Preliminary experiments toward the pre-operational ALADIN 3DVar

by Thibaut Montmerle, Claude Fischer and Loïk Berre
Météo-France/CNRM/GMAP
24/02/2005

1. Introduction

This report presents the aim, strategy and first results for the implementation of an operational variational data assimilation system in the Aladin/France model. Starting from the experience gathered over the last five years with 3D-VAR in Aladin as a research configuration, the goal is to settle a continuous, permanent assimilation cycle with an initial update frequency of 6 hours. Arpège data plus extra data (Meteosat-8/SEVIRI radiances to start with) are considered. The initial choices have been made by performing two test periods of 1 month and 15 days. A discussion concerning the choice of the background error covariance matrix will be presented. Emphasis will furthermore be made on the specific use of SEVIRI data and on their impact on precipitation forecasts at mesoscale.

The goal is to obtain at least as good conventional scores over Western Europe as with Arpège, plus a beneficial effect on short-range wind, temperature and precipitation forecasts. The retained solution offers such an improvement, with better precipitation "scores" at least up to 12 hours. In 2005, this assimilation cycle will be improved by more frequent updates and additional algorithmic facilities.

2. Choice of the background error covariance matrix

2.1 Theoretical aspects

2.1.1 3D-Var formalism

The formalism of the variational analysis follows closely the work in ARPEGE/IFS (Courtier et. al, 1994).

The analysis x^a represents the atmospheric state which is the best fit between the background x^b (usually taken as a 6 hour forecast) and the available observations stored in the y vector. The cost function is written with an incremental formulation :

$$J(x) = \frac{1}{2}\delta x^T \mathbf{B}^{-1}\delta x + \frac{1}{2}\left(\mathbf{H}\delta x - d\right)^T \mathbf{R}^{-1}\left(\mathbf{H}\delta x - d\right) \qquad (1)$$

Where δx is the increment defined by the difference between x^a and x^b; **B** is the background covariance error matrix defined in the next section; **R** is the observation covariance error matrix composed of instrumental errors and errors in the observation operator H . It is usually taken as a block diagonal matrix since observations are supposed to be uncorrelated horizontally (values for SEVIRI are given in section 3.2); d is the innovation vector which represents the departure between the observations and the background model state interpolated in observation space :

$d = y - H(x^b)$

H being the observation operator that represents the model state in the observation space. The latter operator can be strongly non linear and contains for satellite radiances i) the fast radiative transfer model RTTOV-7 (Saunders and Matricardi, 1999), which allows to retrieve the brightness temperature T_b from surface pressure and temperature and from vertical profiles of temperature, humidity and ozone, ii) a horizontal interpolation operator to position the control variable profiles on observation locations, iii) vertical interpolation and extrapolation operators allowing to position these profiles on RTTOV vertical levels. **H** is the tangent linear operator of H in the vicinity of the background state x^b.

The variational problem is solved by calculating iteratively the cost function (1) and its gradient:

$$\nabla_{\delta x} J = \left(\mathbf{B^{-1}} + \mathbf{H^T R^{-1} H}\right)\delta x - \mathbf{H^T R^{-1}} d$$

Convergence is obtained after $\| \nabla_{\delta x} J \|$ reaches a fixed minimum.

2.1.2 The B matrix

As showed in Eq. (1), the background error covariance matrix **B** filters and propagates the information given by the increment. Thus, it plays a key role in data assimilation and needs to be carefully chosen. Berre (2000) has proposed for ALADIN a new multivariate formalism adapted from Parrish et. al (1997) and Derber and Bouttier (1999) for global NWP systems. This formalism uses linear balance relationships between errors of different physical quantities computed from statistical regression, with an extra balance relation for specific humidity. The use of regressions allows to obtain balance relationships which are representative of the area of interest. Thus, this method seems well suited for assimilation purposes at mesoscale in any domain. The statistical relations read :

$$\begin{aligned}
\zeta &= \zeta \\
\eta &= \mathcal{M}\mathcal{H}\zeta + \eta_u \\
(T,P_s) &= \mathcal{N}\mathcal{H}\zeta + \mathcal{P}\eta_u + (T,P_s)_u \\
q &= \mathcal{Q}\mathcal{H}\zeta + \mathcal{R}\eta_u + \mathcal{S}(T,P_s)_u + q_u
\end{aligned}$$

where (ζ, η, (T, P_s), q) are forecast errors of vorticity, divergence, temperature and surface pressure, and specific humidity on model vertical levels ; the subscript u stands for unbalanced (total minus balanced) fields.

$\mathcal{M}$, $\mathcal{N}$, $\mathcal{P}$, $\mathcal{Q}$, $\mathcal{R}$ and $\mathcal{S}$ are vertical balance operators relating spectral vertical profiles of predictors and those of the predictands. They are deduced from homogeneous and isotropic auto-covariance and cross-covariance matrices (the detailed formulation and the impact of this assumption on the spectral covariance formulation can be found in Berre (2000)). $\mathcal{H}$ is a horizontal balance operator obtained by linear regression that transforms spectral coefficients of vorticity ζ into those of balanced geopotential P_b. Balanced geopotential is supposed to be the balanced part of P_t, which is the linearized mass variable deduced from (T, P_s) by the linearized hydrostatic relationship (Parrish et al. 1997). These regression coefficients can be related to the percentage of explained variances of each total field by its predictors and to the strength of the statistical link.

Separating predictors into balanced and unbalanced parts allows to provide an independent set of variables better suited for the 3DVar analysis. The background error covariance matrix $\mathbf{B_u}$ in the predictor space is indeed taken as block diagonal, each block consisting of the vertical error covariances for each predictor and for each wavenumber :

$$\mathbf{B_u} = \begin{pmatrix} \mathbf{C}\varsigma & 0 & 0 & 0 \\ 0 & \mathbf{C}\eta_u & 0 & 0 \\ 0 & 0 & \mathbf{C}_{(T,Ps)_u} & 0 \\ 0 & 0 & 0 & \mathbf{C}q_u \end{pmatrix}$$

The total background error covariance matrix **B** is then retrieved in the analysis space :

$$\mathbf{B} = \mathbf{K}\mathbf{B_u}\mathbf{K}^{\mathrm{T}}$$

With: $$\mathbf{K} = \begin{pmatrix} \mathbf{I} & 0 & 0 & 0 \\ \mathcal{M}\mathcal{H} & \mathbf{I} & 0 & 0 \\ \mathcal{N}\mathcal{H} & \mathcal{P} & \mathbf{I} & 0 \\ \mathcal{Q}\mathcal{H} & \mathcal{R} & \mathcal{S} & \mathbf{I} \end{pmatrix}$$

where the subscript *T* stands for the transpose and **I** is the identity.

2.2 Choice of the B matrix

Three different methods have been used to estimate the regression coefficients defining the background error covariances :

- The standard NMC method (Parrish and Derber, 1992) that computes the model error covariances from statistics on sets of differences of model forecasts for the same validating time (36 and 12 hours forecasts in our case). 3 months of daily ALADIN forecasts (used as a dynamical adaptation of ARPEGE at 10 km horizontal resolution) have been used for that purpose.

