Impact of the Assimilation of CHAMP Refractivity Profiles on Environment Canada Global Forecasts

JOSEP M. APARICIO AND GODELIEVE DEBLONDE

Meteorological Research Division, Atmospheric Science and Technology Directorate, Environment Canada, Dorval, Quebec, Canada

(Manuscript received 10 July 2006, in final form 30 March 2007)

ABSTRACT

The data assimilation system of Environment Canada (EC) is adapted to accept GPS radio occultation (GPSRO) data. Observations of this type are available with extensive coverage from several satellites. In this study, experiments are performed to compare the skill of EC’s three-dimensional variational data assimilation (3DVAR) system (including all data normally assimilated operationally), with and without the addition of radio occultation refractivity data from the Challenging Minisatellite Payload for Geophysical Research (CHAMP). These data were not available at the time studied as near-real-time (NRT) observations. However, data from this and other radio occultation missions are now available as NRT data, and the conditions (latency, reliability) are improving. It is expected that NRT GPSRO data from a number of satellite missions will continue to be available through the following years. The results of the assimilation tests are evaluated against the following three data types: radiosondes (temperature and dewpoint depression), satellite brightness temperatures (from the Advanced Microwave Sounding Unit-A), and GPS radio occultation refractivity profiles. For the 6-h forecasts, the differences between GPSRO observations and forecasts \( O/H11002 F \) are significantly reduced in the experiment that assimilates the GPSRO data. This reduction increases as the experiment proceeds in time, and stabilizes after a transient period of approximately 2 weeks, suggesting that the addition of GPSRO data to the assimilation system has a beneficial, persisting, and cumulative effect. This effect is more pronounced in the stratosphere than in the troposphere. In the stratosphere, the standard deviation of GPSRO \( O - F \) of the experiment that assimilates GPSRO decreases after the initial transient period by approximately 10%. This improvement can best be observed in the southern stratosphere where reductions of the order of 30% are common. This shows that, as a globally distributed and vertically well-resolved source of data, the GPSRO observations are not only useful for assimilation, but also as a tool to quantify the forecast skill of the assimilation system. Comparisons with radiometer and radiosonde data confirm the positive impact in these geographical areas. Longer-range forecasts (up to 6 days) also show a positive impact with similar geographical and altitude distribution.

1. Introduction

The radio occultation (RO) technique for the remote sensing of atmospheric profiles was originally applied to the study of planetary atmospheres other than that of the earth; notably, profiling the atmospheres of Mars (Fjeldbo and Eshleman 1968; Kliore et al. 1972) and Venus (Fjeldbo et al. 1971) with this technique started several decades ago. These measurements could be performed only during the rare visits of interplanetary probes, and their objective was to obtain an overall estimation of the climate of the planets being probed.

The RO technique can also be applied to the earth (Melbourne et al. 1994). However, because the earth’s atmosphere is already quite well known, this is not of practical use unless the radio source and the receiver can be characterized to a very high level of accuracy, effectively measuring not just the mean climate but also the weather. Any practical measurement of the weather must be frequent, with a short latency, and distributed as globally as possible. Because of the development of radionavigation satellite systems, and especially the global positioning system (GPS) (Parkinson and Spilker 1996), these conditions have been met and have provided the required large-scale source of well-characterized signals, justifying the development of programs using radio occultation for weather remote sensing.

The pioneering GPS/Meteorology (MET) mission
(Ware et al. 1996; Rocken et al. 1997) showed that it was indeed possible to place a specially designed GPS receiver in a low earth orbit (LEO) to obtain atmospheric data sufficient enough both in quantity and quality to impact operational meteorology. Later scientific missions such as the Challenging Minisatellite Payload for Geophysical Research (CHAMP; Wickert et al. 2001a), the Satélite de Aplicaciones Científicas-C (SAC-C; Hajj et al. 2004), and the Gravity Recovery and Climate Experiment (GRACE; Wickert et al. 2005; Beyerle et al. 2005), still in operation today, provide large samples of quality data. In particular, the CHAMP mission, with its focus on the completeness and homogeneity of the dataset, offers a nearly uninterrupted series of observations since 2001. The recently launched Constellation Observing System for Meteorology, Ionosphere and Climate (COSMIC) constellation of six satellites (Rocken et al. 2000) and the Global Navigation Satellite System (GNSS) Receiver for Atmospheric Sounding (GRAS) mission onboard Meteorological Operational (METOP) satellites (Loiselet et al. 2000) provide even larger amounts of data. Compared to former missions, these are designed with tighter requirements toward fulfilling operational needs, mainly of timeliness in data downlink, processing, and delivery to the end users. Effort has also been made to reduce the latency of data from CHAMP and GRACE, to meet near-real-time (NRT) requirements (J. Wickert 2006, personal communication).

