

AO 160 - OZVAL: validation of ENVISAT ozone products through assimilation in a CTM; first results obtained with GOMOS

Bertrand Théodore⁽¹⁾, Marielle Guirlet⁽¹⁾, Alain Hauchecorne⁽²⁾, Odile Hembise⁽¹⁾, Antoine Mangin⁽¹⁾

⁽¹⁾ACRI-ST, BP 234, F-06904 Sophia-Antipolis, France

⁽²⁾ Service d'aéronomie du CNRS, BP 3, F-91371 Verrières-le-Buisson, France

Email : ozval@acri-st.fr

ABSTRACT

This paper presents preliminary results obtained by assimilating GOMOS data into the MSDOL assimilation system in the frame of the OZVAL project. This project aimed at supporting Envisat data validation by providing the ACVT group with assimilated fields of ozone. The availability of GOMOS data allowed us to perform assimilation experiments using GOMOS data. We present here some of these results and discuss the use of data assimilation for validation.

1. INTRODUCTION

The OZVAL project aims at providing support to the validation of ENVISAT “atmospheric chemistry” instruments using numerical modeling and data assimilation. These techniques can indeed be used to:

- Estimate the observation and forecast error statistics from the difference between the measurements to be validated and the forecasts calculated by a model; in that case, the measurements are not assimilated (although the model can be constrained by any independent data).
- Provide synoptic views of the measured quantity (ozone in our case) to allow the comparison with external sources. In that case the measurements are used to constrain the model by optimal interpolation. An example of such a comparison is shown on fig. 1 on which two GOMOS profiles are compared with the output of the assimilation system using which Odin/SMR measurements. The assimilated field is used here as an independent reference for all locations and times, and in particular those of the GOMOS profiles.

These ambitious objectives were only partially approached during the commissioning phase. This paper gives an overview of the achievements so far. We indeed used only GOMOS data which were easily available to us. Also, we focused on the period corresponding to the breakup of the Southern hemisphere polar vortex that occurred late September 2002.

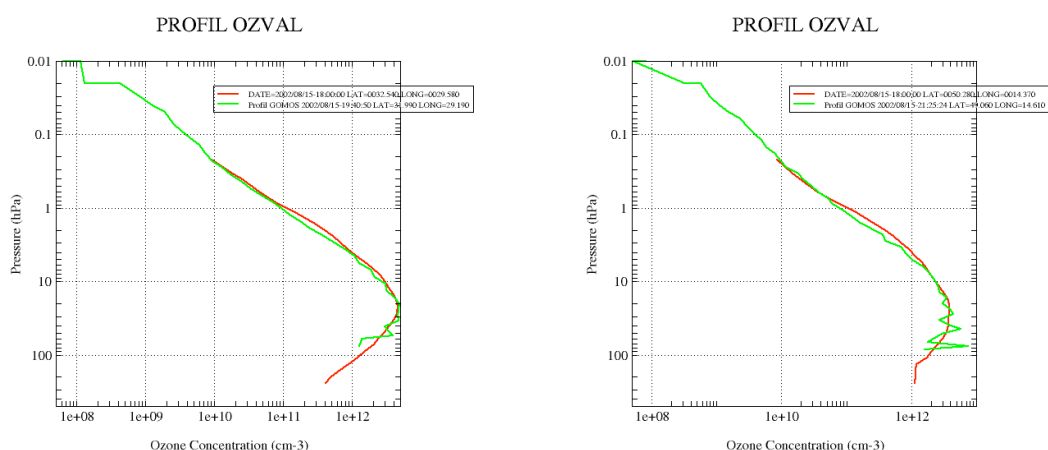


Fig. 1. Comparison between two GOMOS profiles (in green) and the output of the assimilation system MSDOL in which SMR data have been ingested (in red). Both profiles are in reasonable agreement except between 10 and 1 hPa.

The model and the assimilation algorithm will be briefly described as a first step. Emphasis will be given to the pre-processing of the assimilated data and data selection. It indeed turned out that this is quite critical due to the large oscillations present on much of the GOMOS ozone vertical profiles. The results of the assimilations of these data are then presented along with a discussion on the possibility of using data assimilation for instrument validation.

2. THE MSDOL SYSTEM

MSDOL includes both a CTM based on the ROSE model and a sequential assimilation module that combines observations and forecasts of the model by optimal interpolation.

2.1 The CTM

The CTM used in the MSDOL system is based on the model ROSE from NCAR [1]. Its resolution is 2.5 degrees both in latitude and longitude. The domain is divided into 36 pressure levels from 240 to 0.3 hPa, yielding a vertical resolution of about 1.3 km. The horizontal and vertical winds as well as temperature are prescribed from ECMWF analyses.

There are 33 chemical species or families of species interacting through a set of 145 chemical reactions; heterogeneous chemistry is included. The long-reacting species are transported using a semi-lagrangian algorithm that gives accurate results for the long time steps used in the model (30 minutes).