- The "lagged-NMC" method (Široka et al., 2002) which follows the same approach as the standard NMC method, except that the short term run uses almost the same lateral and initial conditions as the long term run. This allows to reduce the large scale variance and to get sharper analysis increments more adapted for mesoscale assimilation

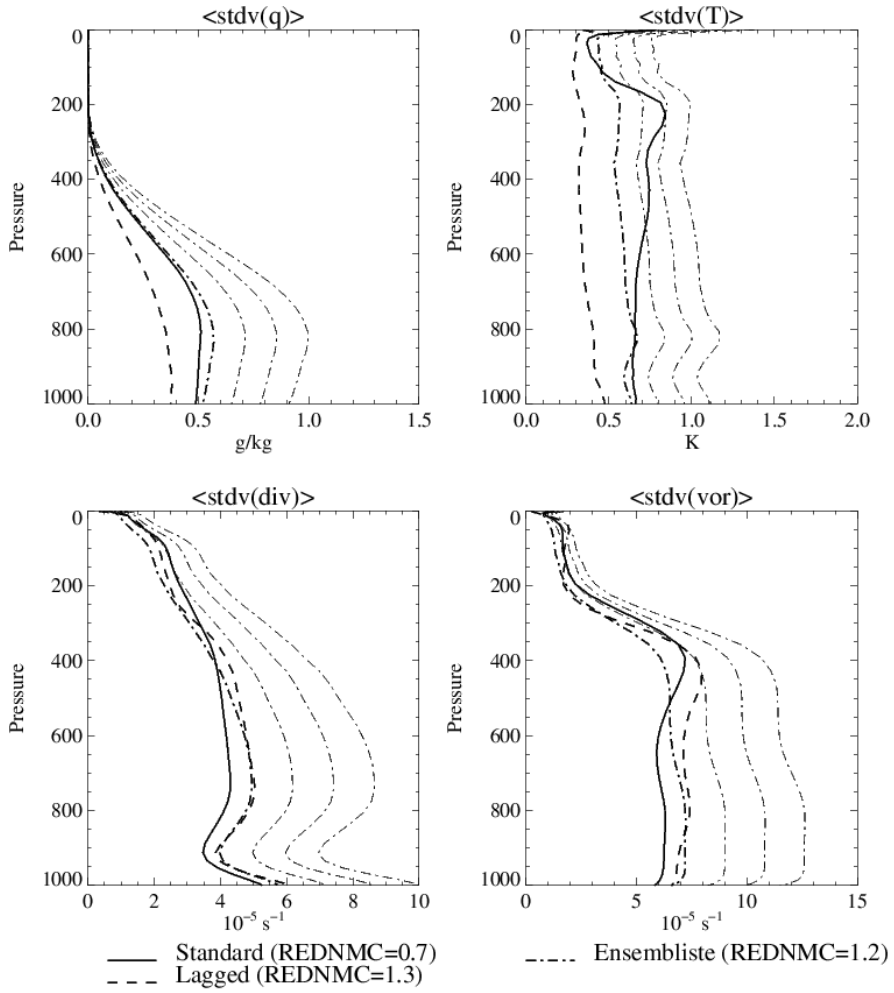


Fig 1 : *Mean vertical standard deviations of the background errors for humidity, temperature, wind divergence and vorticity (from top left to bottom right) for different estimation method (see text for details). The REDNMC factor allows to perform a posteriori callibration. Thin dashed lines represent variances obtained with the ensemble method using different REDNMC factors (1.2, 1.5, 1.8 and 2.1).*

- The "ensemble Jb" (Stefanescu etal., 2005, Berre etal., 2005) computed from an ensemble of Aladin forecasts with initial conditions from an ensemble of Arpège analyses. The sample is composed of two pairs of 6 hour forecasts over 48 days that are extracted from the ensemble. Their difference is then computed which gives 96 elements in the sample. This approach allows to represent the effect of the analyses and of the short range forecasts in a more accurate way.

The three different **B** matrices deduced from these methods have been tested by running three complete assimilation cycles (with 4 assimilations per day) using ALADIN 3DVar during 1 month (june 2003). Forecast scores for temperature (not shown) display a strong negative bias around the tropopause for the lagged-NMC and the ensemble method. Studies have demonstrated that this bias was brought during the model integration, not during the assimilation step. The mean vertical background error variances are plotted in Fig. 1 for each of the control variable of the model. With the appropriate a posteriori calibration, these variances are comparable in shape and in intensity except for the temperature in the high troposphere. This explains why forecast scores on

temperature for the run that uses the standard NMC do not present any bias around the tropopause: its larger background error variances at these levels allow a better correction of the bias with the successive assimilations of high troposphere observations like radiosondings and/or AMSU-A data. However, this method does not represent the effects of the analysis and of the short forecast ranges in an accurate way, and therefore it produces large vertical and horizontal correlation lengths that are not well suited for data assimilation at regional scale. As a consequence, the total error variances of the two other methods have been corrected for further experiments in order to allow the data assimilation to correct the model temperature bias. The resulting experiment that uses the ensemble **B** matrix with this a posteriori calibration showed better forecast scores against radiosoundings than the others (not shown). The ensemble **B** also possesses correlation lengthscales which are in between those of standard and lagged NMC, and therefore the ensemble **B** appears as a reasonable compromise, with the "most likely" structure functions. Thus, this formulation has finally been chosen for ALADIN 3DVar.

Further research will be devoted in future to re-calibrate more thoroughly the error variances, possibly using a more objective a posteriori technique.

