Estimation of the Added Value of the Absolute Calibration of GPS Radio Occultation Data for Numerical Weather Prediction

JOSEP M. APARICIO AND STÉPHANE LAROCHE

Data Assimilation and Satellite Meteorology, Environment Canada, Dorval, Quebec, Canada

(Manuscript received 30 April 2014, in final form 3 October 2014)

ABSTRACT

An analysis of the impact of GPS radio occultation observations on Environment Canada’s global deterministic weather prediction system is presented. Radio occultation data, as any other source of weather observations, have a direct impact on the analyses. Since they are assimilated assuming that they are well calibrated, they also impact the bias correction scheme employed for other data, such as satellite radiances. The authors estimate the relative impact of occultation data obtained from, first, their assimilation as atmospheric measurements and, second, their influence on the bias correction for radiance data. This assessment is performed using several implementations of the thermodynamic relationships involved, and also allowing or blocking this influence to the radiance bias correction scheme.

The current implementation of occultation operators at Environment Canada is presented, collecting upgrades that have been detailed elsewhere, such as the equation of state of air and the expression of refractivity. The performance of the system with and without assimilation of occultations is reviewed under conditions representative of current operations. Several denial runs are prepared, withdrawing only the occultation data from the assimilation, but keeping their influence on the radiance bias correction, or assimilating occultations but denying their impact on the bias correction procedure, and a complete denial.

It is shown that the impact of occultations on the analysis is significant through both paths—assimilation and radiance bias correction—albeit the first is larger. The authors conclude that the traceability link of the ensemble of occultations has an added value, beyond the value of each datum as an atmospheric measurement.

1. Introduction

GPS radio occultation (GPSRO) data have been a successful addition to the sources of observational data for numerical weather prediction (NWP). Positive results were confirmed at various meteorological centers (e.g., Healy and Thépaut 2006; Cucurull et al. 2007; Aparicio and Deblonde 2008; Rennie 2008). GPSRO probes with high vertical resolution regions that are otherwise poorly sampled, or where most data are of low vertical resolution, most notably over the oceans and polar areas. This qualitative diversification of sampling and resolution yields overall positive results through a better description of the global circulation. The improvement is less obvious in regions that were already well sampled, as over the continental Northern Hemisphere.

The potential added value of these data probably goes beyond this improvement in coverage and vertical resolution. The long-term stability of the data and their traceability to fundamental standards are recognized assets (Anthes et al. 2000). Exploiting this property may have a profound influence on an assimilation system, but also poses challenges beyond those of merely assimilating the data.

Several other sources of data, notably satellite radiances, require a bias correction procedure before being ready for assimilation. This bias correction requires reference information that must be provided externally, and normally consists of both the physical processes that determine the climatology of the model and data that do not require calibration. Because of their traceability, radio occultation data are good candidates to participate in this procedure, in addition to measurements by radiosondes, aircraft, and certain surface instruments. The amount of reference calibrating data is only a small fraction of
all available data, and their distribution is particularly uneven. Radio occultations have been a major addition to this subset, both due to their volume and their distribution.

This study describes a series of efforts toward the assessment of this potential as a calibrated source, and its exploitation, on top of already obtained good results derived from the diversification and densification of the sampling of the atmosphere. These efforts are part of a broader upgrade of the assimilation of GPSRO data at Environment Canada (EC). There are strong indications that GPSRO contributes significantly to the estimation of radiance bias corrections, when these are allowed to float and adjust, as shown, for instance, in Aparicio and Deblonde (2008), Healy (2008), Poli et al. (2010), and Cucurull et al. (2014). We will here estimate if this impact translates into measurable forecast value, in the form of better predictability.

The first implementation performed at EC was described in Aparicio and Deblonde (2008). The distribution of the impact indicates that a large fraction corresponds to the qualitatively different properties and distribution of these data, with respect to other assimilated data types. These data offer global coverage and vertically well-resolved profiles, although limited horizontal resolution. It is complementary to radiance data, which offers also global coverage and has good horizontal resolution, but limited vertical resolution. It is also complementary to the radiosonde network, which also provides data vertically well resolved, but its distribution is very irregular at a global scale. Fully exploiting the theoretical added value of an absolutely calibrated and traceable data source, on top of this diversified sampling, and a variety of other data types is challenging (Aparicio et al. 2008). Certain types of measurements assimilated in NWP systems are already considered to be well calibrated and traceable to fundamental standards, notably some measurements obtained from radiosondes, aircrafts and surface stations. Despite being assumed to be well calibrated, none is perfect, and some residual inaccuracy is to be expected. An effective radiance bias correction will depend on the compatibility between the different calibrations of these data types. Radiosondes, aircrafts, and surface stations ultimately depend on laboratory-calibrated conventional local measurements (temperature, moisture, and pressure). This is entirely different from GPSRO, which is nonlocal, remote, and provides a nonconventional measurement (refractivity). Since the traceability chain is very independent, this compatibility requires that each of these two classes of realizations be accordingly accurate.

We describe in this article the modifications made to our first implementation. These are oriented to improve the quality of the realization of the traceability link for GPSRO data, and from it, to provide a larger impact from their assimilation. Portions of these developments have been described in Aparicio et al. (2009), and Aparicio and Laroche (2011, hereafter AL11). This article provides additional details of the present approach at EC, after these modifications are collected into a coherent implementation. Their impact in the performance of EC’s global NWP system is then studied.

The spread of the impact of the same data, assimilated under different implementation choices is estimated. Since EC’s operational system has evolved significantly since GPSRO was introduced, a supplementary objective of the present study is to provide an updated estimation, through denial experiments, of the net impact of GPSRO data under the present operational conditions at EC. Here, we examine the relative impact of GPSRO data through both its assimilation, and its impact in the estimation of radiance bias corrections. We explore partial denials of GPSRO, first in the data assimilation only, while keeping the bias corrections of the control run; and, second, in the bias correction procedure, but keeping the assimilation, and also a complete denial, where GPSRO data have been entirely eliminated from the system.