The radio occultation raw measurements (Kursinski et al. 2000) are based on fundamental physical properties, such as the speed of light or the refractive index of a propagating medium. The postprocessing of the raw measurements involves fundamental relationships between physical magnitudes, such as thermodynamics or hydrostatic equilibrium. These measurements are close to being unbiased because the whole process is essentially free of instrument-response calibrations or ad hoc phenomenological models. The impact of cloudiness, rain, and other adverse meteorological conditions is also very small (Parkinson and Spilker 1996). As a result, it has been argued (Yuan et al. 1993) that this kind of observation is not only useful as an operational source of weather data, but also as a reference (e.g., for long-term climate studies) that requires little or no calibration. This type of data is of particular importance for finding and correcting biases in other sources of data, or in the meteorological models themselves. This does not mean that GPS radio occultation (GPSRO) measurements are absolutely free of bias; some approximations are made during data processing that may introduce bias, in particular, that the atmosphere is locally spherically symmetric. Some atmospheric situations, especially during very moist severe weather episodes, are also more prone to loss of low-tropospheric data, which introduces some sampling bias. The forward operators can also introduce bias associated with model approximations such as the finite size of the horizontal or vertical grid. In general, we should expect that the magnitude of these effects is small when compared to the bias that many other remote sensing techniques present, notably brightness temperature measurements. Indeed, the usefulness of brightness temperature observations critically depends on a good instrument calibration (see, e.g., Rao et al. 1993; Prabhakara et al. 2000).

The radio occultation data can be assimilated under different forms (e.g., Kuo et al. 2000) and with different levels of postprocessing. These levels are from rawer to more processed as follows: excess phase or Doppler as a function of time, bending angle as a function of the signal’s impact parameter, or refractivity as a function of mean sea level (MSL) altitude. It is generally preferable to assimilate a rawer form. Some rawer operators (two-dimensional bending angle) have been explored by some authors (e.g., Zou et al. 1999; Liu et al. 2001), but these are complex for operational applications because the central processing unit (CPU) cost is high for the benefit obtained. Simpler operators (one-dimensional bending angle) can be used instead, which in principle have a comparable impact but at a lower CPU cost, and have been chosen for operational implementation in several institutions (e.g., Healy and Thépaut 2006). Another option is to assimilate vertical profiles of refractivity. The CPU cost is even lower, and the impact could still be beneficial. This option has been chosen in a number of studies on GPSRO assimilation (e.g., Eyre 1994; Zou et al. 1995; Kuo et al. 2000; Healy et al. 2005). Each step of postprocessing includes small but nonnegligible approximations. The evaluation of the bending angle from Doppler data requires assuming spherical symmetry in the atmosphere (e.g., Kursinski et al. 2000). The inversion of bending angle data to refractivity again requires this assumption, this time with tighter constraints. However, despite the bias that these approximations may introduce, the data may still be useful because their accuracy is still very good and has an excellent geographical distribution.

In this study, the impact of the assimilation of the GPSRO data is tested in an environment comparable to that of an operational system, following similar work by Healy et al. (2005). The chosen GPSRO data are the refractivity profiles. The objective of this work is to explore the impact that is achievable at a minimal CPU cost. A number of statistics are evaluated to determine whether the skill of the forecast system improves or degrades after the addition of the refractivity profiles.
A priori (e.g., Hoeg et al. 1996), a positive impact on the analysis/forecast system, if any, can be expected because of the assimilation of high-quality data in regions that are otherwise not extensively sampled, such as ocean areas, particularly in the Southern Hemisphere.

It may not be easy to detect this impact, because the very absence of independent data puts a limit to our ability to observe the impact, be it an improvement or degradation. Indeed, previous works have found that the impact, as measured by radiosondes, is positive but very small (Healy et al. 2005; Healy and Thépaut 2006). In this work, supplementary evaluations are performed to search for impacts that may be difficult to detect with standard radiosonde evaluations. Thus, one of the measures of skill of the assimilation experiments makes use of the GPSRO data, which has an excellent geographical distribution. The statistics of the departure between observations \((O)\) and forecasts \((F)\) are evaluated. Because only one of the experiments assimilates GPSRO data, and these data are also used to monitor the skill, the evaluation is performed with the background 6-h forecast fields, when observations have not yet been assimilated, that is, \((O - F)\), rather than with the analysis \((A)\) fields, that is, \((O - A)\).

In section 2, the assimilation environment, the GPSRO dataset, and the experiments are described. Evaluations of the reference experiments (no GPSRO data) and GPSRO experiments (with GPSRO data) are performed against GPSRO data, satellite radiometer data [National Oceanic and Atmospheric Administration (NOAA) Advanced Microwave Sounding Unit-A (AMSU-A)], and radiosonde data. These are, respectively, presented in sections 3a, 3b, and 3c. A comparison is made between the results of each data type and shows that some features are common. Section 3d presents an evaluation of the performance for longer-range forecasts up to 6 days. Finally, conclusions and a discussion are presented in section 4.

2. Methodology

a. Assimilation system

The assimilation system of Environment Canada (EC; Gauthier et al. 1999), run operationally for the production of EC’s forecasts, was chosen for this study. The system can be run (Laroche et al. 2005) in three- or four-dimensional variational data assimilation (3DVAR or 4DVAR, respectively) mode. Until 15 March 2005, the operational configuration consisted of an incremental 3DVAR system, at which time an incremental 4DVAR became operational. This study is performed with the 3DVAR, replicating the operational configuration during the time periods covered by the experiment in this study (i.e., 2004, see below). EC’s 3DVAR system is a cyclic assimilation system that alternates between 6-h forecasts, performed with EC’s Global Environmental Multiscale (GEM) model (Côté et al. 1998), and variational data assimilation at the time of validity of the forecasts. The system has a globally uniform latitude–longitude grid \((0.9^\circ, \text{or} \sim 100 \text{ km at the equator})\), a terrain-following vertical coordinate \(\eta\) with 28 vertical levels, and a model top at 10 hPa (about 30 km). The vertical levels are not evenly spaced and are denser at low altitude. The approximate vertical spacing in the region where RO is expected to have a larger impact (MSL altitude of 10–20 km) is of the order of 1 km.