2.2 The assimilation module MAD

The assimilation module, referred to as MAD, makes use of the classical equations of optimal interpolation to analyze the ozone field. If \mathbf{x}_a is the analyzed state vector, \mathbf{x}_f the forecast state vector and \mathbf{x}_o the vector of the set of observations assimilated during the considered time step, the following relation holds:

$$\mathbf{x}_a = \mathbf{K}(\mathbf{x}_f - \mathbf{H}\mathbf{x}_o) + \mathbf{x}_o \quad (1)$$

where

$$\mathbf{K} = \mathbf{B}(\mathbf{B} + \mathbf{H}\mathbf{R}\mathbf{H}^T)^{-1} \quad (2)$$

The weight \mathbf{K} depends on the forecast error covariance matrix \mathbf{B} , on the observation error covariance matrix \mathbf{R} and on an interpolation operator \mathbf{H} mapping the 3D field on the observations locations.

The forecast error covariance matrix \mathbf{B} is assumed to be the product of a fixed correlation matrix \mathbf{C} and of a diagonal matrix containing the variances of the ozone field. The latter are assumed to be the sum of the time evolution of the variances from the previous time step (transport + chemistry) and a term representing the model error:

$$\mathbf{B} = \mathbf{C}\mathbf{D} + \mathbf{E} \quad (3)$$

The correlation matrix of the forecast error only depends on the distance. It is a gaussian of 1500 by 500 km in the zonal and meridian directions, respectively. The vertical correlation is such that it is significantly different from 0 for the 2 levels around the considered point.

2.3 Data pre-processing

An IDL routine is used to pre-process data to be assimilated in order to :

- Read Level2 products from their native format;
- Select useable data;
- Convert concentrations to mixing ratios using ECMWF temperature fields;
- Compute matrix : each selected point is located on the model grid and the corresponding interpolation coefficients are computed;
- Save these data in a file such that there is one file per profile.

2.4 Data selection

Data selection is an important issue in every assimilation system. It is even more critical when considering GOMOS data. Indeed, it appeared very early that most GOMOS profiles exhibit strong oscillations and/or unrealistic values. Furthermore, the estimation of the measurement errors seems to be quite whimsical. Both aspects combine to yield wrong values in the assimilated fields. It is thus important to screen the data.

Fig. 2 show some examples of data selection on some GOMOS profiles. Note that this focuses on “bad” profiles.

