Improving global chemical simulations with variational assimilation of GOME data

S. Massart
Centre Européen de Recherche et Formation Avancée en Calcul Scientifique, 42, Avenue Gaspard Coriolis, 31057 Toulouse Cedex 01, France

Abstract

Data assimilation, which plays an important role in the analysis of atmospheric data, in particular in Numerical Weather Prediction (NWP), is increasingly being used to analyze photochemical data. This paper presents how MOCAGE CTM simulations of global ozone have been improved using a variational assimilation scheme. The analysis of the retrieved ozone profiles from the GOME nadir-viewing spectrometer increases the quality of MOCAGE forecasts in terms of ozone profiles and total columns.

1 Introduction

Since the evidence of the hole in the ozone layer over Antarctica presented by British Antarctic Survey in 1985, ozone has became one of the most important trace species in the atmosphere, and its study has generated wide scientific and public interest. Because of its role in the radiation budget, in the tropospheric pollution and its interaction with other meteorological variables, understanding the past and present states of the atmospheric ozone field is critical for the prediction of future ozone changes and the consequent impact on the Earth’s environment. Currently, several Chemical Transport Models (CTMs) have been developed to calculate the chemical evolution of the atmosphere with a dynamical evolution usually taken from Global Circulation Models (GCMs). With the increasing number of space-based remote sensing instruments designed to provide global data sets of atmospheric ozone, models can be validated by comparing observations and simulations. An improving step for validated models consist of using the data
in order to improve model results. These techniques are well known as data assimilation.

Data assimilation experiments have been performed at Météo-France with the MOCAGE CTM using sequential assimilation of MOZAIC data [2]. Measurements of OZone and wAter vapour by alrbus in-service airCraft (MOZAIC) is a European program which produces world wide upper-air ozone and water vapour measurements. A variational assimilation scheme developed in the framework of the ASSET (ASSimilation of EnvisaT data) EC FP5 project is now running with MOCAGE. ENVISAT was launched in March 2002 and data from its suite of instruments are available. We have initially decided to assimilate observations from another instrument, GOME because of all the expert on ozone profile products that includes averaging kernels.

2 Assimilation system

In order to represent global ozone as accurate as possible, we have chosen to assimilate ozone profiles into the MOCAGE Chemical Transport Model. It first consists to integrate an initial condition for global chemical fields through MOCAGE and then compute the misfit, i.e. the difference between each measurements and the model results. The aim of data assimilation is to minimize this distance with the constraint of a priori information. The scheme used for the assimilation of ozone profiles is a variational scheme implemented with the PALM software.

2.1 MOCAGE Chemical Transport Model

MOCAGE is the new Météo-France global three-dimensional CTM dedicated to the numerical simulation of interactions between dynamical, physical and chemical processes in the lower stratosphere and the troposphere [7]. MOCAGE is a semi-lagrangian grid-point model forced each 3 hours by Météo-France General Circulation Model (GCM) analyzed winds. With four levels of nested domains, the horizontal resolution can be increased from the 2° global resolution down to 10 km. The model includes 47 levels from the surface up to 5 hPa. The vertical resolution is about 800m in the vicinity of the tropopause and the lower stratosphere.

The configuration of MOCAGE chosen for this study is the global version with a 2° horizontal resolution and the RACM [9] chemical scheme for troposphere and REPROBUS [5] for the stratosphere. In this form, MOCAGE was validated (and still going-on) at Météo-France and is currently used for a range of applications in the field of chemical weather forecasts.

2.2 GOME data

The Global Ozone Monitoring Experiment is a nadir-viewing spectrometer that was launched on April 1995 on board the second European Remote Sensing Satellite (ERS-2). This instrument measures the solar radiation scattered by the atmosphere in the ultraviolet and visible spectral regions that permits to retrieve global
ozone profiles. In this study, the ozone profiles are retrieved with the algorithm of the Royal Netherlands Meteorological Institute (KNMI), which uses the optimal estimation method based on a priori knowledge and so-called averaging kernels [6].