The variational assimilation step (i.e., the 3DVAR) takes data from a 6-h window centered on the target analysis time. All GPSRO observations collected in this time window are background checked (i.e., observations that are too far from the 6-h forecast are rejected) and included in the assimilation.

b. Radio occultation data

Radio occultation data (with nearly full and continuous coverage) are available since 2001 from the CHAMP satellite. GPSRO data could be assimilated as either phase or Doppler data, bending angles, or refractivity profiles. This work explores the forecast skill improvement under the restriction of a low CPU cost. This excludes the assimilation of phase and Doppler data, as well as two-dimensional (or higher) bending angle operators.

In the model configuration of this study the model top is located at 10 hPa (about 30 km). Because the bending angle operator requires integrating the bending from the tangent point upward, these integrals will be inaccurate if the tangent height is near the model top. The observation operator for the refractivity, however, is not affected by the model top. Instead, the inversion from the bending angle to refractivity, made during preprocessing, assumes that the atmosphere has local spherical symmetry. The accuracy of this assumption is limited by the presence of horizontal gradients, which are normally small in the stratosphere but can be significant in the troposphere (Kuo et al. 2000) and are associated with weather fronts or moisture variations. Thus, taking advantage of theoretically superior bending angle data is limited by the horizontal resolution of the model (here \(0.9^\circ\), which limits the ability to represent sharp horizontal gradients), by the vertical resolution, and by the low model top. To compute the background estimate of the bending angle, which depends on refractivity gradients, requires a better knowledge of the refractivity field than the computation of the refract-
tivity. For all of the above reasons, refractivity profiles, rather than bending angles, were selected as the GPSRO data to be assimilated.

The data were obtained from the postprocessing system set up by the University Corporation for Atmospheric Research (UCAR) Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC) group (e.g., Rocken et al. 2000). Each profile consists of 3000–5000 samples of refractivity (as a function of MSL altitude), located between the near surface and approximately 40 km. Such a vertical resolution is largely in excess of that of the meteorological model used in this study (28 levels), and also of the estimated resolution of a GPSRO profile [200–500 m, or 100–200 vertical levels, e.g., Kursinski et al. (2000)]. The profiles are thus vertically thinned by a factor of 25 (keeping only one vertical sample out of every 25, to a density of about one datum every 200 m). This is still in excess of the vertical resolution of the model, but is comparable to the resolution of the profile. The NRT data that are expected from several institutions [UCAR, GeoForschungsZentrum (GFZ), and the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT)], of which first examples are beginning to be made available, will be delivered with a sparser sampling (of the order of the estimated resolution), and this thinning will no longer be necessary.

The observed quantity is the refractivity at a given height and varies by about two orders of magnitude over a vertical profile of the atmosphere, from about 4 \( N \) units at the model top (10 hPa) to 300–400 \( N \) units at sea level. To compare \( O - F \) uniformly at different heights, these are normalized to \( (O - F)/F \), which is of the same order of magnitude throughout the entire modeled atmosphere. This quantity is usually small. As shown in the global comparison in Fig. 1, the standard deviation of \( (O - F)/F \) is 0.01–0.02 everywhere in the modeled atmosphere (surface–10 hPa). The \( (O - F)/F \) bias is within ±0.005, and its two major features are small offsets in the lower troposphere and near the tropopause. The completeness of observations at a given height is also displayed in Fig. 1. Many profiles do not reach the low troposphere, because it becomes increasingly difficult for the receiver to continue tracking the signal there. The statistics in the low troposphere will thus present a bias resulting from this selection, in addition to other biases that may also be present.

c. Formulation of refractivity

The atmospheric refractivity is a thermodynamic variable. It depends on properties of the air, such as density and temperature. Among the several expressions available for refractivity, the draft recommendations from the International Association of Geodesy (Rueger 2002) are used in this study:

\[
N = 77.6890 \frac{p_d}{T} + 71.2952 \frac{p_w}{T} + 375463 \frac{p_w}{T^2},
\]

where \( p_d \) is the partial pressure of dry air, \( p_w \) is the partial pressure of water vapor, and \( T \) is the absolute temperature. The units of pressure and temperature are hectopascal and Kelvin, respectively. The refractivity is adimensional. Although the effect of the precise atmospheric composition is small, the coefficients of the above expression account for the current composition of dry air, and especially for the concentration of CO\(_2\), assumed for this expression to be 375 ppm.

d. The observation operator

The atmospheric state is described with the following control variables: surface pressure, surface skin temperature, and vertical profiles of temperature, moisture, and wind. The refractivity \( N_{\text{model}} \) at any given point needs to be estimated from one such description of the atmospheric state. The model assumes that the
atmosphere is in local hydrostatic equilibrium. A local quantity, such as the refractivity, then depends on the control fields only through the local value of each of these fields.