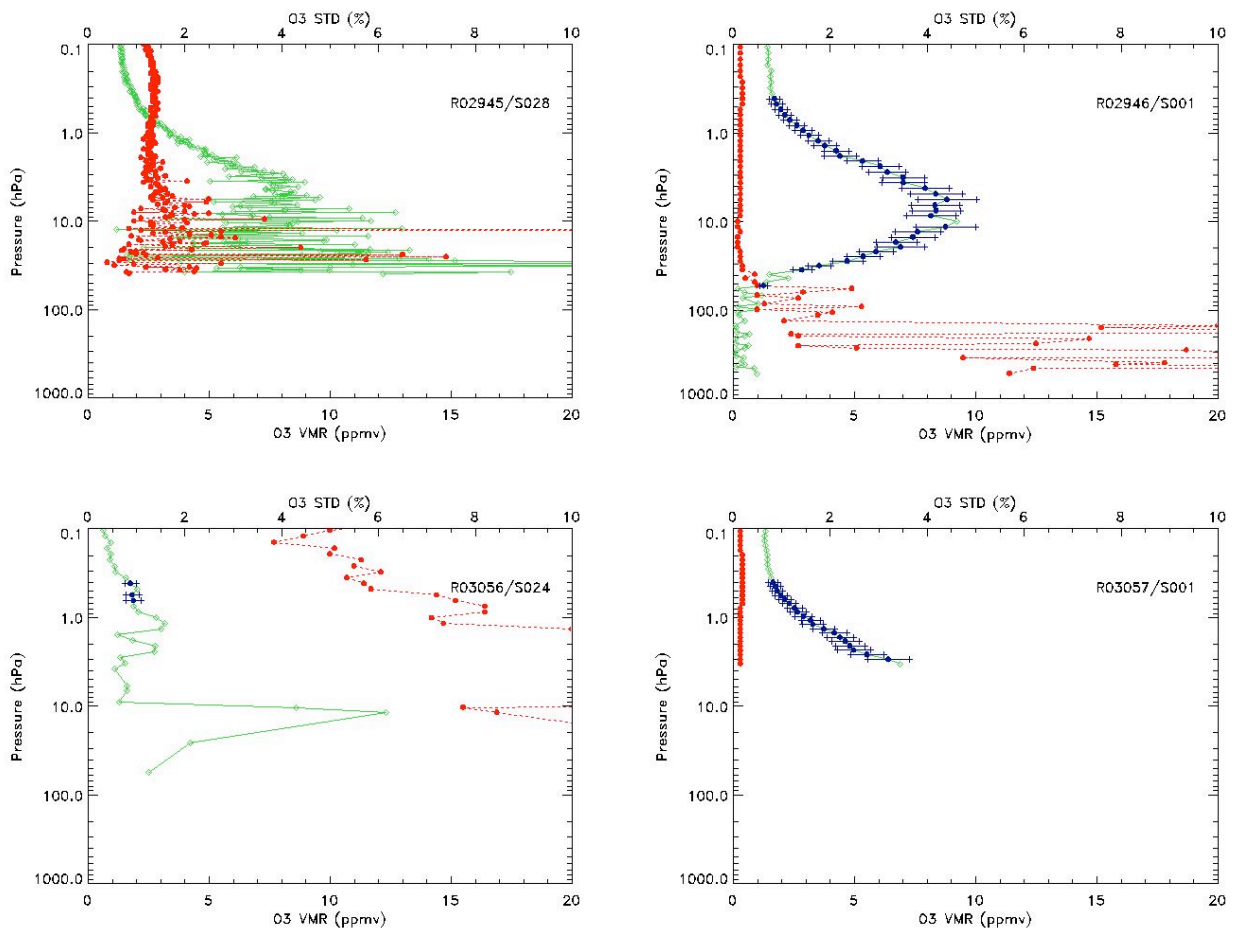


Fig. 2. Illustration of the effect of data selection on some GOMOS profiles. The green line is the actual GOMOS profile, the flagged points, as well as those for which the relative error is greater than 100%, having been excluded. The red line is the error bar as written in the GOMOS product, and for which the scale is at the top of each plot. The blue points are the points that are selected for assimilation. The associated error bar is artificially set at 15 % (see text); indeed, it can be noted that in general the relative error written in the GOMOS products is of the order of 1% or less, even when strong unrealistic fluctuations are observed.

3. ASSIMILATION OF GOMOS PRODUCTS

3.1 A view of the breakup of the southern hemisphere polar vortex

We focus here on the period 16-24 september 2002 during which the Southern polar vortex split down. The location of the GOMOS measurements, shown on fig. 3, indicate that GOMOS did not perform any measurements inside the vortex at that time. It is thus interesting to see how assimilation deals with this regional lack of data.

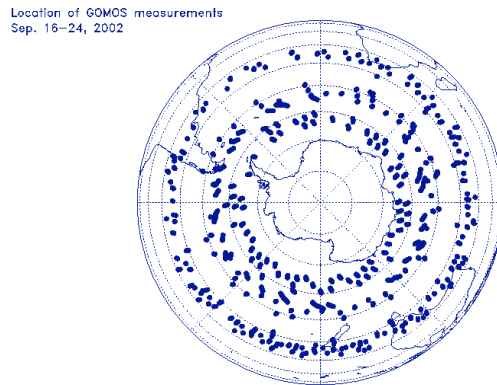


Fig. 3. Locations of GOMOS measurements during the week 16-24 September 2002. Note the absence of profiles obtained above Antarctica due to instrumental limitations.

GOMOS data were associated with an artificial relative error of 15 % (this accounts both for the measurement and the representativity errors). Furthermore, in order to reduce the sensitivity of the assimilation to oscillations in the measurements, the absolute variance used in the assimilation algorithm has been computed by multiplying the relative error of the measurements with the forecast of the model. In a first step, no data selection was applied.

Fig. 4 displays a comparison between the ozone fields at 31.4 hPa simulated by the free running model and by the model constrained with GOMOS. The difference between both simulations is obvious: the ozone hole appears deeper in the second case (values lower than 3 ppmv) and the field is more structured. It is noteworthy that both cases were initialized using a climatological distribution (without any structures in longitude) on September 16th.

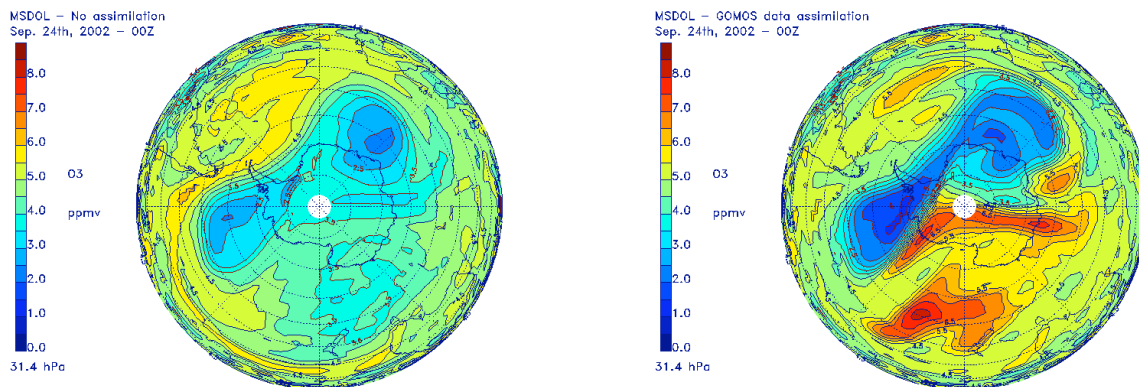


Fig. 4. Comparison between the ozone fields at 31.4 hPa simulated by the free running model (left) and the model constrained by assimilation of GOMOS data (right) on 24/09/2002.

The relative difference between both fields is shown on fig. 5. It can be seen that assimilation brought information over the whole hemisphere and particularly where there were no measurements (at latitude lower than 70S): the difference in the value of the mixing ratio indeed reaches 80% above the South Pole.

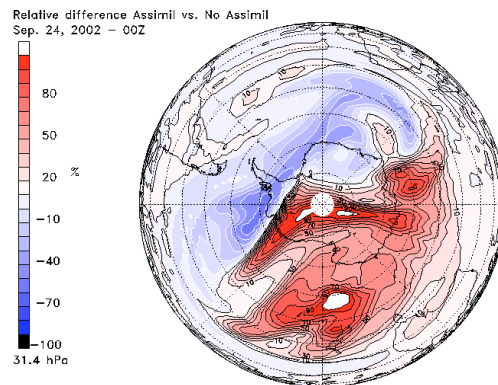


Fig. 5. Relative difference between the free running model and the constrained model at 31.4 hPa.

