Progress in the investigation of problems in the operational 4d-var assimilation at Météo-France

or

" the grey side of 4d-var..."

Dominique Giard . *Météo-France/CNRM/GMAP*

A. Introduction

IFS, ARPEGE and ALADIN share basically the same 4d-var code. However the success in operational implementations is quite different from one configuration to another. For IFS, addressing large scales and medium-range forecasts with the help of powerful computers, the situation looks smiling. For ALADIN, when small scales and short-range forecasts (up to 48h) are considered, 4d-var is still a long-term, and maybe too expensive, goal. ARPEGE is just in-between, for horizontal and time scales, but also for problems. 4d-var assimilation is operational since June 2000, but doesn't prove fully better than 3d-var (as proved by a parallel assimilation experiment covering last winter) and clear deficiencies in the management of humidity (at least) are still to be cured. This paper summarizes the recent investigations of Radi Ajjaji (DMN, Maroc), François Bouyssel and Cécile Loo (Météo-France).

B. Main features of the operational 4d-var assimilation for ARPEGE

The main differences with IFS are :

- a multi-incremental approach
- a generalized use of digital filter (Jc-dfi and incremental semi-external digital filter initialization)
- stretching and tilting at high resolution (as for forecast ...)
- the physics (full or simplified), the observation operators, ...
- a time-window of 6h

The present configuration is the following :

ä High resolution :

- spectral truncation $T_L298c3.5$ *(Δx= 19 / 67 / 235 km, Δt= 830.77 s, ≡ T199c3.5)*
- semi-external incremental initialization (based on finalization)
- 3 steps :
  1. 6h forecast + computation of "obs-guess" + quality control + filtering
  2. update after the first minimization step + 6h forecast +computation of "obs-guess"
  3. update after the second minimization step + 6h forecast + filtering + surface analysis

ä Low resolution :

- spectral truncation T_L161 *(Δx= 125 km, Δt= 1800. s, ≡ T107)*
- 6h forecast and minimization (second step)
- Jc-dfi as a weak constraint
- full simplified physics (vertical diffusion, mesospheric drag, gravity-wave drag, large-scale precipitations, radiation, convection)
- 25 iterations

ä Very low resolution :

- spectral truncation T_L107 *(Δx= 189 km, Δt= 1800. s, ≡ T71)*
- 6h forecast and minimization (first step)
- Jc-dfi as a weak constraint
- dry vertical diffusion (Buizza)
- 50 iterations

The successive operational changes since the launching in June 2000 have been the following :

August 2000

Stronger filtering *(increased weight for q in Jc-dfi and incremental semi-external dfi)*

May 2001

Correct use of ATOVS *(correction of a bug in the observation operator)*

September 2001

Move to ODB

"*Rv=Rd*" trick

Less horizontal diffusion at low resolution *(to reduce incrementality)*

Mesospheric drag in simplified physics

December 2001

Bug correction for horizontal diffusion *(far too weak due to a namelist mistake, leading to an explosion in the T95 forecast on 10/10/2001 !)*

January 2002

Increased resolution :

vertical *(L31* $\rightarrow$ *L41)* and spectral *(T42 /63 /95 /199c3.5* $\rightarrow$ T_L*107 /161 /298c3.5)*

Less incrementality *(most inner loop suppressed, considering the ratio extra-cost / efficiency)*

Semi-Lagrangian advection for low resolutions *(esp. TL/AD models)*

The next step should be the improvement of the vertical diffusion scheme, simultaneously in full and in simplified physics.

C. Investigating problems

The main problem linked to 4d-var is the "Saharian precipitations" symptom, i.e. spurious precipitations over desert areas, as shown in Figure 1. This feature was present with the previous 3d-var assimilation, but significantly amplified by the change to 4d-var, from the very beginning. Some other problems appeared more recently, such as the "24 April 2001" case. At 18 h UTC, the 4d-var assimilation introduced a strong drying over a large area, from Spain to Germany, not justified at all by observations. Six hours later, the reverse correction was introduced, going back to reasonable fields.

In-depth investigations were performed on both situations. For the "Saharian precipitations" case, numerous 10-day assimilation suites and forecasts were performed for the period 1-10 June 2000, testing the impact of :

ä simplified physics,

ä multi-incremental / incremental / non-incremental approach,

ä resolutions,

ä *"Rv=Rd*" trick (described in part D),

...