2. Present implementation of GPSRO assimilation at EC

a. The GPSRO data

Radio occultation observations provide information about the field of index of refraction. This information may be provided in several forms, including the bending angle as a function of the impact parameter, the refractivity as a function of altitude, or others. In all cases, a user must be able to describe and manipulate the scalar field of refractivity $N(x)$, which includes the representations of the position $x$ and the refractivity $N$. The current implementation of GPSRO operators at EC can handle refractivity and bending angle profiles, with refractivity being the operational configuration. For the present purpose, the difference is not relevant, and the implementation of bending angle will be described elsewhere. In the following, we consider the $N(x)$ expression to generically describe any technical representation of refractive observables within a forward operator, regardless of whether refractivity is the final goal, or further transformations, to bending angle or others, are required.

b. General concepts

Let us assume a parcel of moist air of pressure $P$, density $\rho$, absolute temperature $T$, and specific moisture $q$ (mass of water per unit mass of moist air). We also use
The expression \( t \) for the Celsius temperature and \( x \) for the molar fraction of water vapor. We assume that air consists of two fractions: the dry air of molar mass \( m_d \) and the water vapor of molar mass \( m_w \). When necessary, we will use the values \( m_d = 28.965 \, 516 \) and \( m_w = 18.015 \, 254 \) (both in g mol\(^{-1}\)) (AL11). The total density is the sum of the densities of the dry air and water vapor fractions \( \rho = \rho_d + \rho_w \).

Therefore, \( t = T - 273.15 \, K \) and

\[
x = \frac{m_d q}{m_w + (m_d - m_w)q}.
\]

(1)

If air were exactly an ideal gas, its equation of state would be expressed as

\[
P_{id} = (\rho_d R_d + \rho_w R_w) T = (\rho R_d T_v^{id}),
\]

(2)

where \( R_d \) and \( R_w \) are the gas constants for pure dry air and pure water vapor, respectively, associated to their molar mass. Equation (2) defines the virtual temperature \( T_v^{id} \) of the moist air, if assumed ideal, which can be expressed as

\[
T_v^{id} = T \left(1 + q \frac{m_d - m_w}{m_w}\right).
\]

(3)

Moist air is close to ideal, but not exactly, so we express its equation of state as \( P = \rho R_d T_v \), where \( Z \) is the compressibility factor. For this study, following Aparicio et al. (2009) and AL11, we have adopted the compressibility factor given by Picard et al. (2008), which is applicable to atmospheric conditions:

\[
Z = 1 - \frac{P}{T} \left[ a_0 + a_1 t + a_2 t^2 + (b_0 + b_1 t)x + (c_0 + c_1 t)x^2 \right] + \frac{p^2}{T^2} (d + ex^2).
\]

(4)

The constants are listed in Table 1. We therefore express, instead of (2), the nonideal equation of state as

\[
P = \rho R_d T_v,
\]

where we have introduced a nonideal virtual temperature:

\[
T_v = T \left(1 + q \frac{m_d - m_w}{m_w}\right) Z.
\]

(6)

The quantity \( T_v \) absorbs both the moisture and the compressibility factor into an equation of state (5) for air. From the point of view of the hydrostatic behavior of a parcel of air, it can be then treated as if it were dry and ideal with temperature \( T_v \).

c. The position operator

In a generic NWP system, a vertical profile of the atmosphere at a specific location is described by some local state vector, containing several quantities as a function of some vertical coordinate. In the EC’s NWP model, the vertical coordinate is a hybrid quantity, strongly linked to pressure, and altitude is a derived quantity. In a generic case, both altitude and pressure may be functions of the vertical coordinate and other state variables. In all cases, an operator linking altitude and pressure is required. This section presents the altitude operator, regardless of the direction in which it must be applied. We assume hydrostatic equilibrium:

\[
\frac{dP}{dh} = -g \rho,
\]

(7)

where \( h \) is the geometric altitude and \( g \) is the acceleration of gravity. Integrating this equation requires a boundary condition, which a background model normally provides as the surface pressure \( P_s \) at the altitude of the model topography \( h_T \). It also requires some closure through the equation of state, as in (5).

Another element that is required is the quantity \( g \). It includes both gravitation and the rotation of the non-inertial Earth-fixed reference frame. The value of \( g \) is not a constant, and varies with altitude \( h \), and with latitude \( \phi \), both due to the distance to Earth’s center, and to Earth’s rotation. Smaller variations are also found due to local anomalies. The latter are substantially smaller and, hence, neglected. Only the ellipsoidal altitude and latitude dependences are retained. We use the field proposed by the WGS-84 Earth ellipsoid model (National Imagery and Mapping Agency 2000). That report presents an expression for the acceleration at the surface \( g_s \), which is exact for a rotating ellipsoid in equilibrium:

\[
g_s = g e^{1 + k \sin^2 \phi} \left(1 - e^2 \sin^2 \phi\right)^{-1/2}.
\]

(8)
TABLE 2. Constants required for the description of the gravity field of the WGS-84 Earth ellipsoid model.

<table>
<thead>
<tr>
<th>Name</th>
<th>Value</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>6.378 137.0</td>
<td>m</td>
</tr>
<tr>
<td>$f$</td>
<td>1/298.257 223 563</td>
<td>—</td>
</tr>
<tr>
<td>$g_s$</td>
<td>9.780 325 335 9</td>
<td>m s$^{-2}$</td>
</tr>
<tr>
<td>$k$</td>
<td>0.001 931 852 652 41</td>
<td>—</td>
</tr>
<tr>
<td>$e^2$</td>
<td>6.694 379 990 14 × 10$^{-3}$</td>
<td>—</td>
</tr>
<tr>
<td>$m$</td>
<td>0.003 449 786 506 84</td>
<td>—</td>
</tr>
</tbody>
</table>

Constants $g_s$, $k$, and $e$ applicable to Earth, as well as its equatorial semimajor axis $a$ and flattening $f$, are defined by the WGS-84 specification, and summarized in Table 2. For nonzero but small altitudes, $h \ll a$, this specification recommends a second-order expansion, which we assume to be sufficiently accurate to be applied to the entire atmosphere, and is the expression that we retain for the gravity acceleration:

$$g = g_s \left[ 1 - 2 \frac{a}{a} (1 + f + m - 2f \sin^2 \phi)h + \frac{3}{a} h^2 \right]. \quad (9)$$

The hydrostatic equation with boundary condition is therefore expressed as

$$\frac{dP}{dh} = -g \frac{P}{R_d T_w}, \quad (10)$$

$$P(h_T) = P_s, \quad (11)$$

where $T_w$, $h_T$, and $P_s$ are obtained from the background field, and $g$ is from the WGS-84 specification.

d. The forward operator of refractivity

Several relationships between refractivity, pressure, temperature, and moisture have been proposed in the literature. Our former implementation (Aparicio and Deblonde 2008) used Rüeger (2002, hereafter R02). For this study, two other expressions, Smith and Weintraub (1953, hereafter SW53), and AL11 are also tested.

