion process.

We also find that the derivative of the background and observation refractivity lapse rate with respect to height ( $d^2N/dz^2$ ) is useful to locate areas where the physical phenomena considered simulated by the background may differ from the observation conditions. In practice, we screen out all data for which  $|d^2N/dz^2|$  exceeds  $100 \text{ km}^{-2}$  (QC2).

Figure 5a shows the effect of taking into account QC1 and QC2 in the distribution of the bending angle departures between 0 - 2 km altitude. The effect of QC2 after QC1 is smaller than that of QC1 alone. Out of the data that pass both tests, the skew, standard deviation, and mean of the bending angle departures are reduced. The effect of both quality controls is beneficial at other altitudes (Figs. 5b - d).

## 7. BENDING ANGLE ERROR ANALYSIS

### 7.1 Methodology

Desroziers et al. (2005) developed a methodology to evaluate analysis, observation and background error covariance matrices in observation space (written respectively  $\mathbf{A}$ ,  $\mathbf{R}$ ,  $\mathbf{B}$ ) within a variational data assimilation system. For clarity, we remind readers that the analysis is the product of the assimilation of the observations with the background. Note that the observation error includes the measurement error as well as the representativeness error.

The method used by Desroziers et al. (2005) relies on the intrinsic filtering properties of the analysis process. The

assumptions are that the background and the observations present distinct error spectra; otherwise the method may not separate between either sources of error. The calculations involve the expectation value of products of the observation minus background and analysis departures, noted respectively as  $y^o-h(x_b)$  and  $y^o-h(x_a)$ :

$$E\left\{\left[h(x_a)-h(x_b)\right]\cdot\left[y^o-h(x_a)\right]^T\right\}=\mathbf{A} \quad (1)$$

$$E\left\{\left[y^o-h(x_a)\right]\cdot\left[y^o-h(x_b)\right]^T\right\}=\mathbf{R} \quad (2)$$

$$E\left\{\left[h(x_a)-h(x_b)\right]\cdot\left[y^o-h(x_b)\right]^T\right\}=\mathbf{B} \quad (3)$$

where  $h$  denotes the observation operator,  $x_b$  ( $x_a$ ) is the background (analysis) state, and  $y^o$  is the observation vector.

We first run the assimilation system ARPEGE with the addition of F3C bending angle data. We assimilate all the