As a thermodynamic quantity, \( N_{\text{model}} \) depends only on a subset of the control variables—the surface pressure and the profiles of temperature and moisture. The detailed procedure to compute \( N_{\text{model}} \) from 6-h forecasts is as follows:

1) The observed data are obtained as an array of pairs \((h, N_{\text{Obs}})\). For each geometric height \( h \), the corresponding geopotential altitude \( z \) is then evaluated, which is the actual altitude model coordinate. The geopotential altitude (e.g., Holton 2004) represents the gravitational potential energy with respect to the MSL, expressed as a length, that is, divided by a standard value of acceleration. Inasmuch as the acceleration of gravity is not a global constant, but depends on latitude and height (local anomalies are much smaller and will be ignored), the geometric height and the geopotential altitude differ. The geopotential is calculated following the specifications of the World Geodetic System 1984 (WGS-84) World Geodetic Model (NIMA 2004). These account for a geodetic model of the refractivity \( h \), the location of a given observation, which is known in either altitude \( h \) or geopotential height \( z \), can be compared against the ensemble of geopotential heights of the model levels \( z_{\text{model}}(\eta) \). This allows the interpolation of the refractivity from the values evaluated at each level. Because of the behavior of the refractivity, the interpolation is chosen to be linear in \( \log(N) \). We thus have an ensemble \((z, N_{\text{Obs}}, N_{\text{model}})\).

2) At the latitude and longitude of the profile, the model provides a surface pressure \( p_s \), and values of temperature \( T \) and moisture \( q \) at each level. We also know the surface model topography, expressed as the surface geopotential altitude \( z_{\text{MT}} \).

3) The background value \( N_{\text{model}} \) of the refractivity can be evaluated at each vertical level \( \eta \); the local pressure at each level of the vertical coordinate is given by

\[
p = p_T + \eta(p_S - p_T),
\]

where \( p_T \) is the model-top pressure, in this case 10 hPa. Given the pressure, temperature, and moisture, standard formulas allow for the evaluation of the partial pressures of dry air \( p_d \) and water vapor \( p_v \), and hence the refractivity \( N_{\text{model}} \) at each model level.

4) The geopotential altitude \( z_{\text{model}} \) of each model level can be evaluated through the hydrostatic equation.

The surface has a geopotential altitude \( z_{\text{MT}} \), and a discrete integration upward starting there can be performed. The integrand depends on virtual temperature, and the discretization accounts for the gradient of virtual temperature between model levels. The geopotential spacing between two model levels is

\[
\Delta z = \frac{-R}{g_0} \left[ T_v \Delta \chi - \frac{1}{2} \frac{dT_v}{d\chi} (\Delta \chi)^2 \right],
\]

where \( \chi = \log(p) \).

5) To account for the presence of vertical correlation in the observation errors, the contribution \( J_O \) of an observation datum to the cost function, which is normally

\[
J_O = \frac{1}{2} \left( \frac{O - F}{E} \right)^2,
\]

where \( O, F, \) and \( E \) are the observation, its forecast estimate, and the assumed observation error, respectively, is downweighted in order to match the density of observations to the estimated vertical resolution. The latter, which is the vertical error correlation length of the data, is about 200–500 m and of the order of the Fresnel diameter (Kursinski et al. 2000). Our own experience shows that over-weighting the data often leads to degraded performance. There may be unmodeled sources of bias or error correlation that make the data not statistically independent. To avoid over-weighting until practical estimates of the vertical error correlation are available, in this work we take the conservative estimate of 1 km for the vertical correlation length, thus both assuming that the error correlation is slightly larger than the theoretical estimate and slightly under-weighting the GPSRO data. For every single datum, if there are \( N_D \) data in the same profile within less than 1 km vertically, the contribution is reduced to

\[
J_O = \frac{1}{2N_D} \left( \frac{O - F}{E} \right)^2.
\]

With the abovementioned thinning, \( N_D \) typically has a value of 5–10.
e. Observation errors and quality control

Theoretical estimations suggest (e.g., Kursinski et al. 2000, and references therein) that the accuracy of radio occultation data is of the order of 0.2%–0.5% in the upper troposphere and lower stratosphere, degrading to 1%–2% above and below. In the upper region, the degradation is mostly due to contamination of the signal by the ionosphere; in the lower region, the degradation is dominated by the irregular distribution of water vapor. However, given a mismatch between the observations and model, it is difficult to attribute it to either of the above or the observation operators. As a first approach, a statistical evaluation was made based on the difference between the observed and modeled values \((O - F)\), using the entire year of 2004 of CHAMP data (Fig. 1). The model fields are the 6-h forecasts started from the previous analyses, which are the background fields used in the assimilation. The standard deviation of \((O - F)/F\) is always on the order of 1%–2%. The main feature of the results is the presence of an important skewness and excess kurtosis in the distribution of \((O - F)/F\). Although the distribution has a large central bulge, its wings do not tend to zero as quickly as would be expected for a Gaussian distribution (especially in the lowest 5 km). In the assimilation process, it is assumed that the data distribution around the forecast is a centered Gaussian. This is equivalent to assuming that the contribution of a datum to the cost function must be quadratic. The presence of significant bias, skewness, or kurtosis invalidates this. For the case under study, the bias is found to be small, at or below 0.5%. The skewness and especially the kurtosis are more important, and a filter is put in place to limit the data to a subset whose distribution is closer to Gaussian. After some experimentation, it is found that a filter, rejecting data whose \((O - F)/F\) was larger than 5% [i.e., accepting \(-0.05 \leq (O - F)/(F) \leq 0.05\)], eliminated the problem of skewness and kurtosis. The filter eliminates about 1% of the observations. Nearly all of these are in the lowest 4 km, where the filter rejects 5%–10% of the data. The distribution of the remaining data is very close in mean and standard deviation to the unfiltered distribution. The cases with large \((O - F)/F\) are correlated with cases with large moisture and relative humidity. This suggests that many rejected data are not necessarily bad, but result from special meteorological situations where the \((O - F)/F\) distribution is intrinsically different from the Gaussian that describes the rest of the data. Further work is needed to describe their true distribution (out of the scope of this paper) in order to properly evaluate their contribution to the cost function.