There are three pixels per scan along the ERS-2 orbit track, each scan spanning approximately 960 km. The assimilation system assumes that each observation represents a point measurement located at the center of the corresponding pixel. In a one day period approximately, 900 profiles × 22 points by profile are available, giving a total number of observations close to 20 000 (see figure 1).

![Figure 1: ERS-2 orbit track on March 1st 2000 (crosses represent the center of the scanned pixel) and retrieved profile at 10:20 UT (ERS-2 latitude-longitude location is 42°N, 10°E).](image)

2.3 Variational formulation

The strategy we have chosen to improve the simulation of global ozone computed by MOCAGE is to use variational assimilation that take into account all the available information in order to estimate as accurate as possible the state of the chemical atmospheric flow. Given a set of observations \( \{ y_i, i \in [1, M] \} \) over a time interval \([t_1, t_M]\) also referred to as assimilation windows, the first step is to compute for each time \( t_i \) the distance \( d_i \) between the observations and the solution \( x_i \) of the model,

\[
d_i = y_i - H_i (x_i) .
\]

The \( H_i \) observation operator transforms a state computed by the model to an equivalent of observation. In this study, \( x_i \) represent the global ozone fields at time \( t_i \) and \( H_i \) is a bi-cubic space interpolation. The time integration between time \( t_i \) and
time $t_0$ is realized by MOCAGE from the initial condition $x_0$. In this case, the misfit (1) can be viewed as a function of $x_0$.

Talagrand [10] suggested to add another distance which measures the difference between $x_0$ and a prior estimate (or background) $x^b$ of $x_0$. The latter can come from previous MOCAGE simulations. The 3D-FGAT (First Guess at Appropriated Time) consists in finding the best increment $\delta x$ that can be added to background $x^b$ in order to minimize simultaneously all the distances (1) under the constraint of both having the lower increment and the hypothesis of a persistent model [3]. This induces the construction of a scalar function $J$ of $\delta x$,

$$J(\delta x) = \frac{1}{2} \delta x^T \cdot B^{-1} \cdot \delta x + \frac{1}{2} \sum_{i=0}^{M} \left( H_i \delta x - d_i^b \right)^T \cdot R^{-1} \cdot \left( H_i \delta x - d_i^b \right) T \cdot R^{-1} \cdot \left( H_i \delta x - d_i^b \right).$$

(2)

The misfit or innovation $d_i^b$ is the distance between observations and the MOCAGE time integration $M_{i,0}$ of the prior estimate

$$d_i^b = y_i - H_i \circ M_{i,0} \left( x^b \right).$$

(3)

The scalar products are built on correlations of background and observations error, respectively $B$ and $R$ that are important elements of a data assimilation system. Modelling $B$ uses a generalized diffusion equation [11] and $R$ derives from GOME averaging kernel.

Thus, the task of variational analysis is to adjust the increment in order to minimize $J$. Then the analysed increment is added to the background $x^b$ to obtain a new initial state.

2.4 PALM software

The data assimilation algorithm can be viewed as a sequence of elementary operations or elementary components that can exchange data [4]. The PALM software (www.cerfacs.fr/~palm) we have used to build our 3D-FGAT algorithm is based on this paradigm. It aims to implement a general tool in a flexible and extensible way allowing to easily integrate high performance computing applications. The main features of PALM are the dynamic launching of the coupled components, the full independence of the components from the application algorithm, the parallel data exchanges with redistribution and the separation of the physics from the algebraic manipulations performed by the PALM algebraic toolbox [8].

The choice of PALM was done because of the minor modifications to introduce in our source codes and because of the graphical interface that allows us to simply describe algorithms in a parallel way without any expert skill on this topic.

2.5 Assimilation strategy

The main difference between the 3D-FGAT data assimilation scheme and more complex schemes such as 4D-VAR, is the fact that innovation (3), calculated with
the background fields at the correct time, are assumed constant in time and space (within the time window and the advection scale). This means that the assimilation windows have to be short enough compared to the dynamical scales. In this study, the global assimilation window is 1 day, divided into sub-windows of 3 hours. The analysis step is done in the first sub-window and the analyzed increment is added all along the sub-window with a gaussian time evolution [1]. A 3 hours forecast is used to compute the innovation at the correct time for the analysis in the next window. This operation is repeated until the end of the assimilation day.