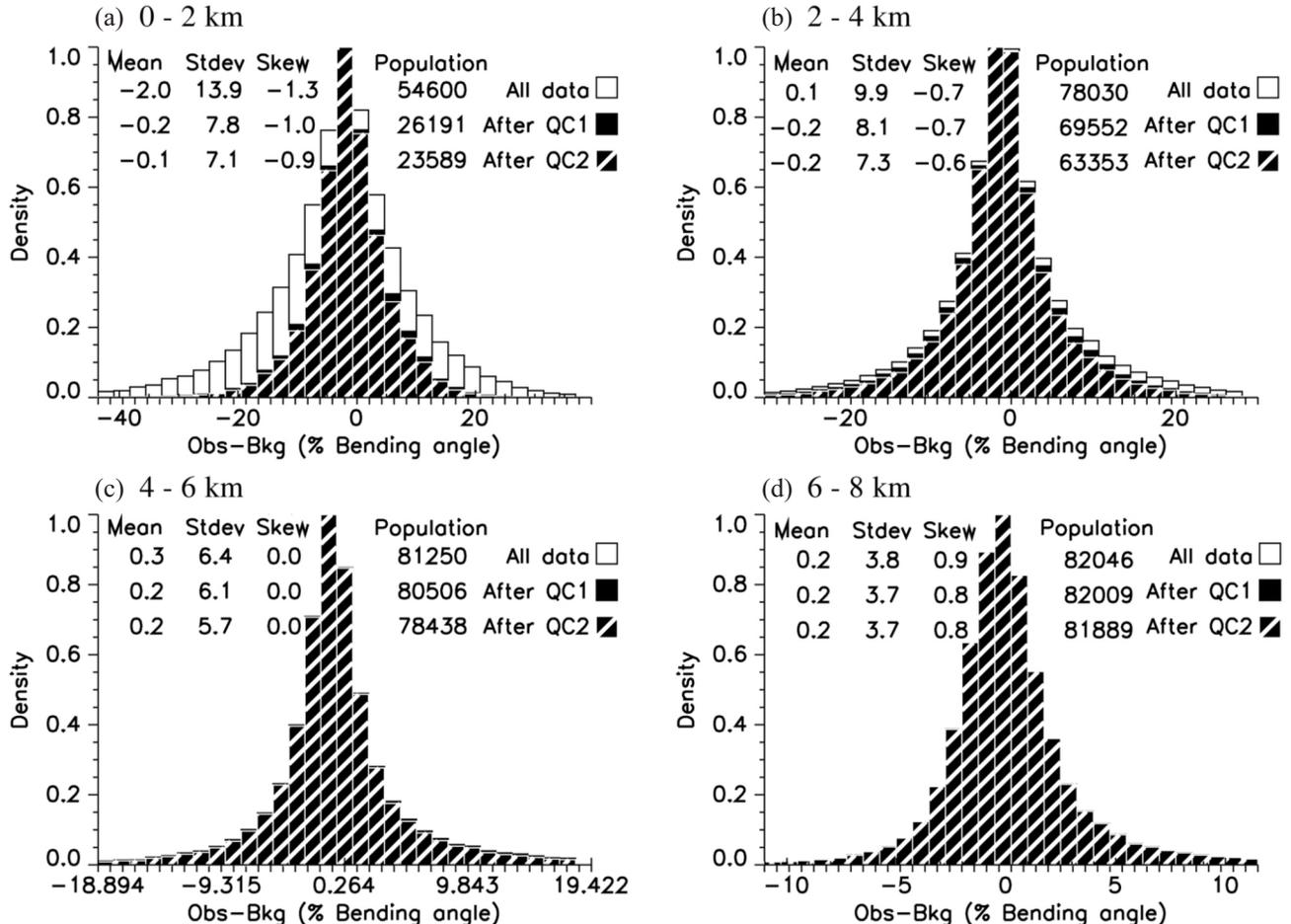


Fig. 5. Histograms of the bending angle departures between F3C observation and background with various quality controls and at various altitude ranges: (a) 0 - 2 km, (b) 2 - 4 km, (c) 4 - 6 km, and (d) 6 - 8 km.

F3C data that pass the quality controls above, for three non-cycling consecutive assimilations (i.e., using the operational forecast as a background). As a first estimate we assign (uncorrelated) bending angle observation errors of 1% above 10 km altitude and increasing linearly to 10% near the surface.

## 7.2 Results

In order to calculate the error covariance matrices we bin all the observations within the vertical to a resolution of 200 m (which is close to the F3C dataset vertical resolution). Figure 6a shows the square root of the diagonal of the  $\mathbf{R}$  matrix. The observation error standard deviation is found to be around 1.5% above 8 km in altitude and increasing toward the surface up to 8%. The background error standard deviation reaches a peak at 3 km altitude and is usually larger than the observation error above 2 km altitude. The analysis standard deviation error is found (as expected) below the background and the observation errors.

The observation error vertical correlation is shown at four different levels in Figs. 6b - e. The correlations in terms of bending angle are fairly narrow for levels above 5 km, about 1 to 2 km thick. The correlation in the lowermost level is found to be fairly broad. We further note that the vertical correlations exhibit anti-correlation patterns for the three levels shown here above 4 km in altitude. These anti-correlations confirm simulation studies by Syndergaard (1999) and Rieder and Kirchengast (2001). The negative correlations originate from the raw measurement of phases and the derivative operation necessary to derive a frequency shift from there. The vertical extent of the correlations may come from the smoothing applied during data processing.

Note that these results are subject to several points of caution. First, the hypotheses used above may not all be true. Second, the method could be applied within another assimilation system, for example, to confirm that observation errors remain the same. Involving other data-types to evaluate the observation errors is also an option.