Compared to the abovementioned estimates of the accuracy of observations, an important fraction of the observed \((O - F)/F\) seems to be due to the model, especially in the upper troposphere and lower stratosphere. Whatever the source, a conservative approach was taken by setting the observation error equal to 1.5% (i.e., \(\sigma_{\text{Obs}} = 0.015F\)). In most of the atmosphere this is an overestimation of both the observed \((O - F)\) difference and the theoretically estimated accuracy. This choice was made to avoid giving a too-large weight to the radio occultations. Also, as observed in Fig. 1, the bias is small and can be ignored, provided that it is a small fraction of the observation error. The bias cannot be ignored if the assumed observation error is smaller than, or comparable to, the bias. This would make the problem substantially more complex, because the bias changes with latitude, altitude, and season, and probably also with other parameters. Our experimentation with smaller assumed error, comparable to theoretical estimates (Kursinski et al. 2000), showed that the larger weight to the data leads to forecast degradation. This suggests that the fine tuning of the assimilation of GPSRO data is a complex issue, and in general we have taken conservative estimates of these and other parameters to avoid overweighting the data, even if this leads to suboptimal results in the data assimilation. Once the impact is shown to be positive, further fine tuning is desirable to reduce the assumed error and maximize the impact of the data. This is considered out of the scope of this work.

Since the initial GPS/MET, profiles (Ao et al. 2003, and references therein) in the lower troposphere have been found to often contain a substantial negative bias in \((O - F)\) (2%–3%). This makes them inappropriate for either assimilation or forecast evaluation. This bias has been reduced substantially in recent years (as a result of improved postprocessing algorithms), and the statistical analysis of \((O - F)/F\) shows that the data rejection filter (section 2e) is sufficiently well behaved and removes most of the remaining bias. The lower troposphere still shows some bias, but only of the order of 0.5% (Fig. 1). There are also systematic biases near the tropopause, which is a feature also found in other models (Rocken et al. 1997; Wickert et al. 2001b) and likely caused by the difficulty in representing the sharp change in temperature gradient at the tropopause.

The vertical GPSRO data range is normally from near the surface to approximately 40 km. In the data, the upper layers are expected to be contaminated (Kursinski et al. 2000) by small residual effects of the ionosphere, which cannot be fully subtracted (even with a dual-frequency receiver). Refractivity data will also be contaminated by the climate approximation...
used as upper boundary condition during GPSRO data processing. It is thus advisable, in general, to ignore or downweight the upper part of the profile. Given that the model top in this work is at 10 hPa, this already happens automatically, because a cutoff ignores all data above the model lid (approximately 30 km).

f. Description of the experiments

Two experiments covering different time periods are considered here (one for summer and one for winter conditions). The first set covers the period from 1 January 2004 to 15 February 2004 (hereafter the winter experiment), whereas the second set covers from 1 June 2004 to 15 July 2004 (hereafter the summer experiment). Each required the evaluation of two assimilation cycles—a reference with all normal operational data, and a GPSRO cycle, where CHAMP refractivity data are added. A number of statistics are then computed, comparing the relative performance of the two cycles in each experiment. Within the normal procedure of assimilation, 6-h forecasts (also referred to as background fields) are sequentially produced. Longer-range forecasts (up to 6 days) are also produced, starting from the analyzed fields at 0000 and 1200 UTC each day, from 15 January 2004 to 9 February 2004 (winter period) and from 15 June 2004 to 9 July 2004 (summer period). In each experiment, the first 2 weeks are discarded as a transient period after adding the GPSRO data. Forecasts are not launched for the last 6 days, because their validity date would be beyond the end of the analyzed period.

3. Evaluation of the forecast performance

For each of the two experiments (winter and summer), it is possible to compare the relative forecast performance of the reference cycles (hereafter Ref) and the ones assimilating GPSRO data (hereafter GPS). Statistical comparisons are made of the comparative behavior of the differences between cycles of \((O - F)/F\) values. The observation data that are used here for the evaluation are the GPSRO profiles, radiosonde data (RS), and satellite radiance data (NOAA AMSU-A).

a. GPSRO evaluation of 6-h forecasts

The GPSRO data are not only useful for assimilation. Their global homogeneous distribution and good vertical resolution make them ideal to evaluate the comparative skill of different experiments. To study the effects on the assimilation cycle of the adjustment to the GPSRO data, statistical analyses of \((O - F)/F\) are produced for each day of the experiment. The modeled portion of the atmosphere (surface to 10 hPa, or approximately 0-30 km height) is here binned into vertical layers of 500 m each. The ratio of the standard deviation of \((O - F)/F\) in the two experiments \((\text{STD}_{\text{GPS}}/\text{STD}_{\text{Ref}})\) for each bin, a measure of the comparative forecasting skill, is shown in Fig. 2 as a function of time. A ratio of one means equal forecasting skill, whereas values smaller than one imply that the GPS experiment has better skill. Lower-altitude bins show little difference between the two experiments—the reference and GPSRO. However, for the higher-altitude bins (in the stratosphere) the experiments assimilating GPSRO show a progressive and systematic improvement over time. This is also shown in Fig. 3 for several height layers. The behavior of the experiments assimilating GPSRO for the higher-altitude bins becomes systematically better than the reference experiments in less than a week. The transient period is recognizable up to about 2 weeks. After that point, the standard deviation statistics, which are averaged over the globe, are about 10% better than for the reference experiment in the entire stratosphere.