The time evolution of the ozone distribution on the level 31.4 hPa is shown between 21/09/2002 and 24/09/02 on fig. 6 where the onset of the vortex breaking has been correctly caught up. Whether such an ozone distribution is realistic when compared to independent measurements is the subject of the next section.

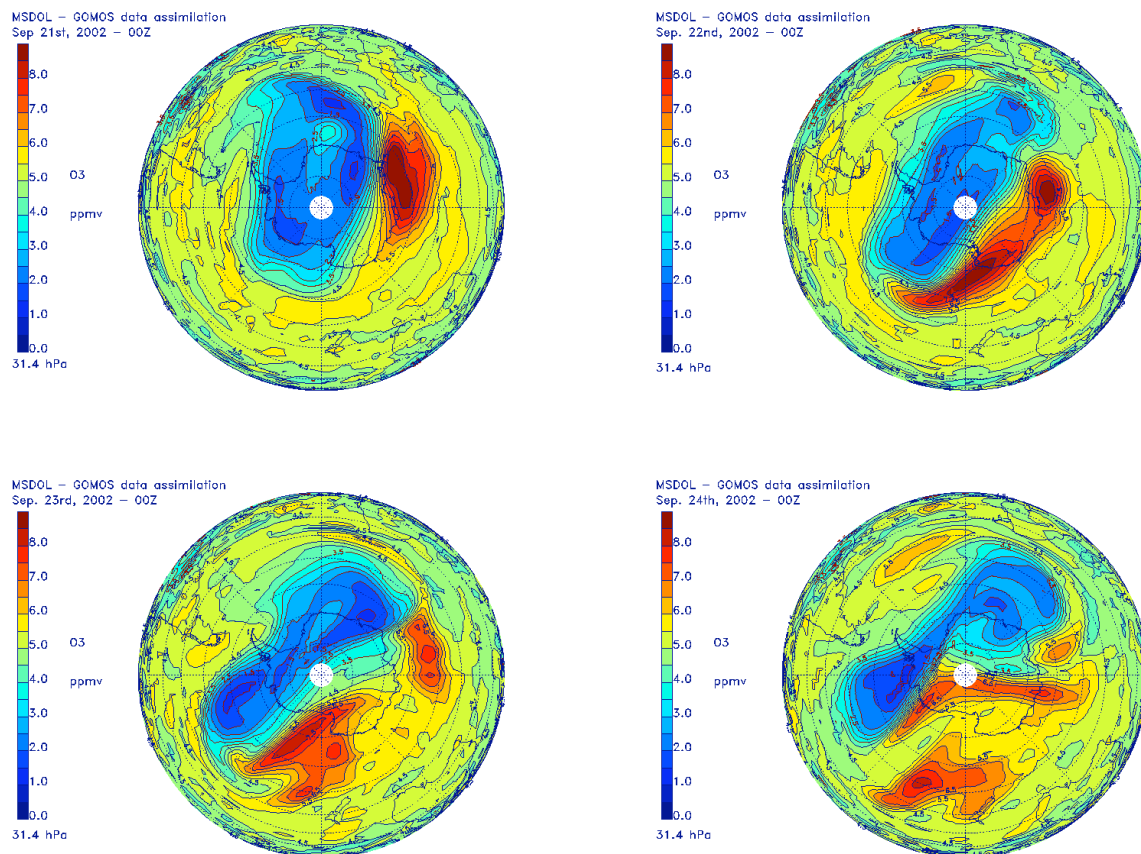


Fig. 6. Sequence of ozone fields at 31.4 hPa at the onset of the polar vortex breakup between 21/09/2002 (top left plot) and 24/09/2002 (bottom right plot).

3.2 Comparisons with external sources

In order to assess the quality of the assimilation results, we compared the outputs of the simulation with external data such as ozone soundings and Lidar measurements. Unfortunately, it was not possible to perform a thorough validation because of the lack of such data.

On fig. 7 is plotted a comparison between an ozone sounding at the South Pole (kindly provided by CMDL on the site: <http://www.cmdl.noaa.gov/ozwv/ozsondes/spo>) and two profiles extracted from the two MSDOL experiments (free running model and assimilation). One can see that assimilation clearly improves the ozone distribution in the lowermost part of the model domain, although the minimum value inside the ozone hole is overestimated. Furthermore a large bias exists at the bottom of the model, probably due to assimilation of wrong values of the mixing ratio created by the strong oscillations of the measured profile.