For completeness, we mention here the different expressions that have been tested for this work. These expressions use hectopascal (hPa) units for pressures, kelvin (K) for temperatures, kilogram per cubic meter (kg m$^{-3}$) for densities, and $N$-units for refractivity. For R02, which had been the initial operational expression at Environment Canada:

$$N = 77.6890 P/T + 71.2952 P_w/T + 3.75463 \times 10^5 P_w/T^2. \quad (12)$$

The SW53 expression is

$$N = 77.6 P/T + 3.73 \times 10^5 P_w/T^2. \quad (13)$$

Our current operational implementation uses AL11, whose expression is

$$N = N_0 (1 + 10^{-6} N_0/6) \quad (14)$$

with

$$N_0 = a_1 \rho_d + a_2 \rho_d \tau + a_3 \rho_w + a_4 \rho_w \tau, \quad (15)$$

where the expression’s parameters are collected in Table 3, and where $\tau = 273.15 K/T - 1$.

It is important to mention here that AL11 suggested that an expression of refractivity should be presented in terms of the partial densities of dry air and water vapor, in contrast to virtually all other available expressions, and in particular those of SW53 and R02, which use partial pressures while implicitly assuming that those partial pressures are well defined, additive, and proportional to the respective mole fractions. The reason provided in AL11 is the nonideal behavior of moist air. The partial densities of a mixture are well-defined quantities, independently of the ideality of the equation of state of the mixture, and are additive. By contrast, partial pressures are clearly defined, observable, and additive only if the mixture is an ideal gas. Air presents only small deviations from a perfect gas. However, by construction and use of the partial pressure of water vapor $P_w$, both SW53 and R02 implicitly assume that air is an ideal gas. For the tests performed here with SW53 and R02, we cannot avoid the evaluation of this quantity, which we choose as $P_w = x_w P$, where $x_w$ is the molar fraction of water vapor.

Equation (14) does not assume that the air is an ideal gas. Its partial densities are related to the total pressure and water vapor fraction consistently with the non-ideality applied elsewhere in this implementation, and in particular during the evaluation of the hydrostatic equation (4).

e. The estimation of observation error statistics

It is difficult to give a proper estimation of the a priori error distribution for GPSRO data. Values formally offered by data providers vary in their characteristics,
including only a declaration of intent, or imprecisely documented values. Some data are provided without an estimation of the error distribution, although the data appear to be valuable. As a result, we do not rely on estimates of error statistics offered by the providers.

However, these values are essential to determine the relative weights given to the background field and observational data. Therefore, a procedure was chosen, which dynamically determines a useful value for the a priori standard deviation of the observation error (Aparicio and Deblonde 2008).

In the specific case of GPSRO, data are naturally associated into vertical profiles. Profiles are measured and processed as an ensemble, and it is, therefore, to be expected that accuracy, precision, and resolution are collective properties of a profile, rather than individual properties of each datum, even if these properties may vary along the profile. We, therefore, estimate the a priori observation error distributions through entire profiles.

Let us assume a profile of observed values \( O_i \). If the observable is refractivity, this means a set \( \{ N_i(h_i); i = 1, \ldots, n \} \), but the procedure may be equally applied to profiles of bending angles \( \{ \alpha_i(a_i); i = 1, \ldots, n \} \), or other variables. The application of the forward model to a background field allows the evaluation of the corresponding background values \( B_i \) for each observation. We can then define the relative increment in observation space, and normalize them to allow a better comparison at different altitudes:

\[
Z_i = \frac{O_i - B_i}{B_i}. \tag{16}
\]

We also define the following weights, with a Gaussian shape as a function of altitude:

\[
w_{ij} = \exp\left(-\frac{(h_j - h_i)^2}{H^2}\right), \tag{17}\]

where \( H \) is some scale height, which we arbitrarily set to \( H = 5 \text{ km} \), and evaluate with them the weighted root-mean square (RMS) of the increments:

\[
z_i^2 = \frac{\sum_j w_{ij} Z_j^2}{\sum_j w_{ij}}. \tag{18}\]

This is an estimate of the average normalized difference between observations and the background field, at altitudes similar to that of each observation \( i \). For refractivity, typical values of \( z_i \) are of the order of 0.01, or 1% relative difference between data and background. Minimal values typically occur in the upper troposphere and lower stratosphere, around 0.005 or 0.5%, but often reach 3%–5% in the lower troposphere, especially if it is moist. Reversing the normalization, we obtain a list of values associated with each datum:

\[
e_i = z_i B_i. \tag{19}\]

We choose to use these moving vertical average RMSs against the background field \( e_i \), as a priori error estimates of the observations. For robustness, a lower bound of 0.002 is applied to \( z_i \) everywhere, before evaluating \( e_i \).

An example of the estimated dispersion of the error \( e_i \) is shown in Fig. 1, where the algorithm is applied to both the bending angle and refractivity data of the same profile. It is to be noted that, by construction, these values are not intrinsic properties of exclusively the input data, but also depend on the background field.
This estimate combines observation error (measurement or representativity), and background error. If the observation contribution is dominant, the procedure leads to weights for the data very close to optimum. Otherwise, the algorithm will attribute a lower than optimum weight, which is conservative.

However, this does not necessarily lead to significantly suboptimal performance. We are here concerned with the optimal use of GPSRO as a stream of data, rather than the optimal use of any individual datum. If the assimilation is suboptimal at any given moment, because the background error inflates the above-mentioned estimate of the data error statistics, the residual suboptimal improvement will at least reduce the background error, and thus the error estimate of future data, progressively increasing the weight of the GPSRO stream until an equilibrium is reached. This evolution toward an equilibrium can be seen through the $O - B$ statistics of GPSRO data, which are the basis for the above estimate. The RMS evolve to become clearly smaller in GPSRO-assimilating experiments than in GPSRO-denied ones (Aparicio and Deblonde 2008), until no further benefit is gained. We assume that at this equilibrium we will be close to the optimum weight for GPSRO data that the system is able to use with net benefit. Once there, tests to improve performance with several criteria based on fixed error statistics, failed to provide significantly better results. Given that the results were good, and not trivially outperformed, this algorithm was retained for operational use (Aparicio and Deblonde 2008).