The bias is also substantially reduced in most of the atmosphere, but particularly in the stratosphere (also shown in Fig. 3). The fact that the bias reduces considerably in the upper layers of the model upon the addition of a comparatively small number of observations (about 50 profiles per assimilation step, and each one containing about 30 independent observations) suggests that the model is being only loosely constrained in those layers by other observations. Because this bias later reduces close to zero after a few days of assimilating GPSRO, the model is then constrained in those layers nearly only by GPSRO. Radio occultation observations are thus closing a gap of data coverage from other data types.

To explore the geographical distribution of this impact, the atmosphere is further binned into height by latitude boxes whose size is 500 m \(\times\) 10°. Statistics of \((O - F)/F\) are evaluated over 1 month of data. Because the statistics of the temporal series suggest that the first 15 days are a transient period of adjustment, the averages are done, respectively, from 16 January to 15 February 2004 for the winter experiments, and from 16 June to 15 July 2004 for the summer experiments. Figure 4 shows the ratio \(\text{STD}_{\text{GPS}}/\text{STD}_{\text{Ref}}\) for each of these bins. The largest improvement [reduction of \((O - F)/F\) standard deviation] takes place in the southern stratosphere (midlatitude and polar), and is most significant during the winter experiment (austral summer). It is also significant in the northern polar stratosphere. The effect appears as an improvement of up to 30% in the 6-h forecast skill, as measured with GPSRO data. Elsewhere, there is a generalized but small improvement.
Only a few boxes show deterioration: these are sparse and have values of a few percent.

b. AMSU evaluation of 6-h forecasts

The evaluation with GPSRO data (section 3a) suggests that there is an improvement in the short-term forecast skill, mainly in the stratosphere, and especially in the Southern Hemisphere and near the North Pole. Some channels of the AMSU radiometers are sensitive to the properties of the same atmospheric regions where the above evaluation shows a significant impact. In particular, the weighting functions of channels 9 and 10 from AMSU-A are sensitive to a broad region around 20 km (weighting function peak around 50–100 hPa). AMSU-A data effectively provide temperature profiles of the atmosphere, but the broad weighting functions imply that only a much lower vertical resolution than that of the GPSRO is obtained.

The potential for impact of GPSRO on the assimilation is seen from the favorable ratio of observations to control variables. The GPSRO observation operator depends on 57 variables: the temperature and specific humidity at each of the 28 vertical levels and the surface pressure, whereas we estimated that there are (conservatively accounting for vertical correlation) approximately 30 independent observations in a GPSRO profile. In contrast, the number of assimilated radiance channels over oceans in EC’s system is eight from an AMSU-A instrument (channels 3–10) and four from an AMSU-B instrument (channels 2–5). The improved profile after GPSRO assimilation should also improve the assimilation of collocated AMSU observations. This could particularly happen for channels with sensitivity above the model top (i.e., AMSU-A channels 9 and 10). Because a fraction of their weighting function is above the model top, an extrapolation is necessary (Matricardi 2003). This extrapolation is likely more accurate if GPSRO data are available. It must be mentioned that radiance data have been calibrated against

Fig. 2. Global evolution of \((O-F)/F\) for GPSRO data \((O)\) in the winter experiments. Data are grouped in height layers of 5 km.
the model, subtracting the relative bias. It is not unlikely that, especially in data-scarce regions, GPSRO will modify the model sufficiently to require a recalibration of radiance data. In this study, however, the standard calibration has not been modified.

A performance evaluation through $O - F$ for some of the upper AMSU observations is thus carried out in this study, and particular attention is paid to the brightness temperature of AMSU-A channels 9 and 10. The evaluation with GPSRO data shows a very different impact in the Northern and Southern Hemispheres. It is thus interesting to find out whether a similar effect appears with AMSU data. Hemispheric daily averages of $(O - F)$ are displayed in Fig. 5 for AMSU channels. There is a significant improvement in both bias and standard deviation in the Southern Hemisphere as a result of assimilating GPSRO. In the Northern Hemisphere the results are positive in the summer experiment, but there is an important degradation in the winter experiment. The amount of AMSU data assimilated is also smaller than that in the reference. Some data had been rejected, which suggests that a recalibra-