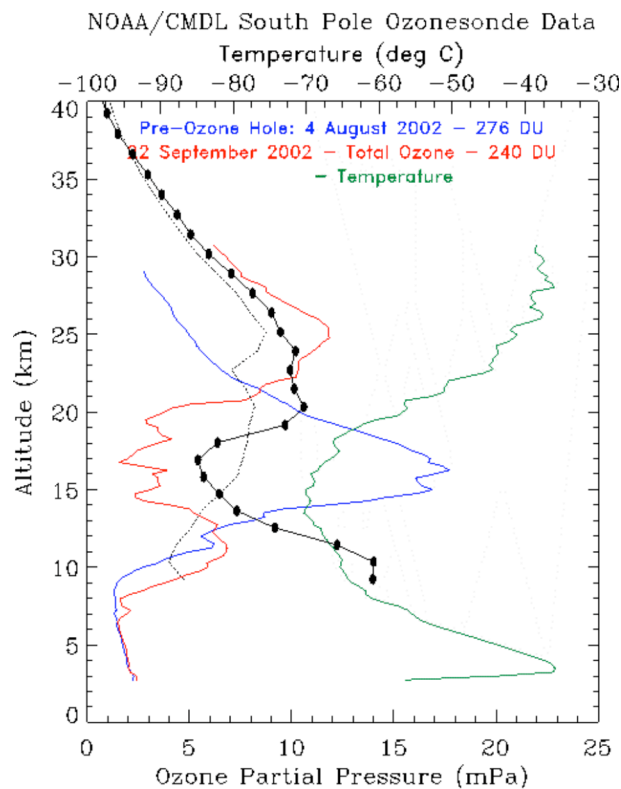


Fig. 7. Comparison ozone partial pressure vertical profile measured by an O₃ sounding over the South Pole (in red) and calculated by the assimilation model (black lines) on 22/09/2002. The solid line with dots corresponds to the simulated profile with assimilation, while the dotted one is the output of the free running model. Although assimilation appears to overestimate the ozone partial pressure over the stratosphere, the shape of the profile agrees quite well with the sounding. Also plotted are the temperature profile (in green) and the ozone profile at the beginning of summer (in blue).

A comparison of the assimilated field with two Lidar profiles obtained at Lauder on 22/09/2002 is displayed on fig. 8. Again, there is a general agreement between modeled and measured profiles even if the small structures observed in the Lidar profiles are not observed in the simulated profiles that, on the other hand, exhibit quite strong oscillations on the lowermost levels. It is noteworthy that Lauder is at the edge of the region where assimilation was performed.

These few examples tend to indicate that assimilation has been beneficial (in the sense that more realistic profile are produced) even if drawing definite conclusions only from them is not possible yet.

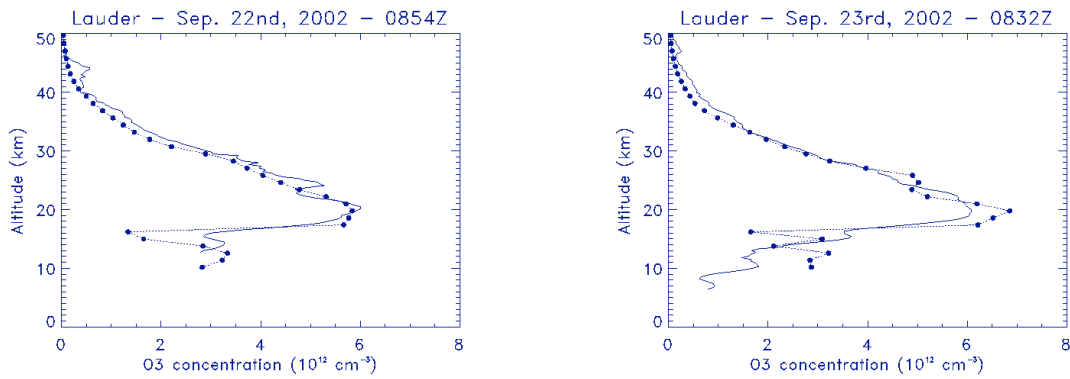


Fig. 8. Comparisons between Lidar soundings at Lauder (solid line) and the assimilated run (dotted line with points).

3.3 Assimilation diagnostics

The performances of the assimilation can be monitored by considering some simple diagnostics. An example is shown on fig. 9 where is plotted for each profile the quantity

$$\frac{\sigma_{\text{innovation}}^2}{\sigma_{\text{actual}}^2} \quad (4)$$

This scalar value can be seen as the ratio of the variance of the quantity $y - \hat{y}$, termed *innovation* or “*OmF*”, to its actual covariance σ_{actual}^2 . This ratio is equal to 1 if both the measurement and the forecast error covariance matrix are correct [2]. It can be seen on the figure that this value is in general of the order of 0.2 or 0.3 indicating that the covariances are overestimated. It seems however that there is a latitudinal effect: indeed the value of this parameter seems to be larger near the edge of the vortex.

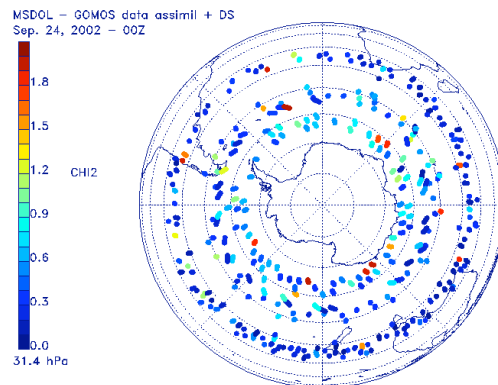


Fig. 9. Map of the ratio of the variance of the innovation to the model covariance matrix. This value is equal to 1 if the measurement and the forecast error covariance matrix are correctly estimated.