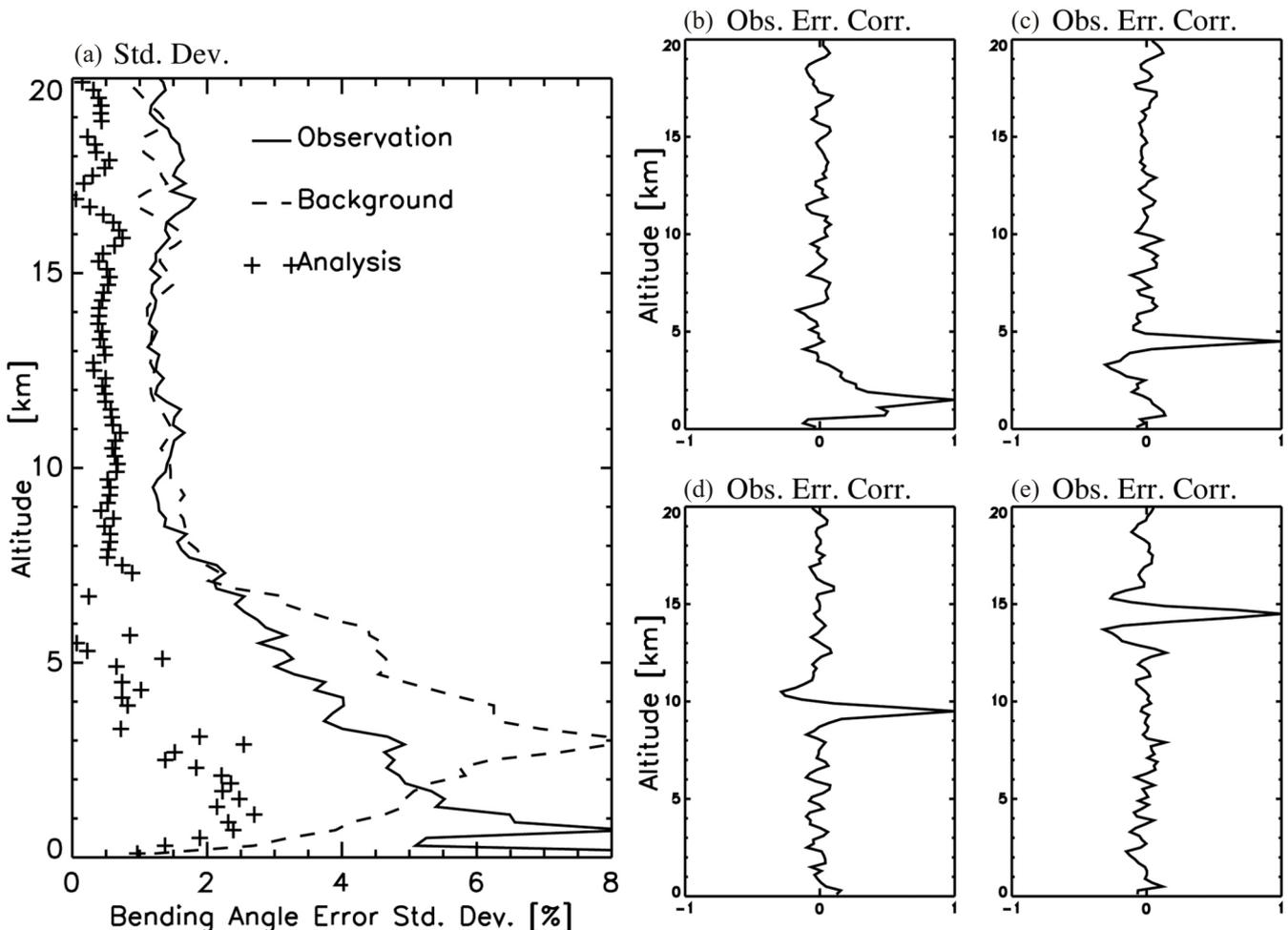


Fig. 6. (a) An estimate of the observation, background, and analysis errors in bending angle space. (b) An estimate of the bending angle error observation correlations at 1.5 km altitude [where (c), (d), and (e) correspond to 4.5, 9.5, and 14.5 km in altitude, respectively].

## 8. ASSIMILATION EXPERIMENT OF F3C BENDING ANGLES

### 8.1 Experimental Setup

We use as baseline the operational configuration of the ARPEGE assimilation and forecast system presented earlier. We add to this reference the assimilation of F3C bending angles between the altitudes of 5 and 15 km. We limit the vertical coverage in this first experiment as we note that adding more data below 5 km modifies the standard deviation of the analysis fit to the other observations (including radiosondes), suggesting that our partitioning between bending angle observation errors and other observation datatype errors needs further work. As for the higher limit (15 km), it corresponds to the upper level of the prime region of interest of the users of the ARPEGE operational model, i.e., the lower stratosphere and the troposphere. Furthermore the vertical resolution of our background becomes coarser (from 2 to 4 km) between altitudes of 15 and 25 km. Finally, we observe that assimilating the bending angle data above 15 km changes significantly the mean background fit to brightness temperature radiance observations. This is explained by our brightness temperature radiance bias correction being not adaptive but could have been solved with an iterative bias correction tuning (a significant undertaking in itself). As an assimilation system with increased vertical resolution and variational bias correction of the radiances is currently being developed, we plan to soon revise the choice of restricting the assimilation to 5 - 15 km in the vertical.