![Fig. 3. Global evolution of the change of $(O - F)/F$ for GPSRO data $(O)$, upon assimilation of GPSRO, as a function of time and altitude. Data are analyzed for each day along the assimilation experiments, binned in height layers of 500 m. The two upper plots are for the winter experiment, and the two lower ones for the summer experiment. There are no data for 28 Jan. The grayscale coding shows the ratio between STD in the GPSRO vs reference experiments. For emphasis, areas with improvements ($\text{STD}_{\text{GPS}}/\text{STD}_{\text{Ref}} < 1$) are shown in the left-hand-side plots, whereas the two-right-hand side panels (notice the inversion of the grayscale) show those with degradations ($\text{STD}_{\text{GPS}}/\text{STD}_{\text{Ref}} > 1$).]
tion was necessary. In Fig. 6, the \((O - F)\) standard deviation ratios are shown as a function of latitude and channel. Channels 3–10 are sensitive to the temperature of layers at different pressure ranges. The channel number is an approximate indicator of the altitude (channel number increasing with altitude). The plots that are shown are thus analogous to those shown in Fig. 4. Some features from Fig. 4 can be identified here, notably an improvement in the winter experiment in the stratosphere of the Southern Hemisphere and the North Pole. Other features are also present, but the results are sometimes significantly negative in areas where the GPSRO evaluation indicates neutral or positive results.

c. Radiosonde evaluation of 6-h forecasts

As noted above, the distribution of radiosondes is geographically very uneven, and it should be expected that the impact of the added radio occultation data is concentrated in regions with sparse sampling by radiosondes. The globally averaged result for the skill of the 6-h forecast of temperature (see Fig. 7) is a small but positive improvement (0.5%–1%) in both summer and winter experiments. These statistics are likely to underestimate the global impact, because the areas of the earth most affected by GPSRO assimilation have a low density of radiosondes. Evaluation by latitude band is thus performed in this study and the standard deviation
of the 6-h ($O - F$) is evaluated for the temperature and dewpoint depression of the reference and GPSRO experiments. The results are shown in Figs. 8 (temperature) and 9 (dewpoint depression). The plots follow the structure and grayscale coding of Fig. 4. The pixels are darkened according to the impact in the ratio $\text{STD}_{\text{GPS}}/\text{STD}_{\text{Ref}}$. Two plots are presented for each experiment. In the left-hand side plots of Figs. 8 and 9, only pixels with improvement (smaller standard deviation in the GPS experiment) are darkened whereas in the right-hand side plots in Figs. 8 and 9 only the pixels with degradation (bigger standard deviation) are darkened (except "missing data" pixels, such as the entire stratosphere in Fig. 9). In both figures, all left-hand-side plots are darker, which means that the results are predominantly positive and, as could have been expected a priori, the Southern Hemisphere is particularly improved. Figure 8 shows that the temperature fields are improved throughout the entire Southern Hemisphere and in the northern polar areas. The impact is neutral in northern midlatitudes, where most radiosonde observations are concentrated, and is slightly negative near the equator. Figure 9 illustrates moisture fields in the form of dewpoint depression. In general, an improvement in the lower troposphere (at or below 700 hPa) is observed. This can be seen at progressively higher altitudes with increasing southern latitudes. A mixed area appears, however, at the upper troposphere in the southern tropical/midlatitude region in the January 2004 experiment. The areas where it is most important to improve the results, however, are the low troposphere and especially the Tropics and midlatitudes, because these contain the largest amounts of moisture. There, the impact is in general positive.

d. Longer-range forecast

The evaluations presented in the previous sections are made against the short-term forecasts that are computed as part of the assimilation process itself. These
forecasts have a range of 6 h. The improvements found after assimilation of GPSRO data suggest that longer-range forecasts may also show a positive impact. In this section, a study is done of this impact for forecasts up to a range of 144 h, or 6 days. Following impact results presented in former sections, a significant signature is expected to appear in the thermodynamic variables.

Two sets of evaluations are performed that are based on 1) the evaluation statistics of the forecasts versus the analyses and 2) the statistics of the forecasts versus the radiosonde data. As in former sections, the grand averages exclude the first two weeks of each experiment. The anomaly correlation of the temperature field is shown in Fig. 10 (winter experiment) and Fig. 11 (summer experiment). In general, the results are positive, and confirm that the impact found in the 6-h forecasts propagates in time to longer ranges. Both hemispheres

---

**Fig. 6.** Change, by latitude band (10° binning), in the statistics of \((O - F)\) for AMSU observations in the GPS experiment vs the reference. The first 2 weeks are ignored as a transient period. The panel distribution and grayscale coding follow that of Fig. 3. Notice some missing data: channels 2 and 11, which are not assimilated, as well as some latitude bands for certain other channels, due to constraints of topography or surface type.
appear to benefit from the inclusion of the GPSRO data. The Northern Hemisphere (not shown) presents a small but positive impact at low altitudes and a gain of the order of 3 ha/td a y for 100 and 50 hPa. The results for the Southern Hemisphere present a larger impact, which can be noticed at all altitudes. Comparisons were also made versus the radiosonde data. The geographical and altitude distribution of the regions that present an impact (not shown) is, for the longer-range forecasts, similar to that of the 6-h forecasts.

4. Discussion and conclusions

The comparative skill of EC’s operational system after the addition of GPSRO refractivity data is tested. The evaluation includes GPSRO data, which are in a global, vertically resolved, and well-distributed database, and thus useful for evaluation purposes. Compared to the reference cycle, the cycle where GPSRO is assimilated has an initial transient period where statistical scores improve progressively. This transient state is followed by an improved stationary situation, which suggests that the GPSRO data are having a beneficial effect. The length of the transient period, an indicator of the persistence of the information, is about 2 weeks. After the transient period, the standard deviation of (O - F)/F of the refractivity in the stratosphere is about 90% of the value without assimilation of GPSRO data. The analysis of the geographical distribution of this impact shows that the improvement is concentrated in the stratosphere of the Southern Hemisphere and the northern polar areas, where the skill improves by up to 30%. In such dry conditions, the refractivity is essentially a measure of the air density, meaning that the forecast accuracy of the air density has improved by 10%-30% in the entire stratosphere.