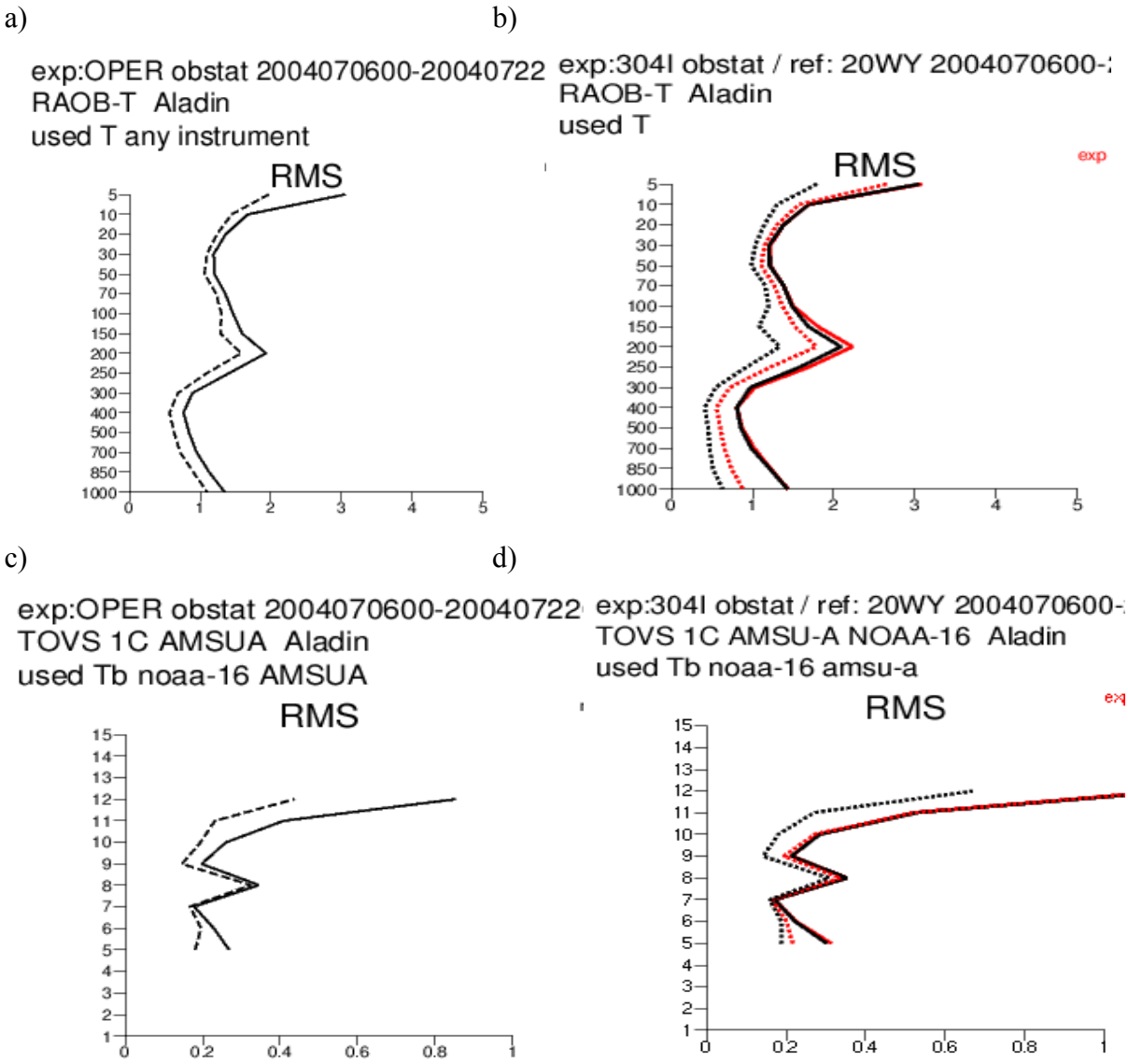


Fig.2 : *Assimilation statistics over the ALADIN domain for the july 2004 test period : rms errors of (obs-guess) (plain lines) and (obs-analysis) (dashed lines) for a) ARPEGE and b) ALADIN against temperature radiosounding and c) ARPEGE and d) ALADIN against AMSU-A onboard NOAA-16 data (the vertical axis denotes channel number). For ALADIN (right panels), black lines correspond to CNTRL and the red lines the experiment that uses a lagged-NMC B matrix.*

2.3 The CNTRL experiment

In order to study the relative impact of SEVIRI data within the ALADIN 3DVar, a control experiment (CNTRL hereafter) has been run. This experiment has the following characteristics:

- Following the results displayed in the previous section, it uses the ensemble **B** matrix.
- The computational domain covers Western Europe with a 10 km horizontal resolution and 41 vertical levels, with layers of 50 m to 2 km depth.
- It assimilates every 6 hours the same complete set of observations as ARPEGE at the time of the experiment within an assimilation window of +/- 3 hours to ensure a maximum of variety and coverage of observations. These observations are among others ground based data,

aircraft measurements, radiosoundings and polar-orbiting satellite radiances from AMSU-A and HIRS.

A test period of 15 days with 4 daily assimilations has been performed from the 6^{h} to the 22nd of July, 2004. 36 h forecasts have been run from each analysis time.

To validate this experiment, statistics on the assimilation have been performed for the whole period of evaluation and compared to statistics obtained with ARPEGE on the same domain. Results for temperature measured by radiosoundings and AMSU-A onboard NOAA-16 are displayed in Fig. 2. The curves of rms errors present the same pattern for the two cases. The (obs-guess) rms error against temperature observed by radiosoundings is slightly larger for ALADIN, indicating that the model errors are propagating more quickly during a 6 hour forecast with a higher horizontal resolution and a smaller timestep. However, more information coming from the radiosoundings are taken into account during the 3DVar assimilation process for ALADIN, the difference between the (obs-guess) and (obs-analysis) rms error curves being larger than for ARPEGE. For ALADIN, these figures confirm also that the choice of a proper **B** matrix is very important: statistics plotted for the experiment that uses the "lagged-NMC" **B** matrix show indeed much larger (obs-analysis) errors. The fact that the two curves of (obs-guess) rms errors are closer for the 2 formulations seems to indicate that a lot of information brought by the 3DVar vanishes after 6 hours of forecast.

This point is confirmed by the forecast scores against radiosoundings plotted in Fig. 3. The scores of CNTRL have been computed relatively to ALADIN used as the dynamical adaptation of ARPEGE (called DA hereafter) (i.e without data assimilation scheme). CNTRL presents a weak bias reduction (1 m) in the stratosphere for geopotential height after 12 h of forecast. The positive impact for the temperature in the analysis, all over the troposphere, is however lost after 12 hours. A weak degradation of the bias of approximately – 0.1 K can be seen below the tropopause and seems to propagates downward with time. Scores on humidity show approximately the same behaviour than for temperature but shifted downward. Finally, the use of cycled 3DVar analyses allows also to reduce the standard deviation of the wind vector up to 0.8 ms^{-1} in the analysis, with a maximum in mid troposphere. For all the parameters, the additional information which is brought by the Aladin analyses (compared with the Arpège analyses information) is lost after 12 hours of integration : the model produces from these initial states forecasts that are not better statistically than ALADIN runs based on the Arpège analyses. Emphasis on precipitation forecasts will be made in the next section.

DA.r12/TP – CNTRL.r12/TP

(15 cases, 06/07/2004 12UTC -> 22/07/2004 12UTC)

