Assimilation of GPS radio occultation measurements at the Meteorological Service of Canada

Josep M. Aparicio, Godelieve Deblonde and Stephane Laroche

Data Assimilation and Satellite Meteorology Division
Environment Canada

Abstract

We present the infrastructure to assimilate Radio Occultations (RO) at Environment Canada (EC), including its historical development, a critical analysis of the choices that were made, the results that have been obtained, and the future development lines that are currently under consideration.

1. Historical outline

Radio occultation work was initiated at Environment Canada after the success of the GPS/Met mission, along the late 1990s. It was however only in 2003, with missions CHAMP and SAC-C providing substantial amounts of data, that a major plan was set to prepare for the upcoming NRT availability of this kind of data. By then it was evident that not only those instruments would be in orbit, but a new upgraded generation would be in place in a short term future, including the COSMIC constellation and the GRAS/METOP mission. Both due to the volume and quality of the upcoming data, there was a qualitative leap with respect to earlier outlooks that made this ensemble of data not just supplemental to existing systems. It was instead evident that this would represent one of the core sets within the available Earth Observation data.

An ensemble of observation operators (forward, tangent linear and adjoint, for both refractivity and bending angle) was prepared along 2003-2004, which were tested in the global analysis/forecast system at EC. At that time, the global model had a horizontal resolution of 100 km, 28 vertical levels and a lid at 10 hPa.

The results were encouraging and positive (Aparicio and Deblonde, 2008). However, they lacked several important elements. One essential mandate for the EC system is to forecast extreme weather events, which may lead to large social impact. This means that, in order to be considered as an improvement from an operational perspective, a modification to the system must succeed particularly in the description of the low troposphere and the weather elements.

However, early RO impact was good at medium and high altitude, yet small or even negative at low altitude. Similarly, the temperature field was improved in general, but the moisture field often presented a neutral or even negative impact. This meant that although it seemed promising within the research team, the technology did not yet meet the criteria for operational implementation.

The problem was not so much one of quality of the data, but of making a good use of them. Indeed, meteorological analysis/forecast systems have a large amount of parameterizations and fine-tuning adjustments, which may be put out of balance and degrade the performance if we inject a substantial amount of new information, as is the case with RO data. It took several modifications to get across this stalemate. Some of these were implemented independently of the RO data, others are more directly related, or were even developed to accommodate RO data:

1. A 4D-Var assimilation system was operationally implemented in 2005.

2. The global forecast system was upgraded in all three spatial dimensions, to a horizontal resolution of 33 km, and 58 vertical levels. The lid is still at 10 hPa.
3. The available RO data increased from just CHAMP, to include GRACE and the COSMIC constellation, an enhancement of the data volume by a factor of 10.
4. The estimation of the observation error (and thus the data weight) was made adaptive, background dependent, which was finally able to capture the full range of situations where a datum is assimilated.

The first two (4D-Var assimilation and higher resolution model) improve the a priori accuracy of the system, so the background state is closer to the true state when a new datum arrives. The assimilation is then more likely to take place within the convergence basin towards the true atmospheric state. The third increases the amount of RO data, whereas the fourth optimizes their weight during assimilation.

It is difficult to say how each of these has weighted in the qualitative improvement of the impact of RO data, but after this ensemble of changes, the footprint of the assimilation of RO has become more attractive from the operational point of view. The impact is then not limited to the high atmosphere and the temperature field, but extends to moisture fields, the low troposphere and the precipitation forecasts. These latter are particularly beneficial for EC’s mandate. The RO data was thus finally accepted by the operational committee June 3rd 2008.

In retrospective, this still points to the current major limitation of RO: although there is now a benefit in the forecast of the energy density field, it still does not seem to be providing all the potential benefit. The energy density field is largely dominated by the atmospheric moisture content, and drives most events of extreme weather. Our focus is currently in

1. Improving the impact in the low troposphere and especially the moisture field.
2. Further optimize the data weight.

2. The observation operator

The RO data is presented at several levels of post processing, which in practice amount to:

- Level 1b: Bending angle as a function of impact parameter.
- Level 2: Refractivity as a function of MSL height.
- Level 3: Dry temperature as a function of dry pressure.

The standard practice in meteorological data assimilation is to use rawer, less processed data, which points to Level 1b. However, there is a fundamental problem: the measurement itself proceeds from the outside towards inner layers (the best phase measurements are those at high tangent heights), whereas the level 1b and 2 operators construct the height coordinate (be it MSL or impact parameter) proceeding from the surface up. This mismatch between an inwards observation profile and an outwards model construct affects data at both level 1b and 2, and ultimately leads to the assimilation being nonlocal. This would rather suggest that despite standard practice we should choose a matching approach (i.e. Level 3 assimilation).

As this was judged too radical, only Level 1b and 2 were considered. However, it was recognized that the coordinate inversion (outward vs. inward) between observation and operator was risky and may lead to a number of problems.

From the point of view of the structure and complexity of the observation operator, there is little difference between a refractivity operator and a 1-dimensional (1-D) bending angle one. The coding would be very similar. However, some practical limitations appear. Whereas it is possible to build an accurate refractivity operator up to the model lid, a bending angle operator requires knowledge, or a sufficiently good extrapolation, of the atmospheric conditions up to infinity. This can be obtained reasonably with a high-lid model, but it is not the case with a low-lid model. With a lid at 10 hPa (~30 km), the bending angle operator does not deliver good accuracy.