With a set of fixed a priori errors, the restoring tendency that pulls the analysis toward the true atmosphere becomes stiffer when the deviation of the background state is larger. By comparison, with the above dynamic algorithm, since the assumed a priori observation error also grows with background error, the restoring tendency provided by GPSRO would still be flexible, even at large background errors. Thus, if this dynamic algorithm is used, GPSRO will guarantee only weakly that the experiment will not diverge from the true atmosphere. Experience, however, has never shown symptoms of divergence, neither in experimental nor in operational data assimilation cycles. This is not surprising since many other observations whose a priori errors are fixed are providing this restoring stiffness. Besides, this property is not new: procedures based on fixed a priori errors, which are applied to other data types, are also normally built to be flexible when the background deviation is large, for instance through quality control softening of the cost function, or deletion of data that present large observation minus background differences. On the other hand, this dynamic error estimation has allowed, without algorithmic modification, an evolution of the average a priori error with the assimilation system, as the forecast skill improved over several upgrades of the NWP system. GPSRO’s accuracy, known only very approximately at the initial implementation, has been demonstrated at progressive levels. Although this good potential was suspected during the initial implementation, it could not be sufficiently substantiated. The assimilation system has, during this evolution, provided increasing proof of the accuracy of GPSRO, allowed smaller a priori errors, and in turn extracted progressively larger benefit from it.

Also, and most importantly, dynamic error estimation has proven to be able to correctly identify, and properly handle, infrequent events where data were abnormally inaccurate, such as periods with significant receiver navigation errors, or large ionospheric contamination, before the providers flag them as below nominal. In these cases, the profile fits poorly with the background, and the data receive an appropriately low weight.

To summarize, we do not claim that dynamic error estimation of GPSRO is statistically optimal, or that it could not be outperformed. It is, instead, an approach that has successfully handled a new technology, and that has adapted to the evolution of the assimilation system, in a way that was sufficiently close to optimal to be useful. We, therefore, report it here for completeness.

f. Data screening

Besides general sanity checks for valid values, the following rules are applied. We mention the rationale behind each of them.

- Height must be at least 1 km above MSL. Data below may suffer from unmodeled phenomena: reflections, superrefraction, etc.
- Height must be at least 1 km above terrain. Data below may suffer from diffraction and reflections.
- Height must be below 40 km above MSL. Refractivity profiles are not expected to be accurate in the upper levels (contamination by ionospheric signatures, receiver navigation error, presence of gravity waves, and upper initialization of the Abel transform).
- Vertical thinning: Proceeding upward, after one value is accepted, no value is accepted during the next 1 km. Data values are probably correlated at smaller scales. Tests of assimilation at higher vertical density have not shown consistently better results.
- Values with absolute deviations larger than 0.05 in $(O - B)/B$ for refractivity are rejected. These values do appear, but are often questionable, and are associated with extremely irregular structures in the low troposphere or gravity waves in the upper stratosphere. In both cases, it is unadvisable to assimilate them.
Assimilation of data in the upper stratosphere is particularly problematic. Since GPSRO is a relatively fast measurement technique, besides other issues that could potentially have a technical solution (such as imperfect ionospheric subtraction, satellite navigation errors, or the upper boundary condition), it measures the instantaneous profile of the atmosphere, with a scanning time smaller than the period of gravity waves. Even if they were free of error, measurements obtained in the presence of gravity waves do not represent the conditions of hydrostatic equilibrium of that column of air. However, the field that the assimilation system is trying to estimate is the state of hydrostatic equilibrium. Therefore, it would not be appropriate to assimilate these upper data, even if the other sources of error were absent.

Concerning the lower height limit, here set at 1 km above MSL, other higher values have been tested. Notably, it is clear that the data show large variations, and the frequent presence of negative bias in the low troposphere is well established (Ao et al. 2003). However, the net effect of increasing the lower height limit, which would systematically withdraw more data in the low troposphere, was found to be negative when verified against radiosondes, in all regions and at all forecast ranges tested, up to 5 days. Other means of filtering the low troposphere, such as the rejection of data with large $O - B$ deviations (set at 5%), and the dynamic error estimation, were found to be better at handling lower-tropospheric structures such as superrefraction than a more strict filter based exclusively on altitude.

3. Data assimilation experiments

a. General setup

A series of data assimilation experiments carried out with the operational EC global forecast system (Charron et al. 2012) is described in this section. The system uses the General Environmental Multiscale model (GEM; Côté et al. 1998a,b). The assimilation is performed with either a 3D-Var scheme, which had been employed operationally before the implementation of 4D-Var (Gauthier et al. 2007). The 3D-Var scheme employed here evaluates the same background trajectory as the operational 4D-Var [i.e., First-Guess at the Appropriate Time (FGAT)], and is chosen for this study because it is computationally cheaper than 4D-Var, although it retains nearly all its properties, and is appropriate for assessing the impact of small modifications. Also, we have tested the system under an ensemble variational (EnVar) data assimilation scheme, which is scheduled to become operational by the end of 2014, and uses information from an external ensemble system to provide flow-dependent correlation information over the assimilation window (Buehner et al. 2013). This test was only performed to verify that the general behavior of the system is consistent under system updates representative of both the current operational system, and the immediate future.

The experiments were tested over several periods: 15 December 2008–1 February 2009 and 31 January–14 March 2011. A third period, 30 June–2 August 2011, was also tested for qualitative agreement over an updated version of the system, but is not shown in this paper. The 3D-Var scheme is used for the first two periods, whereas EnVar is used for the third. These experiments cover several generations of the forecast system and detailed system configurations, including a wide range of model resolutions and assimilated data volumes. This variety of trial periods and data assimilation schemes is expected to underline the consistent impact of radio occultations as anchor data.

Within each group of tests, control experiments have been spun up for several weeks before the start of the test period, to ensure that the dynamic bias corrections for the radiances are in equilibrium at the beginning of the control experiment. All departures from the control start at that point, and are allowed a further spinup to allow a transient period of the state of their respective bias correction procedures. Statistical results are evaluated over the remaining period. The list of experiments is given in Table 4. We explore various refractivity formulations, the role of the compressibility of air, and of the dynamical bias correction for radiances implemented in the EC global data assimilation system.