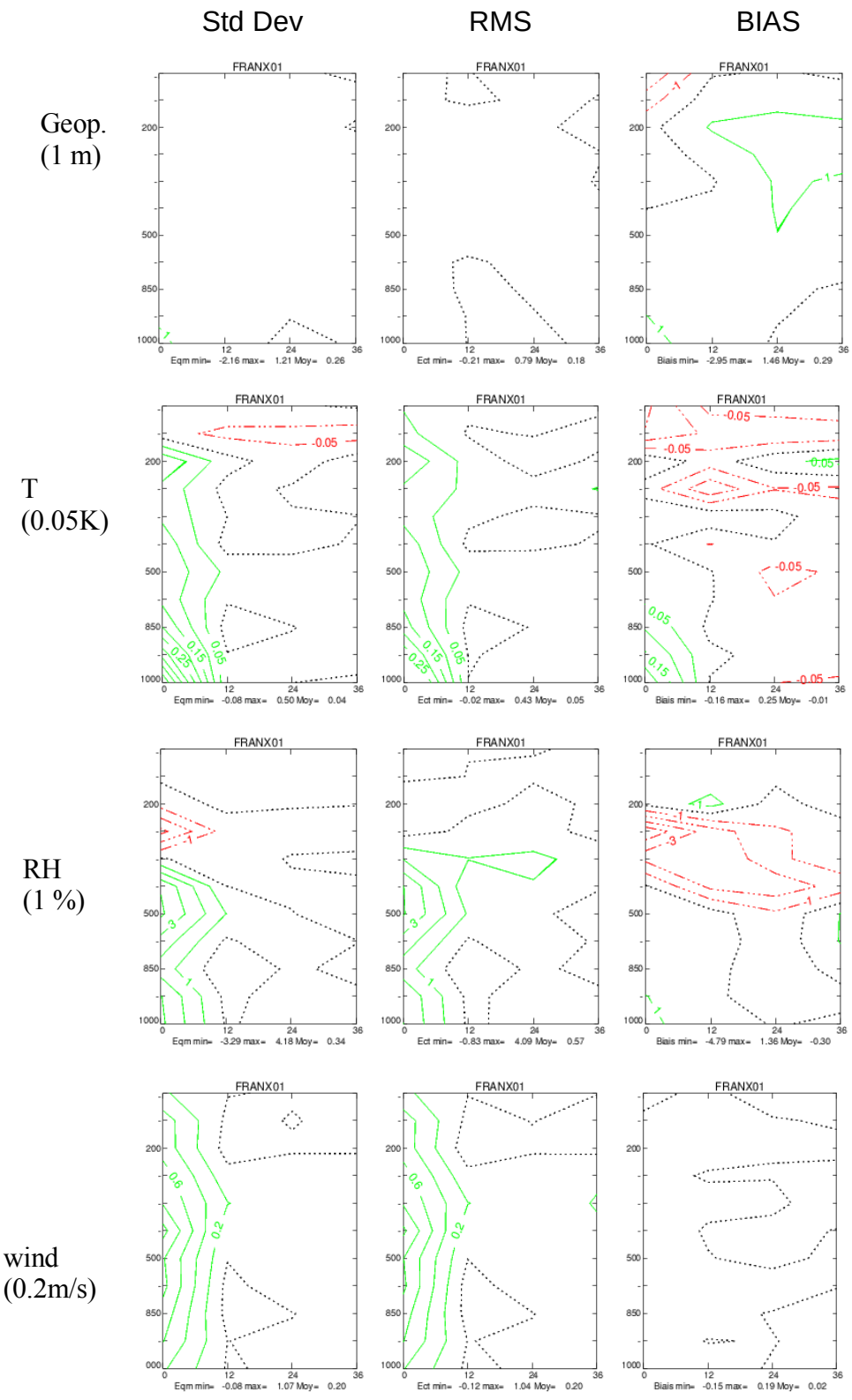


Fig. 3 : *Forecast scores of DA vs. radiosoundings minus CNTRL vs. radiosoundings over the ALADIN domain for the july 2004 test period. Left column is the standard deviation, middle the rms errors and right the bias. Green isocontours denote positive impact of CNTRL.*

3. Impact of Météosat-8 SEVIRI raw radiances

In Montmerle (2004), 4 successive hourly assimilations of SEVIRI raw radiances in ALADIN 3DVar have been performed to study the contribution of its high temporal resolution on cloud cover prediction. It has been shown that it allows to take into account the response of the environment to explosive meteorological systems such as storms. In the present report, we focus on the impact that SEVIRI data could have in the ALADIN 3DVar with an operational configuration, i.e for cycled assimilations every 6 hours.

3.1 Presentation of the product

Since May 2004, the CMS (Météo-France/ Centre de Météorologie Spatiale, Lannion, France) is sending to Toulouse a SEVIRI/MSG product devoted to ALADIN studies. This product is sent in GRIB format every hour and is composed of different fields at full-resolution covering all European ALADIN domains :

- The 8 IR SEVIRI channels, from 3.9 μ to 13.4 μ
- The associated date, lat/lon position, angles of sight
- A cloud type (CT hereafter) and the cloud top pressure with the associated quality flags.

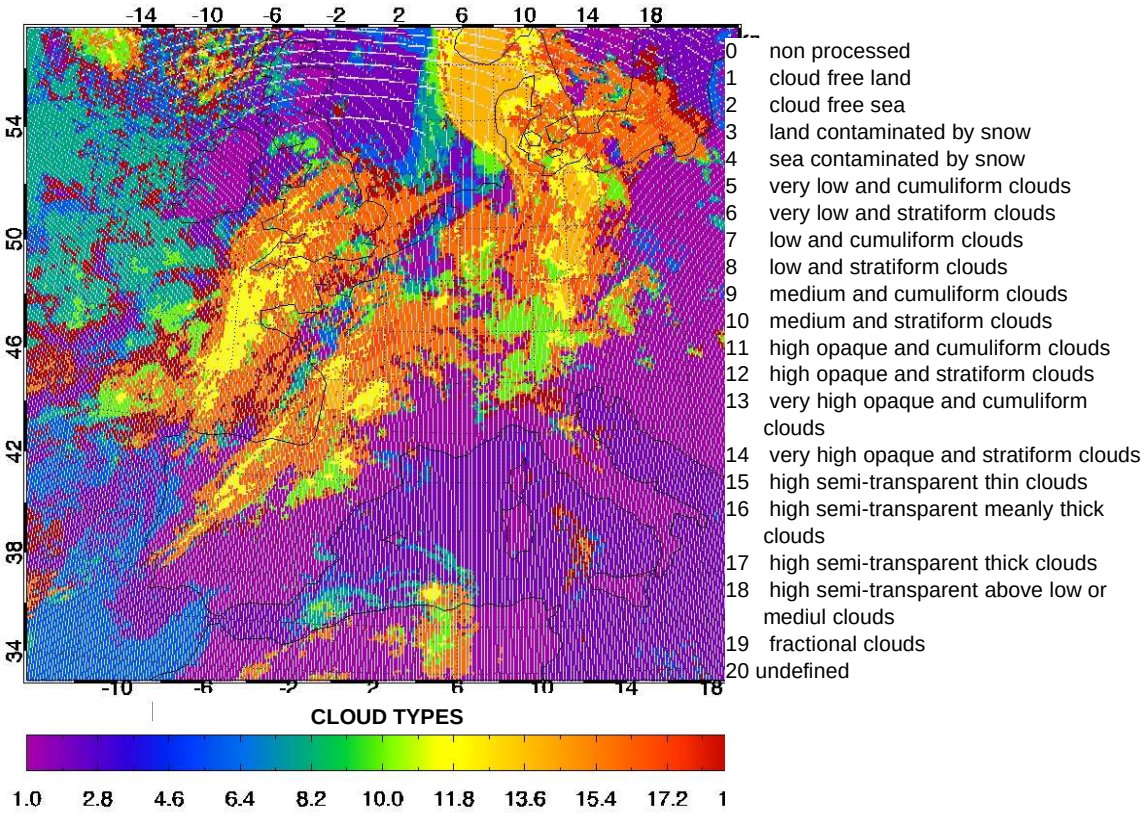
As described in the next section, the latter fields permit to keep in the assimilation process channels whose weighting functions peak above the cloud top. These cloud products have been developed by CMS in the SAF/NWC MSG framework. Complete documentations about this SAF/NWC can be found at http://www.meteorologie.eu.org/safnwc/ . The CT product contains information on the major cloud classes : fractional clouds, semitransparent clouds, high, medium and low clouds (including fog) for all the pixels identified as cloudy in a scene. The set of thresholds to be applied depends mainly on the illumination conditions, whereas the values of the thresholds themselves may depend on the illumination, the viewing geometry, the geographical location and NWP data describing the water vapour content and a coarse vertical structure of the atmosphere.