In line with Fig. 6a, we assume observation error standard deviations of 2% above 10 km altitude and increasing linearly to 3% at 5 km altitude.

A thinning procedure is applied because our assimilation system cannot yet handle spatial observation error correlations. The horizontal thinning is a two-step procedure similar to that employed for other satellite observations. The first step bins all bending angle observations in 0.60 degree latitude and longitude square boxes and retains one observation in each box. The second step repeats this operation but only on those observations that were selected in the first pass, and in 1.25 degree latitude and longitude boxes (about 139 km horizontal resolution at the Equator). After this we also perform a vertical thinning specific to bending angle observations. We retain only one bending angle observation per occultation profile and per background layer index, as determined by the observation operator described in section 4.2. The selection of one observation per horizontal bin and per vertical layer is random so that no preferential selection is applied.

The quality control procedures introduced above are used. Another test consists in removing all the observations whose departure from the background exceeds a certain threshold. That threshold is about three times the observa-

tion error standard deviation, or ~6% above 10 km altitude and increasing linearly to ~9% at 5 km altitude.

The experiment assimilating the F3C bending angles is called ARPF3C, and run from 0000 UTC 4 October 2006 to 1800 UTC 10 November 2006. A total of 33 four-day forecasts are issued daily from the analysis at 0000 UTC.

### 8.2 Results

Figures 7a and b show the background departures statistics for all bending angle data (for the first week of assimilation), before and after application of the quality control. The thinning procedures reduce the data density to a lower quantity. Figure 7c shows that only a fraction of all the F3C data is used in the present study. After assimilation, the standard deviations of the departures are reduced between 5 and 15 km in altitude. The limited reduction is an effect of the conservative observation error estimate.

We now define the following concepts to present the forecast impact results. The ‘forecast RMS score’ is the RMS of differences between verification data and forecast data. The ‘forecast RMS score difference’ is the difference of the forecast RMS score for the reference minus the forecast RMS score for the ARPF3C experiment. A positive forecast RMS score difference indicates that the ARPEGE reference contains larger forecast errors than the ARPF3C experiment (positive impact). Likewise, a negative forecast RMS score difference indicates degradation in the forecasts assimilating the F3C bending angle data.

Figure 8 shows the forecast RMS score difference for the Southern, tropical, and Northern latitudes. The verification data used here are geopotential heights from each analysis. This approach enables to have many more points of comparison and verify the stability of the results with a bootstrap test. The Northern hemisphere shows little or no impact. The Southern hemisphere results are much more pronounced with a gain in forecast RMS of more than 2 mgp around the tropopause at day 4.

Figure 9 shows the forecast RMS score difference using geopotential heights from radiosondes as verification. Again, the Southern hemisphere results are positive, while the Northern hemisphere shows a (generally positive, but small) neutral impact. The tropics are more subject to discussion, as the negative regions of impact seem to have gained room in the plot. The signs of these results are found to be significant for most altitudes and most forecast lead times with a sign test.

Finally, Fig. 10 shows the forecast RMS score difference using winds from radiosondes as verification. The forecast impact measured using that metrics is more visible in the Northern hemisphere, while it remains strong in the Southern hemisphere. The results are more mixed in the tropics.