Similar analyses are performed with other relevant observations, and particularly AMSU-A brightness temperatures and radiosonde measurements of temperature and dewpoint depression. AMSU-A channels 3–10 are sensitive to broad layers of the atmosphere, which are consecutive in altitude. In particular, channels 9 and 10 are sensitive to regions around 20–30 km, where the GPSRO evaluation indicates a substantial impact, especially in the Southern Hemisphere. The geographical distribution of the difference between the reference and GPSRO experiments, evaluated with AMSU-A data, is analyzed, and shows some features in common with the distribution of impacts as measured with GPSRO. In particular, the impact in the southern stratosphere, where improvements of 30% in GPSRO are observed, is also noticeable. Some other areas, however, show a slight degradation. Based on AMSU-A observations, the winter experiment shows predominantly positive results, whereas the summer experiment shows mixed positive and negative results.

A similar evaluation (by latitude band) of the geographical distribution of impact, as measured with radiosonde data, is also performed. The observations analyzed to study the effect of the assimilation of GPSRO data on temperature and moisture fields are temperature and dewpoint depression. The radiosondes are distributed very unevenly on the earth, and a global mean shows little impact, although positive. This is not surprising because the knowledge of the atmospheric state is already good in most of the Northern Hemisphere, where most of the radiosonde launch sites are located. The analysis by latitude band reveals instead many features in common with the GPSRO evaluation, and suggests also that assimilating GPSRO leads to improvements in both the temperature and the moisture fields, especially in the Southern Hemisphere and northern polar areas. The analysis of radiosonde temperature is generally positive. That of dewpoint is often positive, although presents mixed results at tropopause altitudes. The accuracy of radiosonde moisture, however, is known to be poor there (Wang et al. 2002; Turner et al. 2003). The moisture content in the mid- and low troposphere, where water vapor is concentrated, is improved in nearly all cases. A significant...
exception in the generally positive radiosonde evaluation is the equatorial band, which concentrates instances of negative, albeit moderate, impact on both variables. This is consistent with both known weaknesses of the GPSRO refractivity data (i.e., negative bias in the low equatorial troposphere) and the difficulty in modeling this region. Further work is needed to filter or correct GPSRO observations in the low equatorial troposphere.

Assimilation of refractivity data in EC's operational system should lead to predominantly positive results. Although it is difficult to extend this conclusion to other systems, the positive impact in the Southern Hemisphere and in the upper layers, where other data are scarce, suggest that GPSRO is filling data gaps from other data types, which should apply to a broad class of assimilation systems.

The choice of refractivity instead of the theoretically superior bending angle is made because initial analysis suggested that with the values of horizontal resolution

---

**Fig. 8.** Change, by latitude band ($10^\circ$ binning) in the ($O - F$) standard deviation, for RS temperature observations in the GPS experiment vs the reference. The panel distribution and grayscale coding follow that of Fig. 3.
and, most importantly, the low model top in this study (10 hPa), the bending angle data do not present an advantage, and could lead to worse results. However, it is found that refractivity data are still of sufficient quality to allow a positive effect. This is clearly demonstrated in this study using GPSRO as evaluation data, and also with radiosonde data. Satellite radiometer data, though, are less conclusive. In addition to the confirmation of some of the positive results also indicated by GPSRO and the radiosondes, some significant negative results are present. This is exemplified by channels 9 and 10, which confirm the improvement in the southern stratosphere during the winter experiment, but presents problems in the summer experiment. These problems are deeper than a small degradation, and may be related to a bias conflict. It is likely that the impact of GPSRO is sufficient to require a recalibration of the bias of at least some AMSU channels, especially those sounding the stratosphere.

For all three data types, the equatorial region ap-

Fig. 9. Change, by latitude band (10° binning) in the \(O - F\) standard deviation, for RS dewpoint depression observations, in the GPS experiment vs the reference. The panel distribution and grayscale coding follow that of Fig. 3.
pears as the one where less benefit is obtained compared to a priori expectations. In addition to the larger horizontal gradients that are normally present, some systematic negative bias remains in the equatorial troposphere. The data filter could be adjusted for this and is part of ongoing research. On the other hand, some data now rejected should probably be included with a more precise evaluation of their statistics, yet to be determined.

A longer-range forecast evaluation, of up to 6 days, is
performed. The results are, in general, positive and the quality of the forecast is improved. The distribution of the impact is geographically similar to that found in the short-term comparisons. It is particularly encouraging that this takes place with a moderate amount of radio occultation data, which suggests that the addition of larger amounts of data, such as COSMIC in the immediate future, may have a larger positive impact.

FIG. 11. Anomaly correlation (each experiment’s forecasts against its own analysis) of the temperature field, as a function of the forecast range, for the Southern Hemisphere in the summer experiment. The reference experiment is represented with the dark line (SummerRef), whereas the GPSRO experiment is the light gray line (SummerGPS).
Acknowledgments. The authors thank GFZ’s CHAMP and UCAR’s COSMIC teams for making their GPSRO data available. We also thank two anonymous referees for their comments.

REFERENCES