Briefly, The CT classification algorithm is based on the following approach :

- Main cloud types are separable within 2 sets: the fractional and high semitransparent clouds from the low/medium/high opaque clouds. Spectral and textural features are used to separate these 2 classes: ($T_{10.8\mu} - T_{12\,\mu}$) and ($T_{3.9\,\mu} - T_{10.8\,\mu}$) brightness temperature difference (in night time conditions only), $R_{0.6\,\mu}$ visible reflectance (in day time), variance of $T_{10.8\,\mu}$ coupled to variance $R_{0.6\,\mu}$ in daytime conditions.
- Within the first set, the fractional and high semitransparent clouds are split using ($T_{8.7\mu} - T_{10.8\,\mu}$) and in $R_{0.6\,\mu}$ daytime conditions.
- The opaque clouds are distinguished through the comparison of the window channel $T_{10.8\,\mu}$ to ARPEGE forecast brightness temperature at several pressure levels using RTTOV radiative transfer model (Saunders and Matricardi, 1999).

Fig. 4 gives an example of this product and its associated cloud top pressure for the 18th of July 2004. It has to be noticed that no separation between cumuliform and stratiform clouds is currently done in the CT product.

a)



b)

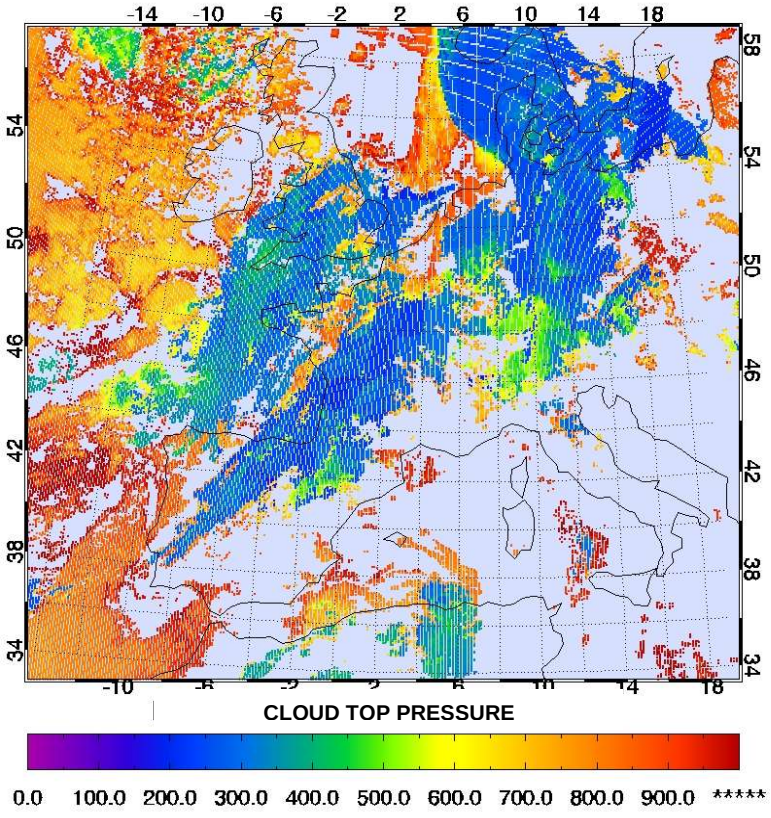


Fig. 4 : *a) Cloud types and b) cloud top pressure for the 18th of July, 2004.*

3.2 Pre-processing of the data

The SEVIRI radiances assimilated in the configuration of ALADIN 3Dvar presented in this report are pre-processed in the following way :

- To keep the observations relatively uncorrelated, one pixel out of 5 is extracted from the database, which gives approximately a 25 km horizontal resolution over France, and thinning boxes of 66*66 km^2 are applied during the screening
- The near IR 3.9 μ and the ozone 9.7 μ channels are blacklisted. The broad 3.9 μ channel is not used because RTTOV has troubles to simulate it (Roger Saunders, personal communication)
- Since the domain of interest is relatively small, a constant bias is used as a first hypothesis for the remaining channels. The following values are used :

  Bias = (0, -2.8, 0, -0.6, 0, -0.05, 0, -0.3)

- The observed brightness temperature error for each channel has an empirical value, based on measurement errors and errors due to RTTOV :

  σ_o = (1.05, 1.7, 1.7, 1.05, 1.05, 1.05, 1.05, 1.05)

The uncertainty of the humidity estimation in the troposphere leads to take a larger σ_o for the two WV channels.

- A quality control is applied to reject data whose (obs-guess) value exceeds the sum of the background and the observation error variances (σ_b and σ_o respectively) times an empirical constant α :

$$\left(\left(y - H\left(x^b\right)\right)/\sigma_b\right)^2 > \alpha\left(1 + \sigma_o^2/\sigma_b^2\right)$$

- The CT product presented in the previous section is used to select channels following its value : the low peaking channels IR 8.7 μ , 10.8 μ and 12 μ are kept only in clear sky conditions, the 13.4 μ is also kept above very low clouds and the two WV channels are considered even above mid-level clouds.

3.3 Impact on analyses

To compare with the CNTRL experiment presented in section 3.3, an experiment that presents the same characteristics than CNTRL but with SEVIRI radiances added (SEV hereafter) has been run during the same period of July 2004. Assimilation statistics plotted in Fig. 5 show that a lot of information coming from the 6 assimilated channels is taken into account in the analyses. The (obs-analysis) rms errors over the whole test period are indeed much lower than the (obs-guess) ones. The relative error decrease is however less pronounced for the 13.4 μ channel as noted in M2004 which is probably due to the broader shape of its weighting function and/or the choice of a non optimal observation error variance. The mean biases have values less than 0.2 K which seems to justify the values of the constant bias correction.