3 Results

A direct simulation of 8 days with MOCAGE was done as “run without assimilation” (or free run) from March 1st, 2000 with chemical initial conditions based on climatology and forcings from Météo-France GCM Arpege. The assimilation strategy was carried out for each day based on the previous day analysis, except for the first one for which chemical initial conditions were the same than the run without assimilation. The further results present the difference on ozone fields between a 1 day forecast after assimilation (there is no assimilation for this forecast day) and the free run.

![Figure 2](image_url)

**Figure 2:** Difference in parts per million by mass (ppmm), between the GOME ozone observation and the co-located MOCAGE ozone forecast with assimilation (line) and without assimilation (dotted line) for 2 pressure levels : (a) 100 hPa, (b) 200 hPa. For each forecast day over the period March 2nd-8th, 2000, the difference is averaged on ERS-2 tracks.
Figure 3: Root Mean Square difference, in parts per million by mass (ppmm), between the GOME ozone profile observation and the MOCAGE ozone profile forecast with assimilation (line) and without assimilation (dotted line), averaged on ERS-2 tracks, over the period March 2nd-8th, 2000.

Figure 4: GOME ozone observations versus MOCAGE ozone forecast for 2 forecast days: March 2nd, 2000, before (a) and after (b) assimilation of March 1st, 2000 GOME data; March 8th, 2000, before (c) and after (d) assimilation of March 1st to March 7th, 2000 GOME data. The black line represents the perfect line (i.e. without error on model nor observations). The dotted line represents the trendline.
Figure 5: Latitude-longitude distribution of total column ozone (DU) on March 2nd, 2000 at 0 UT forecast by a free MOCAGE run (a) and (b) a MOCAGE run after assimilation of March 1st, 2000 GOME data. (c) Difference (%) between (a) and (b). (d) Total ozone column measured by TOMS at local noon on March 2nd.

3.1 Impact on MOCAGE ozone representation

Let first compare the GOME observations and the MOCAGE forecast ozone values with and without assimilation at a given pressure level for each forecast day. Because of the great number of measurements the difference is meant along each ERS-2 track from north to south. Therefore, we have around 12 values per days, which are plotted on figure 2 for the pressure levels 100 hPa and 200 hPa. This figure first shows that the difference between ozone observations and MOCAGE forecast is lower after assimilation for the two pressure levels. Moreover, the average of the difference is closer to zero after assimilation, which means that the model bias has been reduced.

Nevertheless, we have to remark that during few hours of the March 4th, the free run forecast is better than the run with assimilation.

In order to have the impact of assimilation all over ozone profiles we have compute the RMS (Root Mean Square) difference between the GOME ozone profile observation and the MOCAGE ozone profile forecast (see figure 3). The conclusion remains the same as the later two profiles. Figure 4 also suggest that the behavior of assimilation do not depend on ozone concentration (and as a consequence do not depend on pressure levels).
3.2 Impact on ozone total columns

Another way to view the impact of data assimilation on ozone forecast is to compute ozone total columns after and before assimilation. Figures 5 and 6 show the global distribution of ozone total columns, respectively for March 2nd and March 8th, 2000, i.e. the beginning and the final assimilation days. The comparison with TOMS observations which are not assimilated, points out the improvement due to assimilation even for the first forecast day. In particular, the free run simulated ozone over Europe was not so good with strong values. The assimilation has reduced these values and ozone concentrations are closed to TOMS observations.

4 Conclusions and future research

In this paper, we have presented results from a variational data assimilation system based on the Météo-France CTM. GOME ozone profiles have been assimilated over a period of one week on March 2000. The paper suggests that MOCAGE global ozone one-day forecast is improved with assimilation except for few hours of March 4th. This case still needs explanations.

The assimilation scheme will be soon extended in several directions. Being one
part of ASSET EC project, the next steps of this work will be the assimilation of ENVISAT data. This means that the assimilation scheme has to be able to analyze both ozone profiles and total columns.

Moreover, atmospheric chemistry and dynamics are strongly linked in many respects and influence each other. The development of coupled CTM-GCM data assimilation systems is therefore the logical future step.

References