Overall, the maximum positive impact is found in the

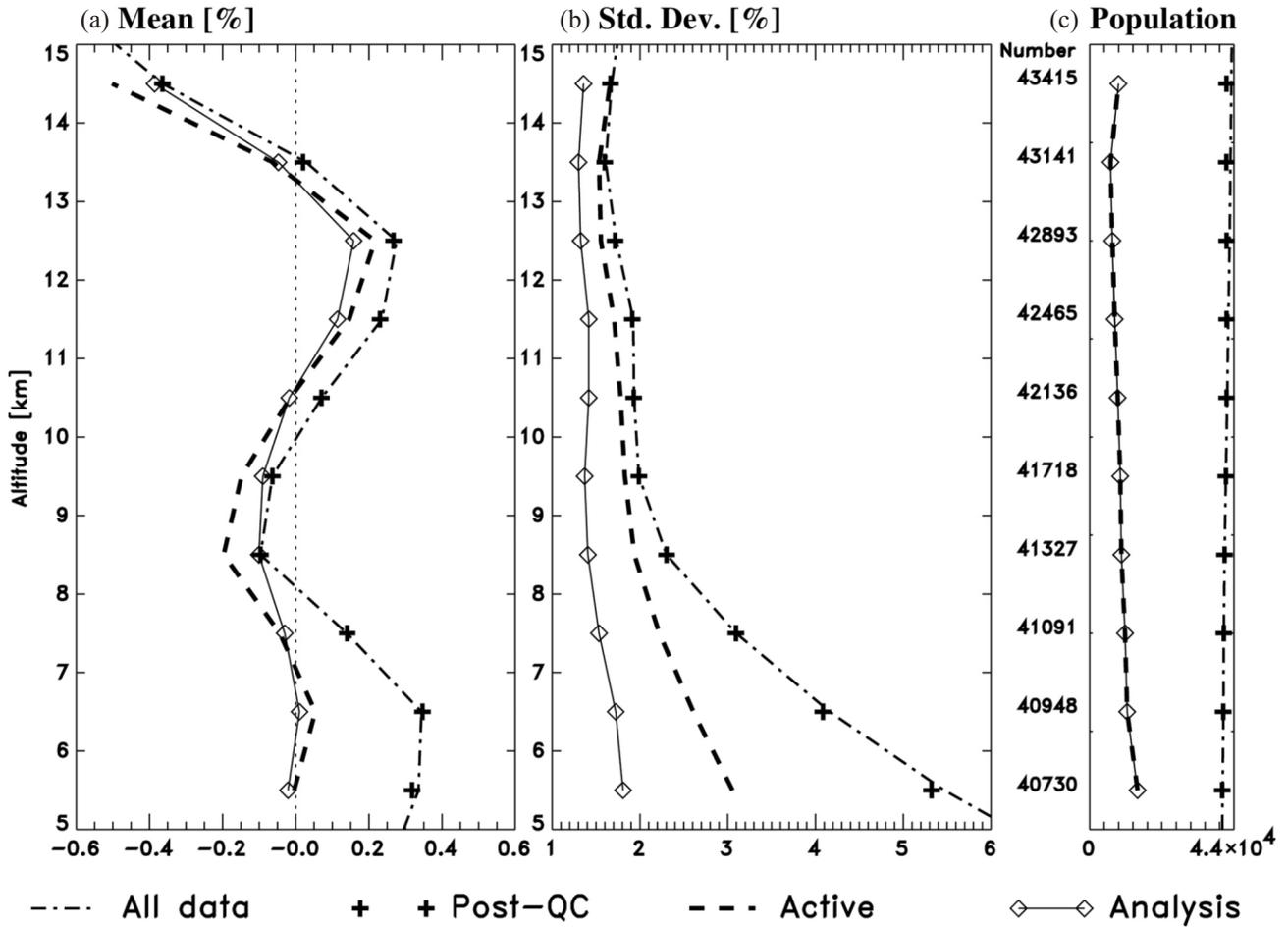


Fig. 7. Global statistics of differences between F3C bending angle observations minus background ('all data' 'post-QC': only those data selected by the QC 'active': only those data selected for assimilation) or analysis in the ARPF3C experiment. The list under 'numbers' indicates the size of the total population ('all data'), binned to 1 km vertical resolution.