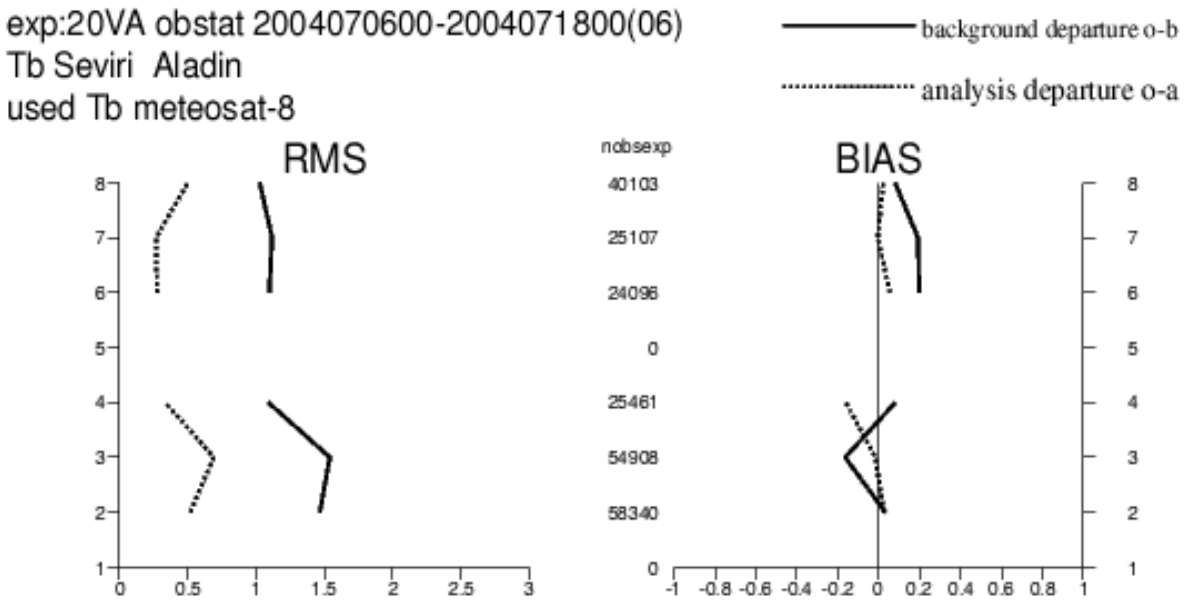


Fig. 5 : *Assimilation statistics for the SEV experiment for the July 2004 test period. The vertical axis denotes the channel number, the left panel the RMS error and the right panel the bias between (obs-guess) (plain line) and (obs-analysis) (dashed line) brightness temperature.*

3.4 Monitoring

Monitoring has been performed and results are plotted on Fig. 6 for the six assimilated channels. It shows firstly a strong negative bias of about –2.6 K for the WV 6.2 μ channel which is well corrected by the flat bias correction. As for the 3.9 μ channel which is blacklisted, this bias is due to its broad spectral resolution that is badly taken into account by RTTOV. The bias corrected channels present very stable features during all the period. A diurnal cycle is visible for the biases and the number of active data for the low peaking channels. For each analysis time, about 1500 observations from the WV channels, 1000 for the 13.4 μ and between 500 and 1000 for the three other channels are considered in the variational process. The (obs-guess) rms error presents also a weak oscillation for the WV 6.2 μ that coincides with two peaks of convective activity at the beginning and at the end of the test period.

3.5 Impact on forecast

3.5.1 Forecast scores

As for CNTRL, forecast scores have been computed relatively to the dynamical adaptation version of ALADIN (DA) and are plotted in Fig. 7. For the geopotential height, the assimilation of SEVIRI data reduces the bias against radiosoundings of about 4 m between 12 and 24 h of forecast above 600 hPa, and increases it slightly below. This impact induces a negative bias on sea level pressure during the forecast which is difficult to explain since no negative bias is present in the analyses. For the temperature, SEV implies a decrease of rms error on all vertical levels before 12 h and up to 24 h near 300 hPa. A negative bias is however present from the start above 300 hPa and in the middle troposphere after 6 h. The rms error for humidity is slightly improved before 6 h of forecast and up to 12 h near 400 hPa. The analyses for this quantity show however small biases at all vertical levels that propagate downward with time. Finally, SEV exhibits a larger analysis fit to the observations than CNTRL in mid to high troposphere, but slightly worse wind forecast scores.

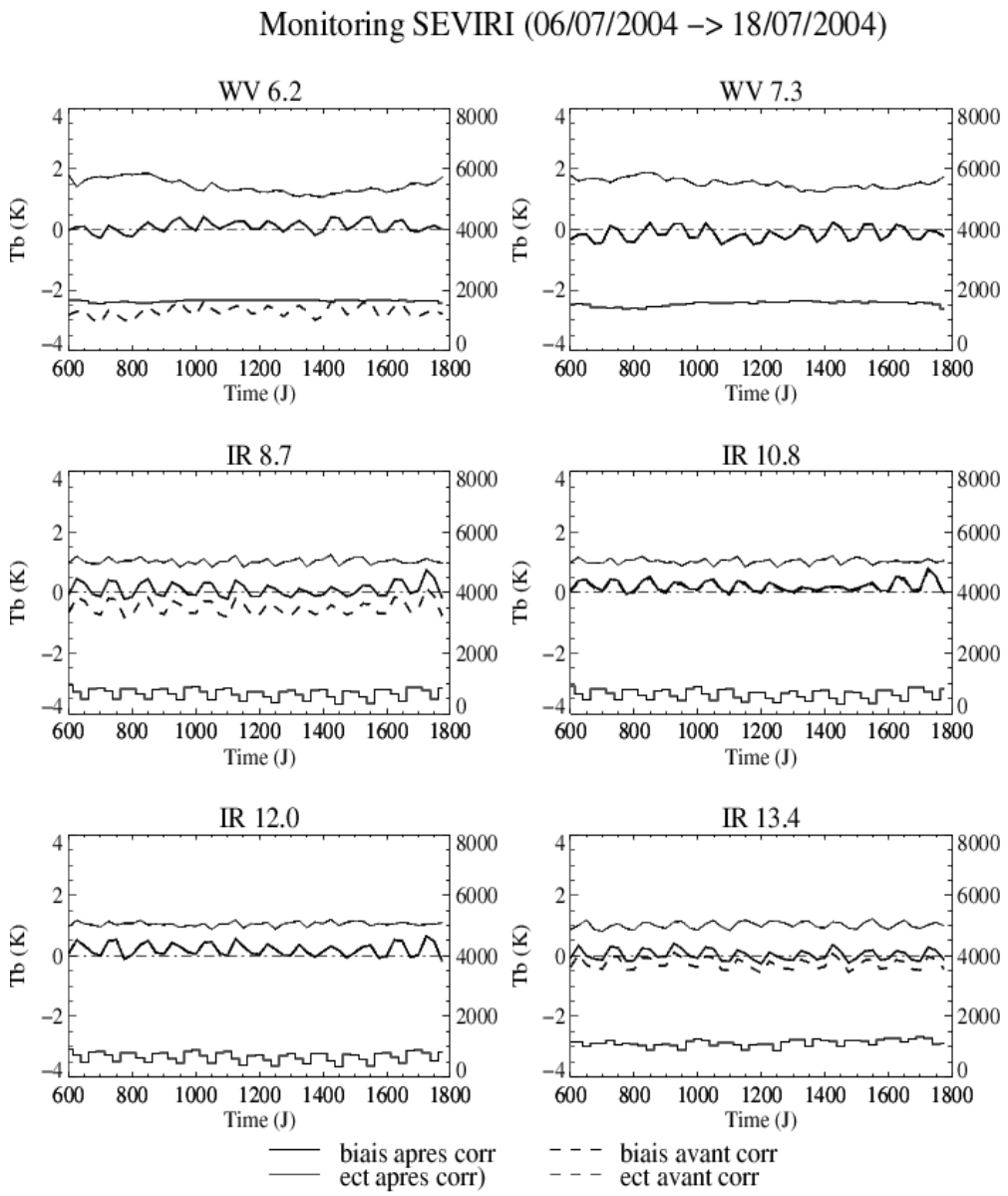


Fig. 6 : *Monitoring for the 6 assimilated SEVIRI channels for the SEV experiment from the 6th to the 18th of July, 2004. Histograms on bottom of figures (associated with the right vertical axis) represent the number of active data that enter the minimization.*