The observations used in these experiments are those that were operationally assimilated at EC during each period. They are from radiosondes, dropsondes, aircraft reports, land stations, buoys, ships, National Oceanic and Atmospheric Administration (NOAA) wind profilers, atmospheric motion vectors from geostationary satellites, and from the Moderate Resolution Imaging Spectroradiometer (MODIS) on board Aqua and Terra polar-orbiting satellites, GPSRO, Advanced Scatterometer (ASCAT) data, operational geostationary satellite imagers (GOES, Meteosat, MTSAT), the atmospheric infrared sounder (AIRS), and Advanced Microwave Sounding Unit (AMSU) A and B sensors. All were kept identical for the respective control experiments of each group (CTL, CTLB, and SL03). Variations are introduced only in the handling of GPSRO data.

The second period tested includes, with respect to the first, a number of system upgrades, and notably for this paper, a substantial increase (about twofold) in the volume of radiance data assimilated, largely due to modifications of their thinning, allowing a larger data density in time and space. The third period’s main
upgrades are the switch to the EnVar assimilation scheme, and an increase in model resolution from 25 to 15 km.

b. Bias correction

As for all major global forecast systems, the largest volume of data for EC’s global system are satellite radiances, which are required to be bias corrected to provide a good contribution to forecast skill (Eyre 1992). We specify EC’s details below, but the main principles that are relevant here are common to implementations in other major centers (e.g., Derber and Wu 1998; Harris and Kelly 2001; Dee and Uppala 2009; Auligné et al. 2007) that use dynamic radiance bias correction. We summarize here the main general issues. The bias correction procedure assumes that for several reasons radiance data are biased, and tries to fit a number of floating parameters, to reduce or eliminate systematic biases between radiances and model fields. These parameters are frequently updated, although some inertia is introduced in this update, so the parameter evolution will be slow compared to weather patterns. If radiance data show bias against the model fields, the fit slowly adjusts to reduce it. Therefore, bias-corrected radiance data are effectively allowed to float. This adjustment of the fit also affects the assimilation of future radiance data, producing a feedback over the following days. This feedback, with model adjusting to data (assimilation), and data adjusting to the model (bias correction), could potentially drift indefinitely. Instead, this feedback damps toward a floating equilibrium where new radiance data, once bias corrected, show very small bias, and the fit is stable over time. This equilibrium depends on properties of the model such as its climate trends, and also of other data also assimilated that are not allowed to float. The ensemble of model properties and of those other data, which damps the drift and determines this equilibrium, is commonly said to anchor the system.

At EC, this radiance bias correction procedure is dynamic (Garand et al. 2005), with its parameters updated at every step of an assimilation sequence. For each analysis, the previous few days, in this case 7, of innovations (i.e., observations minus short-range forecasts during each 6-h assimilation window) are used to estimate a fit to the observation biases, following the multiparametric fit approach proposed by Harris and Kelly (2001). The resulting parameters of these bias corrections, therefore, depend on the data assimilation experiment where the innovations are evaluated, and evolve with time. The bias corrections for the higher-peaking AMSU-A channels are however obtained from a static bias correction scheme, which follows the same methodology, but do not evolve with time. They had been fixed before each test period and, within each period, remain identical in all experiments considered here. These static channels are AMSU-A 11–14 in the first two periods, and 13–14 in the third.

The upper statically corrected channels do not ever produce feedback. The control experiments presented here were already in equilibrium with the bias correction fits. Since the modifications applied in all experiments shown here are small, we assume that after allowing a small spinup period, the bias correction evolution and feedback will also be near equilibrium. Previous experience shows that this is largely sufficient for small modifications like the ones explored here.

Under normal operation, a cycle evaluates its own fit for the evaluation of bias correction. This is what exposes it to potential drift, and why an anchor is important. We can alternatively use, in a given experiment, the history of dynamic bias correction parameters evaluated by a second experiment for the same period. If the two experiments are similar, the bias correction parameters of the second will also be applicable to the first. These bias correction parameters also evolve with time, but in this latter case the first experiment cannot suffer any feedback, as the bias correction parameters are already fixed by the second experiment.

### Table 4. List of experiments carried out in this study that explore the best physical implementation. The summarized properties are the assimilation scheme (4D-Var, or 3D-Var with FGAT), the presence of radio occultation (RO) data, the refractivity expression, the air equation of state (EOS, ideal or nonideal), and the experiment where the dynamic bias corrections (DynBC) for radiance data are evaluated (itself, or another experiment). Refractivity and EOS are only applied in the GPSRO operators, and thus they do not apply to GPSRO-denied experiments.

<table>
<thead>
<tr>
<th>Expt</th>
<th>Scheme</th>
<th>RO assimilation</th>
<th>Refractivity</th>
<th>EOS</th>
<th>DynBC</th>
<th>Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operational</td>
<td>4D-Var</td>
<td>Yes</td>
<td>R02</td>
<td>Nonideal</td>
<td>Itself</td>
<td>Winter 2009</td>
</tr>
<tr>
<td>CTL</td>
<td>3D-Var–FGAT</td>
<td>Yes</td>
<td>R02</td>
<td>Nonideal</td>
<td>Itself</td>
<td>Winter 2009</td>
</tr>
<tr>
<td>AL11</td>
<td>3D-Var–FGAT</td>
<td>Yes</td>
<td>AL11</td>
<td>Ideal</td>
<td>CTL</td>
<td>Winter 2009</td>
</tr>
<tr>
<td>SW53</td>
<td>3D-Var–FGAT</td>
<td>Yes</td>
<td>SW53</td>
<td>Ideal</td>
<td>CTL</td>
<td>Winter 2009</td>
</tr>
<tr>
<td>R02</td>
<td>3D-Var–FGAT</td>
<td>Yes</td>
<td>R02</td>
<td>Ideal</td>
<td>CTL</td>
<td>Winter 2009</td>
</tr>
<tr>
<td>AL11 + CMP</td>
<td>3D-Var–FGAT</td>
<td>Yes</td>
<td>AL11</td>
<td>Nonideal</td>
<td>CTL</td>
<td>Winter 2009</td>
</tr>
<tr>
<td>AL11 + CMP + DYN</td>
<td>3D-Var–FGAT</td>
<td>Yes</td>
<td>AL11</td>
<td>Nonideal</td>
<td>Itself</td>
<td>Winter 2009</td>
</tr>
<tr>
<td>noGPSRO</td>
<td>3D-Var–FGAT</td>
<td>No</td>
<td>—</td>
<td>—</td>
<td>CTL</td>
<td>Winter 2009</td>
</tr>
</tbody>
</table>