The main results are summarized by Figure 2.

On the "24 April 2001" case, a detailed investigation of the impact of observations led to the following remarks :

ä no TEMP is available within 300 km of the maximum increment (missing observation at that time);

ä the anomaly is suppressed by adding an artificial TEMP observation (derived from the forecast);

ä suppressing AIREPs and / or SYNOPs suppress the anomaly;

ä 2 AIREP around Baléares are at the origin of most of the problem;

ä the problem on *q* is replaced by a problem on *T* if suppressing the temperature observation in the 2 AIREPs.

Some other tests were performed, such as : *"Rv=Rd*" trick , *"Rv=Rd*" trick but not in the observation operators, impact of horizontal diffusion, ... <u>The evolution of the problem along all the steps of 4d-var was carefully examined.</u>

This led to the following diagnostic. The multivariate assimilation tries to fit observations by adjusting ***q*** first, even with very strong increments, especially when there are very few observations of humidity to constraint it. This happens since the very first integrations of the TL and AD models, at the lowest resolution, and increases regularly along the minimization steps.

D. A temporary solution : the "*Rv*=*Rd*" trick

The linearized equations read :

$$\partial_t V' + V \cdot \nabla V' = \ldots - R\,T'\,\nabla(\ln p) - \frac{RT}{p}\nabla p' - T(R_v - R_d)\,q'\,\nabla(\ln p)$$

$$\partial_t T' + V \cdot \nabla T' = \ldots + \frac{A}{T}T' + \frac{A}{R_d}(R_v - R_d)\,q' - \frac{A}{C_{pv}}(C_{pv} - C_{pd})\,q'$$

$$\partial_t q' + V \cdot \nabla q' = 0 \qquad \text{with } R = R_d + (R_v - R_d)\,q$$

Considering the adjoint of this system, it is clear than specific humidity (*q*) will be modified to fit wind observations, unless *Rv* and *Rd* keep identical. The suggested trick was to set artificially *Rv* to *Rd* in the TL and AD models. This was very efficient in reducing problems, as illustrated by Figures 1 and 2. But that is not a long-term solution, since :

- not scientifically based
- setting *Rv*=*Rd* everywhere may be not necessary and even detrimental
  - ä useless for observation operators + maybe problems for wind
  - ä dangerous in equations for *T*, because of the balance with (*Cpv*-*Cpd*)
- not only *q* is concerned !

E. Other investigations : comparison between forecasts at high and low resolution

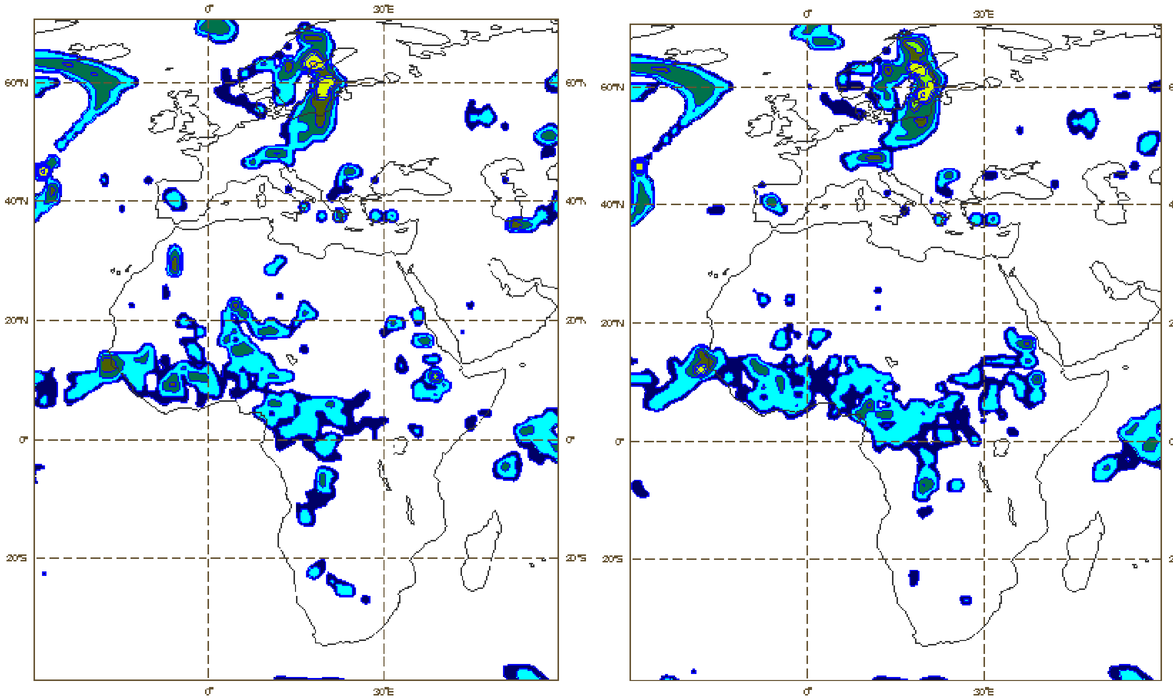
This study was motivated by problems in the convergence of minimization at resolution T_L161 (intermediate one). Thus 6 h forecasts at resolutions T_L298 and T_L161 were carefully compared, which led to the following points :