Globally, the scores of CNTRL against radiosoundings have been slightly degraded. This can be explained by the fact that the large amount of SEVIRI data added in the assimilation process has slightly taken away the analysis from radiosoundings observations which are the main source of observation in CNTRL. To give better weight to the different observation types, the tuning of the observation error variances will be undertaken in the near future, following the work of Chapnik (2004) .

DA.r12/TP – SEV.r12/TP

(15 cases, 06/07/2004 12UTC -> 22/07/2004 12UTC)

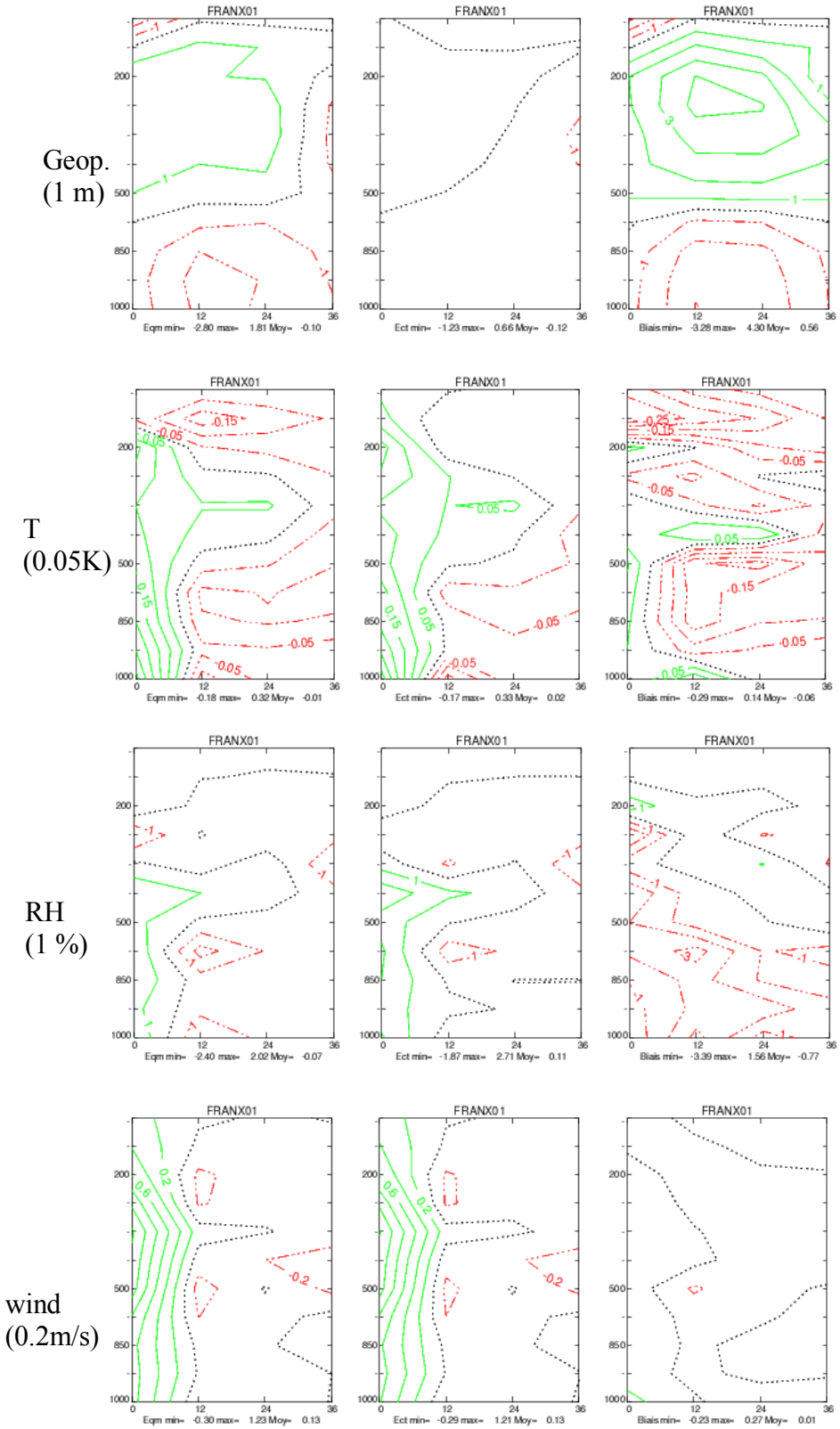


Fig. 7 *: as Fig. 3 but for the SEV experiment.*

3.5.2 Total rainfall forecast

- Case of the 18th of July 2004 :

The total rain forecasted between 6 and 12 hours of lead time by DA, CNTRL and the SEV experiments from the 00 UTC analysis are plotted on Fig. 8 and compared to rain gauge values over France. DA missed the NE/SW orientation of the main rain band. CNTRL produces the good orientation and the SW part of the line seems realistic, although a little bit too south. The use of SEVIRI data allows to forecast the observed 2nd cell of intense precipitations located in the NE part of the line with a slightly overestimated amount (> 20 mm). The maximum over the Bordeaux region is however located too south but with an amount of 40 mm comparable to rain gauge observations. The secondary line of precipitation is also quite well captured over the NE of France with realistic shape and amount.

To understand why SEV produces the observed second cell of intense precipitation in the north-eastern part of the line contrary to CNTRL, increments of humidity and temperature at 700 hPa of the 00 UTC analysis have been plotted for the two experiments in Fig. 9. The most striking difference between the two is that increments produced by SEV present more realistic and mesoscale patterns than for CNTRL where the main source of information seems to come mainly from radiosoundings. In particular, SEVIRI data are cooling and humidifying the mid to low troposphere pre-convective area located upstream of the frontal rain band over western France which produces intense rain 6 hours later.