For example, the operational 4D-Var experiment uses observations from RO data. The dynamic bias corrections (DynBC) for radiance data are evaluated (itself, or another experiment). Refractivity and EOS are only applied in the GPSRO operators, and thus they do not apply to GPSRO-denied experiments.
Some of the experiments described here evaluate their own radiance bias corrections. In others, the corrections are taken from another experiment. We use this to allow or block the feedback mentioned, and explore its impact. We will explore the impact of some modifications. The ability to use a fixed history of bias correction allows the separation of the impact of this modification between their direct contribution, at fixed bias corrections, and their indirect contribution through the dynamic bias correction, such as the following:

- System A: Performs its own dynamic bias correction.
- System B: Like A, plus some modification. Takes bias correction from A.
- System C: Same modification as in B. Performs its own dynamic bias correction.

c. Test of different refractivity expressions

The refractivity expressions proposed by AL11, SW53, and R02 for the assimilation of GPSRO data were considered for several experiments, hereafter labeled as AL11, SW53, and R02, respectively. Air is here assumed to be an ideal gas (its compressibility factor $Z$ has the trivial value of 1). These experiments were run over the first period mentioned, in 2009 (see Table 4 for details). These experiments use the same history of dynamic bias corrections for radiances estimated and applied in the 3D-Var control system, which in turn closely follows EC’s 4D-VAR operational global forecast system at that time. Since the experiments do not calculate their own bias corrections, the bias correction feedback mentioned above is blocked.

The operational system at that time used the R02 refractivity expression, and applied a realistic (nontrivial) compressibility factor. Figure 2 shows the mean innovations for temperature from the global radiosonde network for January 2009 from AL11, SW53, and R02. Most of the impact is below 400 hPa. The results for AL11 and SW53 are very close, whereas the refractivity formulation in R02 leads to a colder bias in the forecast temperature, which was described in Aparicio et al. (2009), and confirmed by Healy (2011) and Cucurull (2010). The expression R02 provides a slightly larger dry refractivity for given conditions. The analysis fits this with a denser (colder by about 0.1°C) low troposphere. This modifies the vertical thickness of the low troposphere. In altitude or geopotential level, the column appears to be shifted downward. In temperature versus pressure, only the low troposphere is affected. Following Aparicio et al. (2009), the atmosphere would indeed be denser, but attributes the reason not to a colder temperature, but to nonideal gas effects. Pressure exerted by air is slightly smaller, for a given molar density, than an ideal gas. The best results are obtained with the AL11 experiment. The warm bias peaking at 200 hPa in all experiments is due to the aircraft temperature observations. Above 150 hPa, the GEM model is known to be systematically colder than radiosonde data.

Concerning the choice of the physical constitutive relations, namely, the equation of state and the refractivity expression, this indicates that results are improved when both the AL11 expression and the real-gas compressibility factor are used in the observation operator. This is consistent with the studies of Aparicio et al. (2009) and AL11 that suggested, independently of an NWP system, that these constitutive relations are of higher quality. It is also consistent with explorations of NWP performance under several constitutive relations by Aparicio et al. (2008), Healy (2011), and Cucurull (2010).

We note that not only the source dataset is identical in all these runs, but also the assimilated dataset, after any filter, quality check, and bias correction. These are taken from an external control run. Thus, only the GPSRO forward operator differs between them.
d. Test of nontrivial compressibility

As shown by Aparicio et al. (2009), the deviation of the compressibility factor of air, from the trivial value of 1, should be taken into account in the estimation of observation heights. The nontrivial compressibility is smaller than one: a parcel of air of given molar density and temperature exerts a pressure slightly smaller than an ideal gas. When this is neglected, the system fits the observations reducing the temperature in the lower troposphere. This is mostly a hydrostatic effect. From (6), we can see that the virtual temperature and the hydrostatic pressure depend on the compressibility factor. By incorrectly setting the latter to 1, but telling the system that GPSRO is a trustable source of thermal data, we effectively force the system to lower the temperature and fit the occultation profiles. An experiment (AL11 + CMP) was therefore prepared, which combines the AL11 refractivity and considers the real-gas compressibility factor in the hydrostatic equation of GPSRO’s height operator. See Table 4 for details. This is shown in Fig. 3, and confirms the proposal by Aparicio et al. (2009) to introduce nonideal gas behavior to evaluate the hydrostatic profiles. In the AL11 + CMP experiment, GPSRO introduces a near-neutral thermal bias with respect to radiosondes.

Both AL11 and AL11 + CMP use the bias-corrected radiances of the control experiment (which is R02 + CMP + DYN). Since neither AL11 nor AL11 + CMP experiments evaluate their radiance corrections, these modifications in the treatment of GPSRO refractivity or height cannot have any impact on the radiance bias correction. The difference shown in Fig. 3 between these two experiments is exclusively due to the direct impact of GPSRO on the system.

We will then keep the treatment of GPSRO to the AL11 refractivity, nonideal gas (CMP), and see if allowing the adjustment of the radiance bias corrections to the new conditions, which is the normal mode of operation, leads to significant differences. This will be AL11 + CMP + DYN.

e. Test of dynamic bias correction feedback

Since both the choice of the refractivity expression and the compressibility have an effect on the mean temperature profile, particularly in the troposphere, the bias corrections for radiances may also be impacted if these corrections were estimated dynamically from previous innovations in the data assimilation cycles. We explored the size of this impact launching an experiment that evaluates its own dynamic bias corrections, instead of using the history from the control (CTL) experiment, therefore, allowing the cycle to receive a feedback from previous innovations of radiances, in addition to the data assimilation itself. We label this new experiment as AL11 + CMP + DYN. With this notation, the CTL experiment corresponds to R02 + CMP + DYN. Figure 3 shows the mean innovations for the AL11, AL11 + CMP, and AL11 + CMP + DYN experiments. The difference between AL11 + CMP and AL11 + CMP + DYN is representative of the size of the bias correction feedback to a difference in the implementation of the GPSRO operator (again, the control experiment used R02, not AL11), everything else being equal.