A scatter plot of the innovation and/or the quantity *OmA* (“*Observation minus Analysis*”) can also help characterizing the assimilation algorithm. On fig.10 are plotted *OmA* versus *OmF* using a color scale depending on the altitude and latitude, respectively. Ideally, all the points should lie between the bisecting line (assimilation is driven by the measurements) and the horizontal line (assimilation is driven by the model). This is globally the case, although a large spread can be observed: there is a wealth of outliers due, again, to the outliers in the measurements.

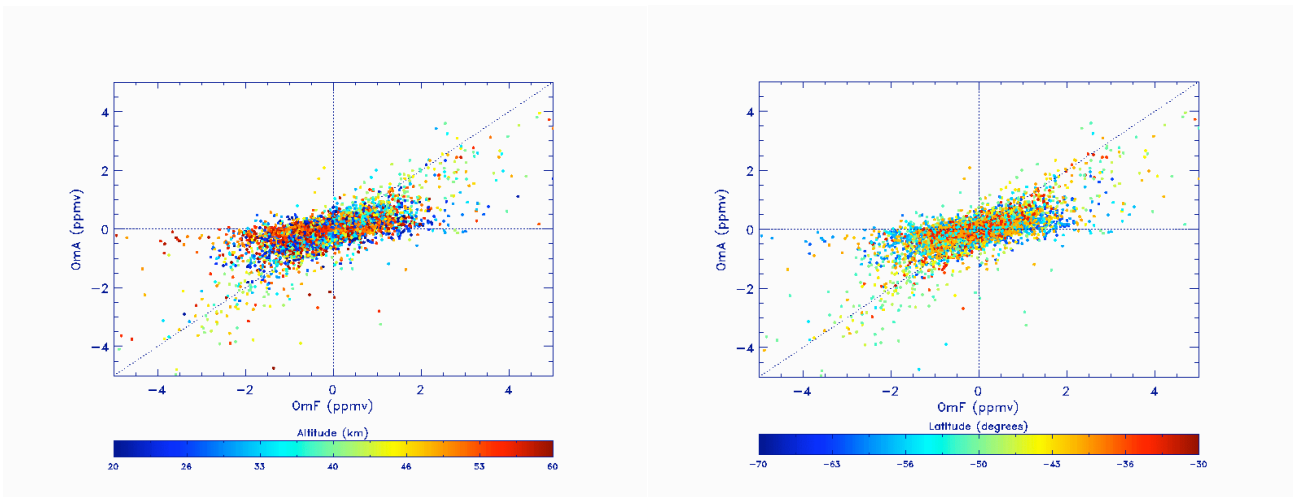


Fig.10. Scatter plots of the quantities OmA versus OmF (see text) as a function of altitude or latitude. Ideally, the points should lie between the bisecting line and the horizontal line.

The time evolution of both of these quantities can help detecting possible problems in the assimilated profiles. This is illustrated on fig. 11 that shows that:

- The distribution of the innovations seems to be globally unbiased (note that the mean is -0.048)
- There are a number of outliers corresponding to bad GOMOS measurements.

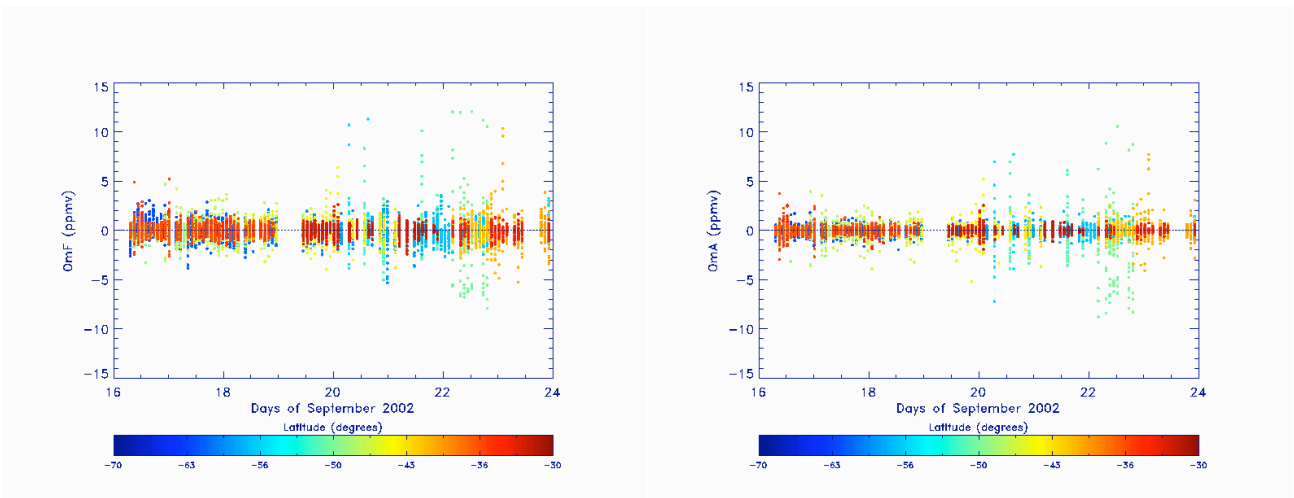


Fig. 11. Time evolution of the quantities OmF and OmA. Note the outliers corresponding to bad GOMOS profiles.