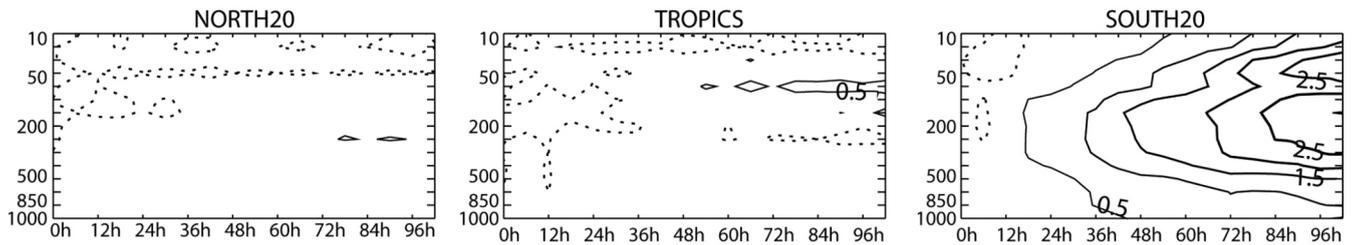


Fig. 8. Forecast RMS score differences of the ARPF3C experiment as compared to the operational ARPEGE system for geopotential height. Analyses from each run are used as verifications (contour 0.5 mgp). The solid (dotted, dashed) contours indicate areas of positive (respectively: neutral, negative) impact. X-axis: forecast lead time (hours), y-axis: pressure level (hPa).

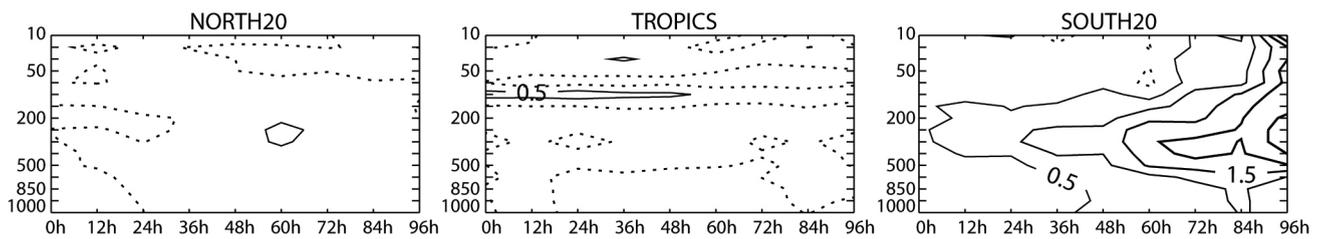


Fig. 9. Same as Fig. 8 but using radiosondes as verifications.

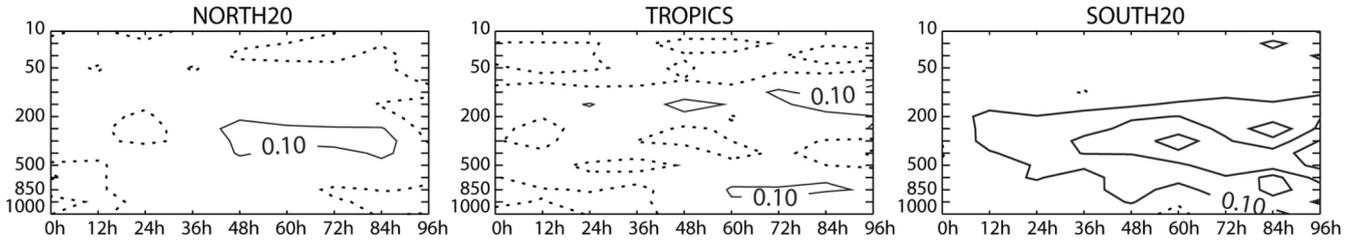


Fig. 10. Same as Fig. 9 but for wind (contour  $0.1 \text{ m s}^{-1}$ ).

upper troposphere and around the tropopause. The results in the tropics need further investigation and/or experimentation to be conclusive.

## 9. CONCLUSIONS AND FUTURE WORK

The F3C mission has been providing a global dataset of GPS refraction observations since mid-2006. We have assessed these data processed by UCAR in a comparison with the global operational assimilation and forecast system of Météo-France (ARPEGE). Looking at four different levels (bending angle, refractivity, refractivity lapse rate, temperature), we have found that the F3C data are close to the forecasts issued by ARPEGE except in the moist troposphere. We propose thresholds for quality control procedures when investigating the distributions of refractivity lapse rates. These checks enable researchers to reduce the skew and standard deviation in the bending angle observation minus background distributions. Capitalizing upon a method recently developed by Desroziers et al. (2005) we derive estimates of bending angle observation error standard deviations and vertical correlations. We confirm the presence of negative correlations in bending angle observation errors (as suggested by previous authors in their simulations) and the overall rather sharp vertical correlations.

Using these elements as well as a horizontal and vertical thinning we run an assimilation experiment of F3C bending angle data for about a month. The forecasts issued by this experiment prove closer to radiosondes and verifying analyses in the southern hemisphere for geopotential heights. We also find an impact on the prediction of winds in the Northern and Southern hemispheres' upper troposphere. The overall forecast impact in the tropics is more mixed.