Allowing the bias correction scheme to operate self-consistently tends to reduce the overall bias in the entire column. This is possible because radiosondes and GPSRO data are in good agreement, which allows a better bias correction. This agreement also supports the choice of the constitutive relations. The thermodynamic biases that remain are dominated by known issues with other data types (e.g., aircraft temperature reports) and known forecast model weaknesses.

A better refractivity expression, as well as the implementation of the nonideal compressibility in the
observation operator also leads to forecast skill improvements, as shown in Fig. 4 for the geopotential height at 250 hPa over the globe. This field is particularly relevant in this study since it is representative of the air mass in the troposphere where most of the impacts of the compressibility factor $Z$, and the refractivity expression examined here, are found. Note that the compressibility factor does not affect as much (Aparicio et al. 2009) the interpretation of data at layers where air is nonideal, as it affects the altitude of the column above, at layers where air is very close to ideal. In the top panel, day-1 to day-5 anomaly correlations are plotted for all mentioned 2009 experiments, as well as for an experiment where the GPSRO data are not assimilated (noGPSRO). In this latter experiment, the radiance bias corrections are still those from the control experiment. The bottom panel shows the anomaly correlation loss incurred by the noGPSRO run, when it is compared against each of the runs that implement GPSRO assimilation. A larger loss of denial implies that the assimilating run had better performance. In these comparisons, it can be seen that at equal data (including GPSRO data), model, and general setup, different implementations of GPSRO assimilation (refractivity expressions, equation of state of air) lead to a very substantial spread in skill gain, as much as a factor of 2 at 5 days, and even larger at shorter lead times. The best results are obtained with the AL11 + CMP + DYN experiment, which has a higher anomaly correlation, and where a denial of GPSRO leads to larger loss. This spread is a large fraction of the total impact from assimilating GPSRO data, and implies that the net impact of GPSRO to the forecast skill of the system does not depend only on the amount of data available, but also on the fine details of the thermodynamic interpretation of these data.

Both the compressibility and the use of a self-consistent dynamic bias correction scheme for radiances improve the forecast skill, whereas the choice of the refractivity expression plays a minor role. It is interesting to note [see Fig. 4, curves (noGPSRO) − (AL11 + CMP) and (noGPSRO) − (AL11 + CMP + DYN)] that a substantial fraction of the anomaly correlation gain available from the assimilation of GPSRO through the best choice of equation of state (nonideal) and refractivity (AL11) is obtained through the use of dynamic bias correction. This means that a substantial part of the benefit of GPSRO appears through their participation in the calibration of radiances, and not exclusively through their direct assimilation.

It is interesting to note, in view of both the anomaly correlations of different cycles, Fig. 4, and the agreement with radiosonde data, Figs. 2 and 3, that the forecast skill is better in those cycles where the agreement with radiosondes is better. As should have been expected, this suggests that agreement between anchor datasets has a significant forecast skill value. This consistency should be understood as including not only the data, but also the physical relationships between the different kinds of data. In the GPSRO case, these are notably the equation of state and refractivity expression. Elsewhere in the system, this may suggest that the quality of other constitutive relationships may have a similar impact on skill.

4. Direct and indirect impact of GPSRO data

Having reached the conclusion that for both theoretical reasons, and practical tests, the best anchoring is obtained with the constitutive relations AL11, for the refractivity, and with nonideal gas pressure in the hydrostatic equation, we explore the paths for this anchored information to flow into the NWP system, and impact its performance. This was done launching GPSRO denial...
experiments, which also allowed the evaluation of the general impact of GPSRO under conditions representative of the updated operational system.

As indicated by the experiments mentioned above, GPSRO not only has a direct impact through its assimilation, but also an indirect impact through the estimation of radiance bias corrections, as suggested by the difference between AL11 + CMP and AL11 + CMP + DYN.

### a. Experimental setup

A set of four experiments was prepared, covering a 2 × 2 contingency table: with and without assimilation of GPSRO data, and with and without the impact of GPSRO data on the radiance bias corrections. That is one control, two partial denial experiments, and one full denial experiment. To verify the robustness of the above results, this was done in a new period, in the boreal winter of 2011, and with an updated version of the NWP system and data ensemble, to verify that the qualitative results are maintained. See Table 5 for the set of four experiments.

The NWP system normally evaluates the bias correction coefficients for radiance data from its own observation minus background statistics. However, we may use the coefficients stored from another experiment.

1) CTLB: follows the operational cycle of that time, but using 3D-Var as data assimilation scheme. Its radiance bias correction coefficients are stored.
2) noGPSRO: denies GPSRO data for all purposes, but is otherwise identical to CTLB. Its radiance bias correction coefficients are also stored. Full denial.
3) noGPSRO-X: without GPSRO assimilation, but takes the history of radiance bias correction coefficients from the recorded CTLB. Assimilation denial.
4) GPSRO-X: with GPSRO assimilation, but takes the history of radiance bias correction coefficients from the recorded noGPSRO. Denial of the bias correction feedback.

### b. Results

The robustness of the result was tested and confirmed by preparing the above mentioned sets of four experiments on two different periods, including full implementation of GPSRO, full denial, and two partially denied experiments. The general setup of the NWP system was comparable, and was representative of current operational standards. Between each block, significant upgrades have been implemented.

In Fig. 5, the evolutions of two particular channels (AMSU-A 8 and 10, in NOAA-16) over the course of two experiments are shown. These are the histories of the self-consistent experiments CTLB (with GPSRO) and the denial (noRO), from the second block of experiments (boreal winter, 2011). See Table 5 for details.

The departure between the two histories of bias correction estimations can be seen. As with other channels, the approximate size of the difference between experiments is of the order of a few hundredths of a kelvin. The same behavior can be seen in a similar set of experiments with an updated version of the system (for boreal summer 2011, not shown). It should be noted that the departure between experiments is very slow, and on time scales substantially longer than the internal variability within each experiment. This internal variability shows most activity at daily to weekly time scales. The difference of bias correction between the GPSRO allowed and denied experiments is very close to linear over the entire period. With no evident saturation, it is to be expected that longer experiments would show an increased departure. Since the departure was still linear after 40 days, the experiments were not continued, as it would have been prohibitively expensive (several months) to wait for saturation.