- ä there is a strong increase of moisture in the lowest levels at low resolution
- ä this is due to the combination of no change in vertical diffusion and a very long time-step (1800s), with additional contributions of the "shallow convection" and "anti-fibrillation" schemes

The retained solutions are :

- ä an additional "smoothing" in shallow convection
- ä a retuning of the anti-fibrillation scheme (+ corrections)
- ä an increased consistency between the full and simplified parameterizations of vertical diffusion

Figure 1 : The "Saharian precipitations" symptom : impact of the "*Rv*=*Rd*" trick
Cumulated (6h) precipitations in the ARPEGE forecast starting on 27/08/2001 at 00h UTC, valid at 18h UTC. Left : operational suite (without). Right : parallel suite (with)



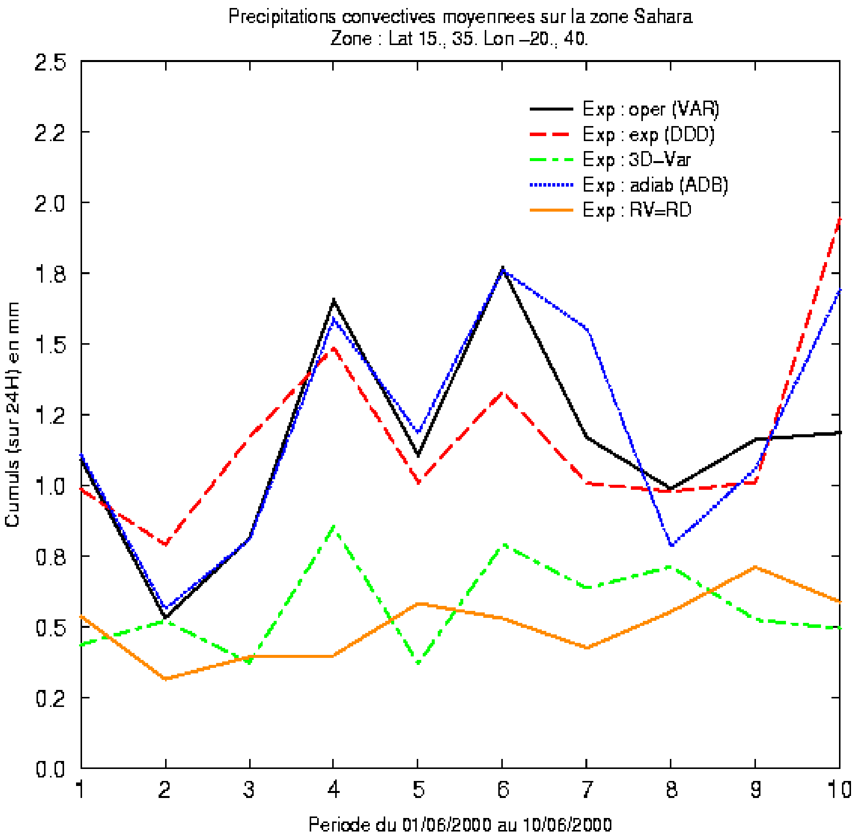


Figure 2 : The "Saharian precipitations" symptom : impact of a few modifications on convective precipitations
Average of convective precipitations over Sahara, along a 10-days assimilation experiments
black: initial 4d-var; red: non-incremental 4d-var (T95); green: 3d-var; blue: adiabatic 4d-var; orange : new 4d-var ("Rv=Rd" trick)