3.4 Assimilation with data selection

As an illustration of the effect of removing some measurements before assimilation, we have performed a run in which severe data selection was applied, resulting in the removal of about half the measurements on a profile. Fig. 12 show the resulting ozone field at 31.4 hPa, as well as the difference with the free running model. Surprisingly enough, the ozone hole is not as deep as in the previous case (see fig. 4 and 5). A comparison with the O₃ sounding profile at South Pole on 22/09/02 is given on fig. 13.

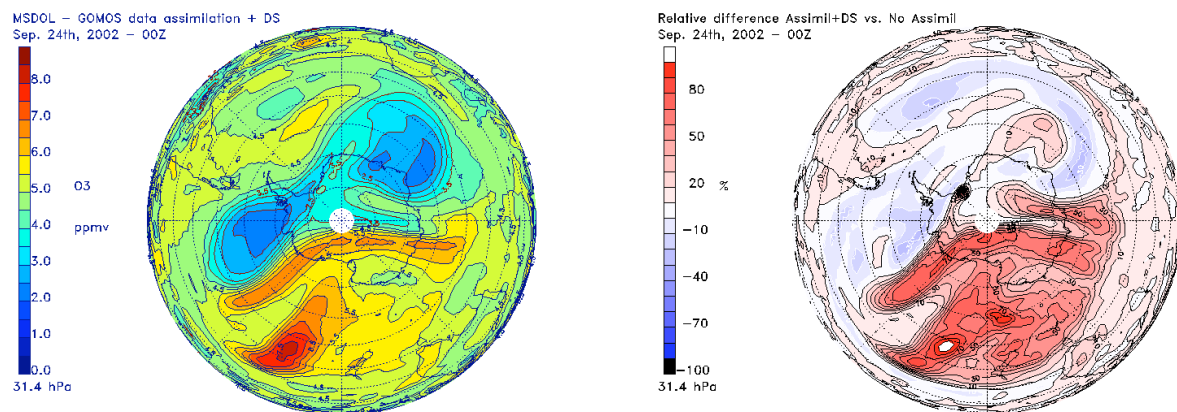


Fig. 12. Ozone distribution at 31.4 hPa on 24/09/2002 when performing assimilation with serious data selection (left) and difference with the free running model (right). These plots can be compared with those on fig. 4 and 5.

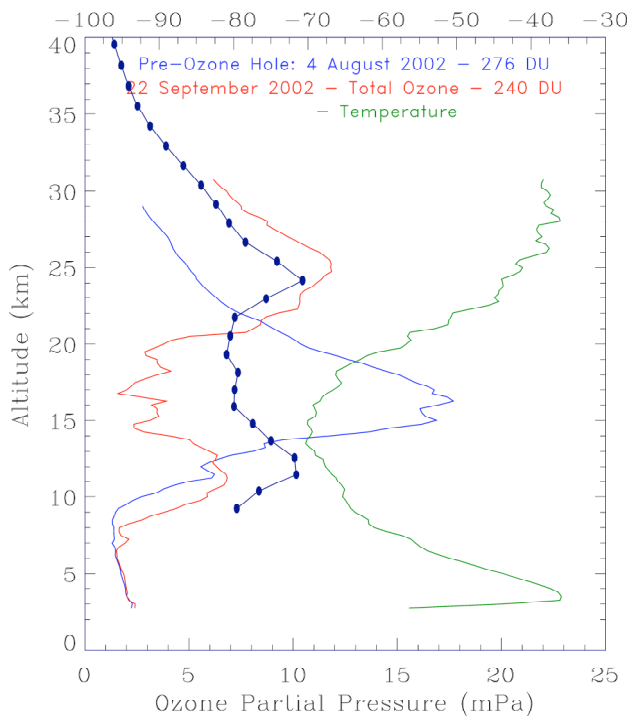


Fig. 13. Comparison between an ozone sounding at South Pole (in red) and the simulation by the model (solid line with dots). This plot can be compared to the result shown on fig. 7: data selection improves the retrieval of the profile in the lowermost layers of the model, even if the ozone partial pressures calculated from the model in the lower stratosphere are still larger than measured values.

3.5 Discussion on the use of data assimilation for instrument validation

Data assimilation has proven a powerful tool for scientific studies of the physics of the atmosphere. Application of this technique to atmospheric chemistry is however more recent. In particular, it has never been used in the frame of validating an instrument.

When the OZVAL project has been proposed, the primary goal was to assimilate the unvalidated ozone measurements. The idea behind this proposal was that data assimilation would be used to interpolate ENVISAT ozone measurements in order to ease the comparison with external data. The assimilated field would then be used to compute what would have been measured by one of the ENVISAT instrument at any place and time, in particular those from independent measurements. It is however difficult to estimate the amount of external information brought by merging measurements and model forecasts.

Another possibility is then to assimilate independent (validated) measurements into the models such as to produce a reference atmosphere as accurate as possible. Comparisons between ENVISAT and this reference is then possible. Furthermore, studying the difference between the ENVISAT observations and the forecast by such a model and/or the analyses can give insight on the forecast skill of the model as well as on the statistics of the observation errors. As an illustration, the distribution of these differences is shown on fig. 14 and 15 for the whole period of the run with data selection. Note that the standard deviation of the distribution of the difference between observations and analyses is significantly lower than the one of the difference between observation and forecast, as expected. Some studies have been devoted to extracting quantitative information about the observation and model errors from these quantities [3, 4]. It is planned to assess the use of such techniques.