This constrained us to use the level 2 option as the best choice, at least until the next model upgrade at EC, in which the lid will be at 0.1 hPa. This model is currently under advanced stage of testing and should be operational by spring 2009. The level 2 (refractivity as a function of MSL altitude) was thus the selected form for the RO data.

However, there are other reasons that in retrospective would have led us to equally select level 2: It was found that under many circumstances the presence of oscillating temperature structures at high altitude (above 25 km) was detectable and actually rather common. These structures are not necessarily artifacts of the GPS RO instruments or data processing. They can be seen also with other fast scanning technologies. They are often interpreted as gravity waves, but may also be related to other structures such as mountain waves, convection etc. It is relatively simple, from Brunt-Vaïsala frequency evaluations, to estimate the timescale of these structures, and they are too short to be viewed as part of the hydrostatic atmospheric structures. However, current variational data assimilation systems cannot represent the instantaneous hydrodynamic state. The profiles containing these oscillations, although do represent the true instantaneous state of the atmosphere, are thus not good candidates for assimilation. We want to assimilate the stable hydrostatic profile.

These oscillations are stronger when viewed in L1b space, than in L2 space. It is thus in this sense preferable to assimilate L2 data, as the result is less sensitive to the existence of those oscillations.

3. The assimilation impact

A number of tests have been performed, of progressively broader scope, finally leading to the full scale runs of four months that have recently been used to support operational acceptance. The tests consistently show, even from the earliest tests, that stratospheric temperatures and the associated winds (when in geosptrophic balance) are improved with the assimilation of RO data. It is however non-trivial to improve tropospheric temperatures, winds and moisture. These latter require, on one hand, being sufficiently close in the background state to make a good use of the data. Both the better results that are obtained with the upgraded high resolution model, as well as the improvement when the analysis scheme is 4DVar, instead of 3DVar, point in that direction. On the other hand, the low height data contain both very useful information on these aspects of the atmospheric state, and larger and more systematic observation errors. It is difficult to distinguish both, and careful weighting was needed to obtain improved results.

It is now unmistakable that RO data is filling some important gaps in the former data set. Notably the southern hemisphere, both polar regions, the stratosphere and the open oceans, are now much better covered. However, although midlatitude moisture contents are improved, the tropical and subtropical moisture structure is still insufficiently probed. Large amounts of small scale structures which are known to exist are not yet sampled, and this region corresponds also to the lowest density of RO observations.

4. Current status

The RO data has been accepted (June 3rd 2008) for operational use at EC, and is pending implementation and testing by the operations team, which should be completed within this summer. We believe that the approach has been very conservative, carefully analyzing the results and the causes for each paradoxical behavior. We must mention notably:

- The importance of variable gravitational acceleration within the atmosphere.
- The importance of variable gravitational acceleration on the effective topography.
- The relevance of the mismatch between the construction of the observable (inwards) and the construction of the observation operator (outwards).
- The importance of the non-ideal behavior of air.
- The critical impact of data weight (thus of the error estimate).

The RO data has also already been extensively tested, with very good results, in the upcoming version of EC's analysis/forecast system, which has a higher lid at 0.1 hPa (~65 km) and 80 vertical levels, and it is in all tests and plans assumed that RO data will be part of the near future operational implantation of this system.

5. Critical retrospective and future outlook

The choice of data that, when modeled, is built from the surface up (both the case for L1b and L2 data), implicitly contains the seed to many problems. In fact, a choice of L3 data may have permitted a much faster implementation with many of the benefits for the upper atmosphere, although at the cost of greater difficulty extending these benefits to lower levels.

The choice of L2 vs. L1b is still viewed favorably, as a low lid model was not appropriate for L1b data assimilation. With the upcoming version of the analysis/forecast system, the L1b assimilation becomes again a possible direction. However, it is only perceived as desirable over the current and extensively tested L2 assimilation provided operators beyond 1-D are prepared, as otherwise would not justify the known drawbacks of L1b data assimilation (larger sensitivity to hydrodynamic state of the atmosphere). There has been some work in this direction, with some low-cost 2-D approaches (within a factor of 3 of the 1-D cost) having been designed and preliminarily tested. This approach, however, would require, to be implemented operationally, of supplementary infrastructure of the assimilation system, which is still under development, and not available in the short term.

The low-cost 2-D operators would be based on the representation of the bending angle as a Hermite Gaussian quadrature along the line of sight. A 1-D operator is equivalent to a 1st order quadrature along the line of sight. Low cost operators are possible with increases of the quadrature order to some small integer. 3rd order quadratures have been tested in forward mode, but its implementation in the operational system would require the addition of extensive operator infrastructure to allow the implementation of the associated tangent linear and adjoint operators. This is currently ruling out the use of 2-D operators at EC for data assimilation purposes.

Beyond this, the most important reason why a 2-D operator may be superior is to represent the fine structure of the atmosphere along the line of sight. However, the data assimilation at EC corrects only large horizontal scale structures (above 150 km).

6. References

Aparicio, J. M., G. Deblonde, 2008: Impact of the assimilation of CHAMP refractivity profiles on Environment Canada Global Forecasts, *Mon. Wea. Rev.*, **136**, pp 257-275.