This first study indicates that the F3C data do contain the promise to help improve numerical weather prediction forecasts. Based on the positive results shown here and thanks to the recent delivery of F3C data on the GTS and in near-real time we plan to make use of these observations as a part of the Météo-France operations in the near future.

It must be noted that the work presented here was largely enabled by the dense daily coverage of observations allowed of the F3C constellation. This has made it easier to identify, investigate (solve) the benefits (challenges) of these data.

This work will help the assimilation of data from single-satellite GPS radio occultation missions such as GRACE, CHAMP, and GRAS on MetOp.

Another anticipated application of F3C data in our operational assimilation system (besides assimilation) is the use of the F3C temperature retrievals to help monitor the upper-air fields, for which we have otherwise few verification data at high vertical resolution.

Regarding future developments, higher resolution experiments of the ARPEGE system are currently being performed with 60 levels instead of 46. These might help lower the lower altitude limit at which we assimilate the F3C observations in the future.

**Acknowledgements** The authors would like to thank the FORMOSAT-3/COSMIC program and funding agencies for making the mission possible and the products freely available within a few weeks after launch. Contributors at NASA/JPL are also acknowledged for their scientific support of the various radio occultation missions that have preceded FORMOSAT-3/COSMIC. The CDAAC retrievals were used in this work. The observation operator for bending angles was provided by the EUMETSAT GRAS-SAF. The paper benefits from constructive remarks from attendees at the Third CNES Earth-Space Propagation Workshop held in Toulouse, France on 25 - 27 September 2006 and at the FORMOSAT-3/COSMIC Workshop 2006: Early Results and IOP Campaign held in Taipei, Taiwan on 28 November - 1 December 2006. Thanks are extended to two anonymous reviewers whose comments helped improve the initial manuscript, and to Piin-Shiow Wang for careful proof-reading of the final manuscript.

## REFERENCES

- Beyerle, G., J. Wickert, T. Schmidt, and C. Reigber, 2004: Atmospheric sounding by global navigation satellite system radio occultation: An analysis of the negative refractivity bias using CHAMP observations. *J. Geophys. Res.*, **109**, D01106, doi: 10.1029/2003JD003922. [[Link](#)]
- Desroziers, G., L. Berre, B. Chapnik, and P. Poli, 2005: Diagnosis of observation, background, and analysis error statistics in observation space. *Quart. J. Roy. Meteor. Soc.*,