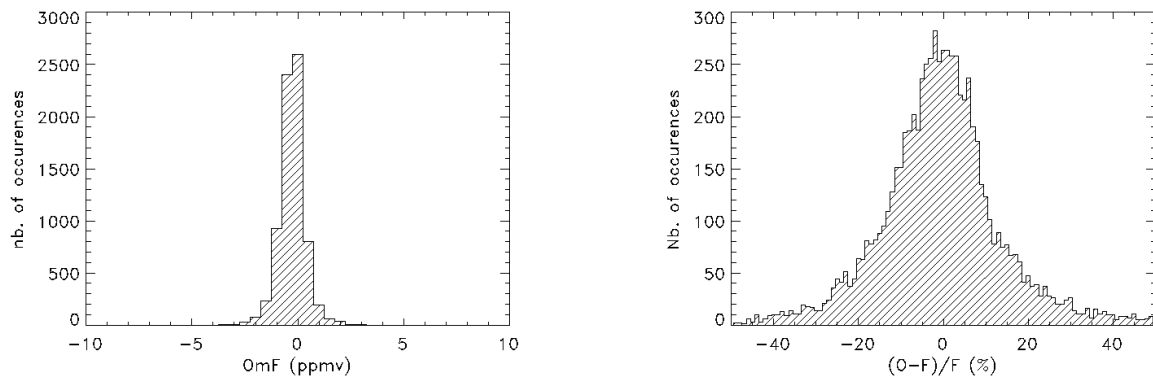


Fig. 14. Histogram of the values of the innovations in absolute (left) and relative values (right). This can be used to estimate the forecast skill of the model.

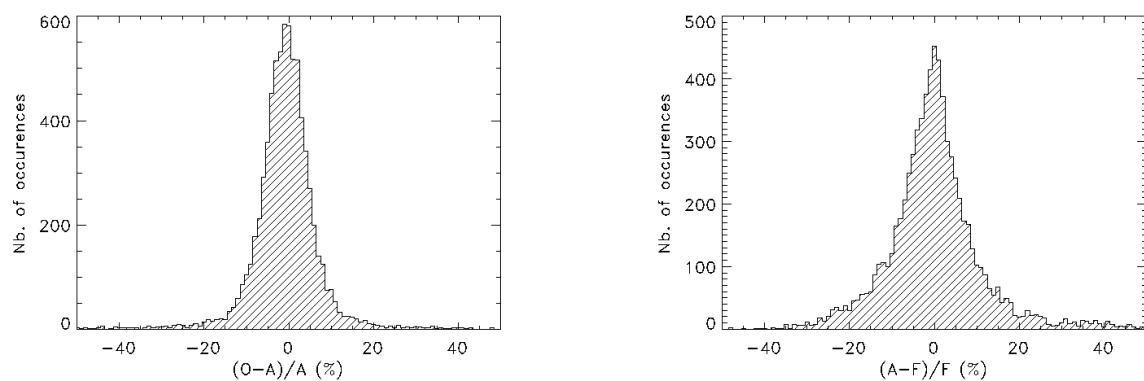


Fig. 15. Histograms of the relative difference between observations and analyses $(O-A)/A$ and of the relative difference between analyses and forecasts $(A-F)/F$. The former basically measures the analyses bias while the latter may be used to estimate the model bias.

4. SUMMARY

In the frame of the Ozval project, we have performed assimilation experiments of GOMOS data. We focused on late September 2002 when the polar vortex split. These experiments showed that assimilation of GOMOS data is possible and gives good results (in good agreement with external data) providing that careful data selection is applied and that the error bar of the measurements is fixed and not taken from the GOMOS products as the latter seem to be underestimated at the time of the end of the commissioning phase. A strong recommendation to the GOMOS team is thus to revised the estimation of the error bar and also to adapt the resolution to reflect the actual accuracy of the measurements.

The promising results obtained incite us to go further by:

- Assimilating data over upcoming validation campaigns (as well as reprocessing some periods when updated L2 products will become available), allowing further comparison with external data
- Developing the assimilation of GOMOS “line densities” in order to overcome the problem of oscillations in the profiles.
- Using the assimilation residuals (O-F, O-A, A-F) to characterize the observation and model errors.
- Assimilating MIPAS data (assimilation of SCIAMACHY column densities is hampered by the fact that the MSDOL model domain does not contain the troposphere).

5. REFERENCES

1. Rose K. and G. Brasseur. A three-dimensional model of chemically active trace species in the middle atmosphere during disturbed winter conditions. *J. Geophys. Res.* **94**, 16387-16403, 1989.
2. Dee D. On-line estimation of error covariance parameters dor atmospheric data assimilation. *Mon. Wea. Rev.* **123**, 1128-114, 1995.
3. Daley R. Estmating observation error statistics for atmospheric data assimilation. *Ann. Geophys.* **11**, 634-647, 1993.
4. Dee D, *et al.* Maximum-likelihood estimation of the forecast and observation error parameters. Part I: methodology. *Mon. Wea. Rev.* **127**, 1822-1834, 1999.

6. ACKNOWLEDGMENTS

This work has been supported by the French space agency (CNES) as the project Ozval. We would like to thank Philippe Ricaud from Observatoire de Bordeaux for fruitful discussions and for providing us with Odin/SMR data used, in particular, for fig. 1.