Figure 6 shows the skill loss of GPSRO denial over several experiments, including full denial. The skill loss was compared over the different periods and detailed configurations, and found to be consistent, up to and including the most recent experiments in a one month period of boreal summer 2011 (not shown), which use an upgraded NWP system with EnVar assimilation, and that is an advanced candidate to become operational during 2014.

It can be seen in Fig. 6, that most of the skill impact from GPSRO data results from their participation in the
However, a significant fraction stems from the impact on radiance bias correction, and there is a significant skill dependence on the history of radiance bias corrections, even if the set of assimilated data is kept constant.

Besides GPSRO there are many other calibration constraints available, but these data are concentrated in the Northern Hemisphere, and at lower altitude, such as radiosonde, aircraft, and surface stations data. As could be expected, this dependence on the use or denial of GPSRO data is especially large where these bias corrections are less constrained, namely in the Southern Hemisphere, and at a higher altitude. Skill impact of GPSRO use or denial is smaller in the Northern Hemisphere, and virtually all of this impact stems from the participation of GPSRO in the direct assimilation (i.e., in the cost function). Denial of GPSRO in the assimilation leads to measurable skill loss, but has a near-neutral impact on the bias correction history of the Northern Hemisphere.

Upper layers are less constrained worldwide, but especially so in the Southern Hemisphere. Denial of GPSRO can be seen to lead to skill loss not only through the assimilation, but also through their participation in the estimation of radiance bias corrections. Indeed, where these corrections are otherwise less constrained, GPSRO contributes a significant amount of skill by having merely contributed to the determination of the radiance bias corrections, even if GPSRO profiles have not been assimilated. In the Southern Hemisphere, both...
the experiments where GPSRO data are used or denied in the assimilation show a clear dependence on the history of bias correction that is applied. The GPSRO-denied experiment is improved by using a bias correction history evaluated with GPSRO data. Also, the GPSRO-allowed experiment shows a degradation if the bias correction applied stems from a GPSRO-denied experiment.

5. Conclusions

A number of data assimilation experiments were conducted to evaluate the impact of several physical constitutive relationships that are involved in the use of GPSRO data in NWP as an absolutely calibrated data source, notably the equation of state of air and the refractivity expression. Nonnegligible gains were obtained with the use of the relationships that are presumed to be most accurate, indicating, as had previously been suggested (Aparicio et al. 2008), and further explored in several directions (Aparicio et al. 2009; Healy 2011; Cucurull 2010; AL11), that the impact of the NWP assimilation of GPSRO data is sensitive to the accuracy of these expressions. We have presented our updated implementation, which is presently operational at EC, and that is based on these presumably best physical constituent
relationships. We find that these are the ones which lead to better agreement with radiosondes and to the best forecast results. Comparing against a GPSRO-denied cycle, it is shown that the forecast skill that GPSRO adds to the system may vary by a factor of about 2 between the best and worst choice of constituent relationships and bias correction feedback, everything else being equal (NWP system, assimilated dataset). The best skill is associated, as should be expected, with the cycles where the different anchors are in better agreement. We, therefore, recommend an investment of effort to verify that the different anchor technologies inject coherent information to the system, in order to maximize the return of the investment in the deployment of receivers.

We have explored the impact of GPSRO data under conditions representative of present operational setups at EC, using denial experiments to explore the loss of skill that the system would incur if GPSRO data were denied. A group of experiments compared full use of occultation data (as in the operational setup), and several denials of occultation data: denial in the bias correction procedure, but not during assimilation; denial in the assimilation, but not in the bias correction procedure; and complete denial. It is shown that, although the contribution through assimilation is larger, GPSRO contributes a significant fraction of its added forecast skill through its participation as anchoring data in the calibration of the radiance bias corrections. This contribution to skill gain through calibration is particularly significant in the Southern Hemisphere, where GPSRO has been a major addition to the ensemble of data assumed to be traceable, and that are used as anchors to the NWP system.

During the different denial experiments applied, it was seen that the information of GPSRO data has a nonnegligible participation both as assimilated data during the variational minimization procedure, and through the determination of the radiance bias correction. Therefore, we must conclude that there is a non-negligible amount of calibrating information, provided by GPSRO, which is imprinted in the radiance data during the procedure of dynamic bias correction. This information is provided to the variational assimilation system, and participates in the observation cost function, through the radiance data. This participation is separate from the direct participation of GPSRO observations in the cost function. This imprinting appears to grow very slowly. This is likely due to slow accumulation, since the bias correction procedure of our system is otherwise able to respond at time scales of less than a week. The accumulated effect of the denial of GPSRO continued to grow linearly, and showed no sign of saturation after 40 days, where it already shows an impact in the ability of radiance data to contribute forecast skill to the system, independent of the forecast skill contributed by GPSRO itself. We assume that the differences between experiments would only grow larger in the direction already indicated by the existing series.

The above results underscore the importance of a good physical calibration of the system, and particularly of the data, in this case GPSRO profiles, but likely applicable to other sources that provide traceability links to the NWP system. This is particularly relevant as the forecast skill improvement obtained from these data can range by approximately a factor of 2 between different implementations of the calibration links, including the physical relationships of calibrated data, and the bias corrections of floating data.

Acknowledgments. The authors are grateful to Louis Garand (EC), and two anonymous reviewers, for useful comments, and to the following providers of GPSRO data for offering access to their data products, including the University Corporation for Atmospheric Research (UCAR) and the National Space Organization of Taiwan (NSPO) for COSMIC data, the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) for METOP/GRAS data, and the German GeoForschungsZentrum (GFZ) for GRACE data.

REFERENCES


——, G. Deblonde, and S. Laroche, 2008: Assimilation of GPS radio occultation measurements at the Meteorological Service of Canada. GRAS SAF Workshop on Applications of GPSRO Measurements, Reading, United Kingdom, ECMWF, 95–98.


Buehner, M., J. Morneau, and C. Charette, 2013: Four-dimensional ensemble-variational data assimilation for global deterministic


Garand, L., G. Deblonde, D. Anselmo, J. M. Aparicio, A. Beaulne, J. Halle, S. Macpherson, and N. Wagneur, 2005: Experience with bias correction at CMC. Bias Correction Workshop, Reading, United Kingdom, ECMWF, 153–162.


Rennie, M., 2008: The assimilation of GPS radio occultation measurements at the Met Office. GRAS SAF Workshop on Applications of GPSRO Measurements, Reading, United Kingdom, ECMWF, 84–94.