- 131, 3385-3396, doi:10.1256/qj.05.108. [\[Link\]](#)
- Doviak, R. J. and D. Z. Zrnic, 1984: Doppler radar and weather observations, Academic Press, Orlando, 458 pp.
- Gauthier, P. and J.-N. Thépaut, 2001: Impact of the digital filter as a weak constraint in the preoperational 4DVAR assimilation system of Météo-France. *Mon. Wea. Rev.*, **129**, 2089-2102, doi: 10.1175/1520-0493(2001)129<2089:IOTDFA>2.0.CO;2. [\[Link\]](#)
- Hajj, G. A., M. de la Torre Juarez, B. A. Ijima, E. R. Kursinski, A. J. Manucci, and T. P. Yunck, 2002a: GPS radio occultations coming of age: Spacecraft launches add two new instruments for climate monitoring. *EOS Trans. AGU*, **83**, 37 pp, doi: 10.1029/2002EO000028. [\[Link\]](#)
- Hajj, G. A., E. R. Kursinski, L. J. Romans, W. I. Bertiger, and S. S. Leroy, 2002b: A technical description of atmospheric sounding by GPS occultation. *J. Atmos. Sol.-Terr. Phys.*, **64**, 451-469, doi: 10.1016/S1364-6826(01)00114-6. [\[Link\]](#)
- Healy, S. B. and J.-N. Thépaut, 2006: Assimilation experiments with CHAMP GPS radio occultation measurements. *Quart. J. Roy. Meteor. Soc.*, **132**, 605-623, doi: 10.1256/qj.04.182. [\[Link\]](#)
- Janiskova, M., J.-N. Thépaut, and J.-F. Geleyn, 1999: Simplified and regular physics parametrizations for incremental four-dimensional variational assimilation. *Mon. Wea. Rev.*, **127**, 26-45, doi: 10.1175/1520-0493(1999)127<0026:SARPPF>2.0.CO;2. [\[Link\]](#)
- Kursinski, E. R., G. A. Hajj, T. Schofield, R. P. Linfield, and K. R. Hardy, 1997: Observing Earth's atmosphere with radio occultation measurements using the global positioning system. *J. Geophys. Res.*, **102**, 23429-23465, doi: 10.1029/97JD01569. [\[Link\]](#)
- Rieder, M. J. and G. Kirchengast, 2001: Error analysis and characterization of atmospheric profiles retrieved from GNSS occultation data. *J. Geophys. Res.*, **106**, 31755-31770, doi: 10.1029/2000JD000052. [\[Link\]](#)
- Smith, E. K. and S. Weintraub, 1953: The constants in the equation for atmospheric refractive index at radio frequencies. *Proc. IRE*, **41**, 1035-1037, doi: 10.1109/JRPROC.1953.274297. [\[Link\]](#)
- Syndergaard, S., 1999: Retrieval analysis and methodologies in atmospheric limb sounding using the GNSS radio occultation technique. Ph.D. Thesis, Danish Meteorol. Inst., Copenhagen, Sci. Rep., **99-6**.
- von Engel, A. and J. Teixeira, 2004: A ducting climatology derived from ECMWF Global analysis fields. *J. Geophys. Res.*, **109**, D18104, doi: 10.1029/2003JD004380. [\[Link\]](#)
- Ware, R., M. Exner, D. Feng, M. Gorbunov, K. Hardy, B. Herman, Y. H. Kuo, T. Meehan, W. Melbourne, C. Rocken, W. Schreiner, S. Sokolovskiy, F. Solheim, X. Zou, R. Anthes, and S. Businger, 1996: GPS sounding of the atmosphere from low earth orbit: Preliminary results. *Bull. Amer. Meteor. Soc.*, **77**, 19-40, doi: 10.1175/1520-0477(1996)077<0019:GSOTAF>2.0.CO;2. [\[Link\]](#)

- Xie, F., S. Syndergaard, E. R. Kursinski, and B. M. Herman, 2006: An approach for retrieving marine boundary layer refractivity from GPS occultation data in the presence of superrefraction. *J. Atmos. Ocean. Technol.*, **23**, 1629-1644, doi: 10.1175/JTECH1996.1. [\[Link\]](#)

## APPENDIX A

Observation operators for refractivity and refractivity lapse rate as a function of impact parameter

The impact parameter (noted  $a$ ) is used as the independent vertical coordinate. First we find the background layer in which  $a$  is located. That background layer is bounded by two interfaces with refractivities  $N_1$  and  $N_2$  (and refractive indices  $n_1$  and  $n_2$ ), with radii  $r_1$  and  $r_2$ , such that

$$n_1 r_1 < a < n_2 r_2 \quad (\text{A1})$$

As noted earlier, this assumption only works as long as the products ( $n \cdot r$ ) decay with altitude in the background (for altitudes below ducting layers, the observation operator cannot be applied). The refractivity  $N(a)$  is then calculated by assuming an exponential decay with respect to the impact parameter:

$$N(a) = \exp \left[ -\ln \left( \frac{N_1}{N_2} \right) \cdot \frac{a - n_1 r_1}{n_2 r_2 - n_1 r_1} \right] \quad (\text{A2})$$

The altitude  $z$  enters this equation via the impact parameter definition

$$a = [1 + 10^{-6} N(a)] (R_c + z) \quad (\text{A3})$$

using the notations defined in section 2.2. Replacing  $a$  by its definition and taking the natural logarithm of (A2), we can calculate the total derivative on both sides

$$\frac{dN}{N(a)} = -\ln \left( \frac{N_1}{N_2} \right) \cdot \frac{1}{n_2 r_2 - n_1 r_1} \cdot \left[ dz + 10^{-6} N(a) \cdot dz + (R_c + \Delta + z) \cdot dN \cdot 10^{-6} \right] \quad (\text{A4})$$

We finally rearrange this expression to yield the refractivity lapse rate

$$\frac{dN}{dz} = \left\{ \frac{n_2 r_2 - n_1 r_1}{\ln \left( \frac{N_2}{N_1} \right) \cdot N(a) \cdot [1 + 10^{-6} N(a)]} - \frac{10^{-6} a}{[1 + 10^{-6} N(a)]^2} \right\}^{-1} \quad (\text{A5})$$