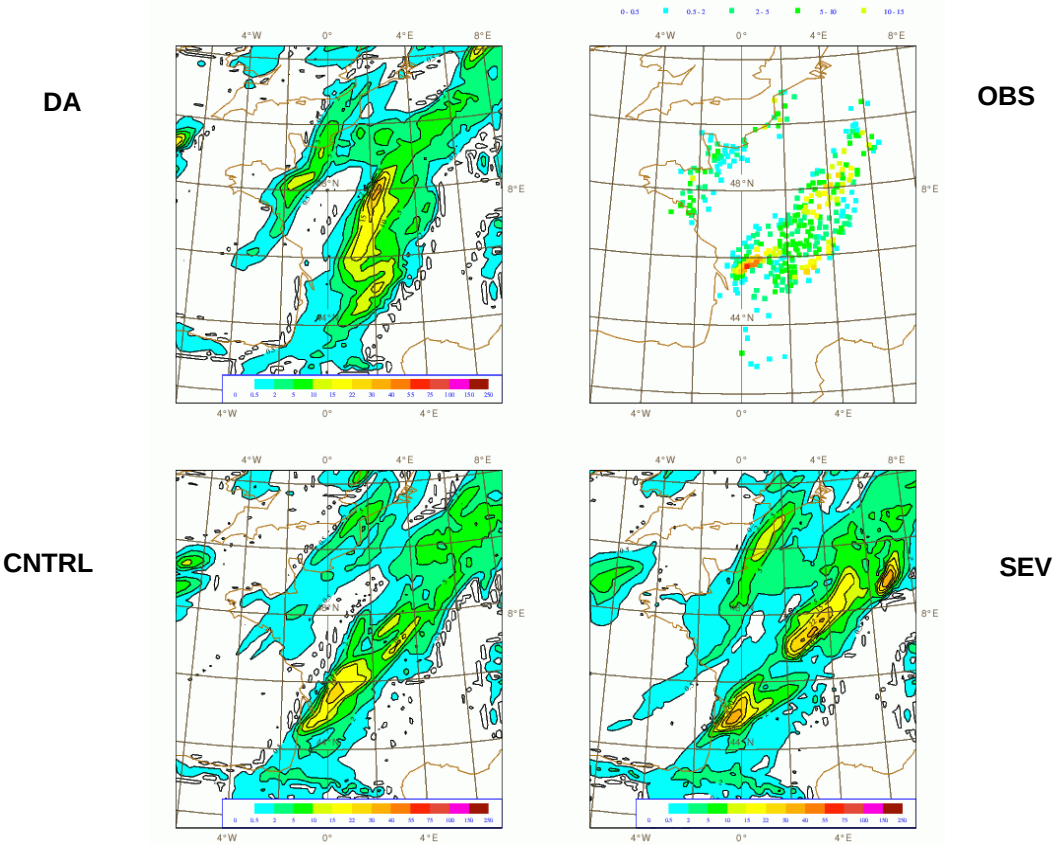


Fig. 8 : *Rain gauges observation (top right) and simulated total rainfall between 12 and 6 h of forecast for the 18th of July 2004.*

As plotted on Figs. 10 and 11, a large amount of IR radiances coming from SEVIRI is taken into account in the assimilation process compared to ATOVS data for example. Depending on the analysis time, the ratio of the number of data that enters the screening for these two observation types is varying between 10 and 25 (Fig. 10). It has however to be noted that for the CNTRL and SEV, ATOVS data have been assimilated using the screening features of ARPEGE. For HIRS for

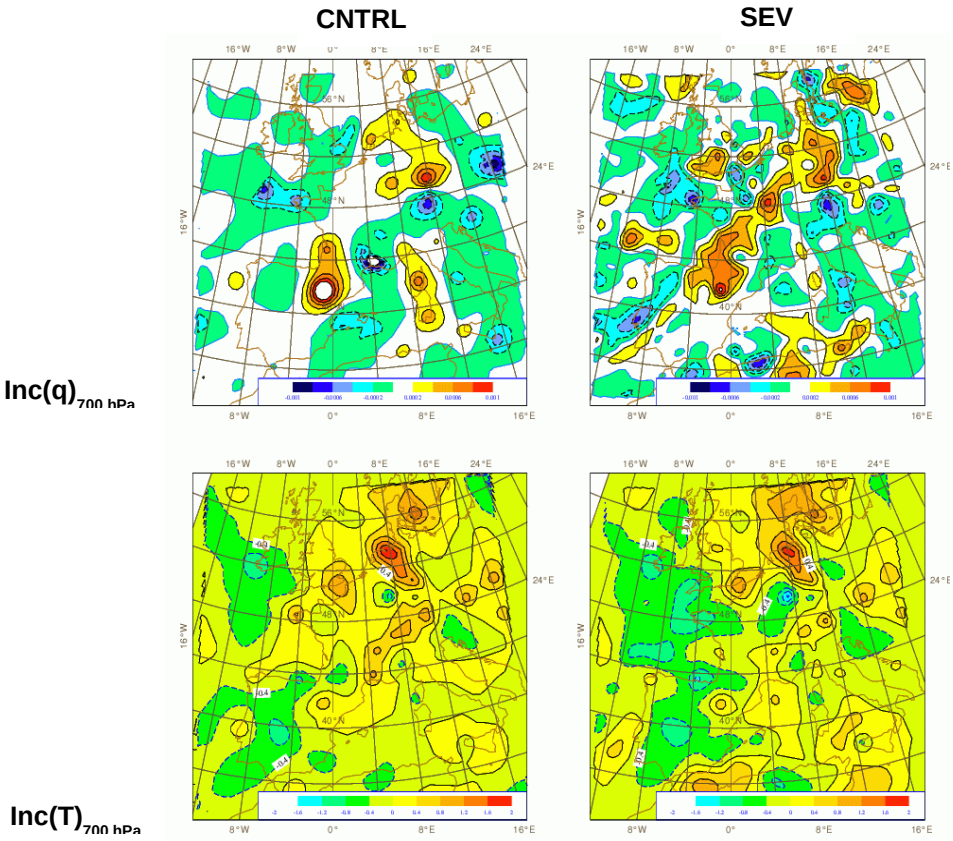


Fig. 9 : *Humidity increments (top panels) and temperature increments (bottom panels) for CNTRL (left) and SEV (right) for the 18h of July 2004 at 00 UTC.*

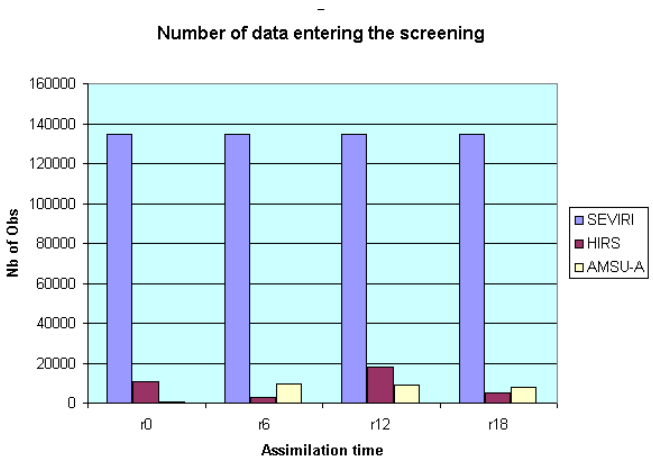


Fig. 10 : *Number of radiances entering the screening for the 18h of July 2004.*

instance, 1 pixel out of 5 have been extracted and thinning boxes of 250*250 km^2 have been applied which is not comparable to SEVIRI (Fig. 11).

In the near future, the impact of higher density ATOVS data will be tested in the 3DVar using two complementary approaches:

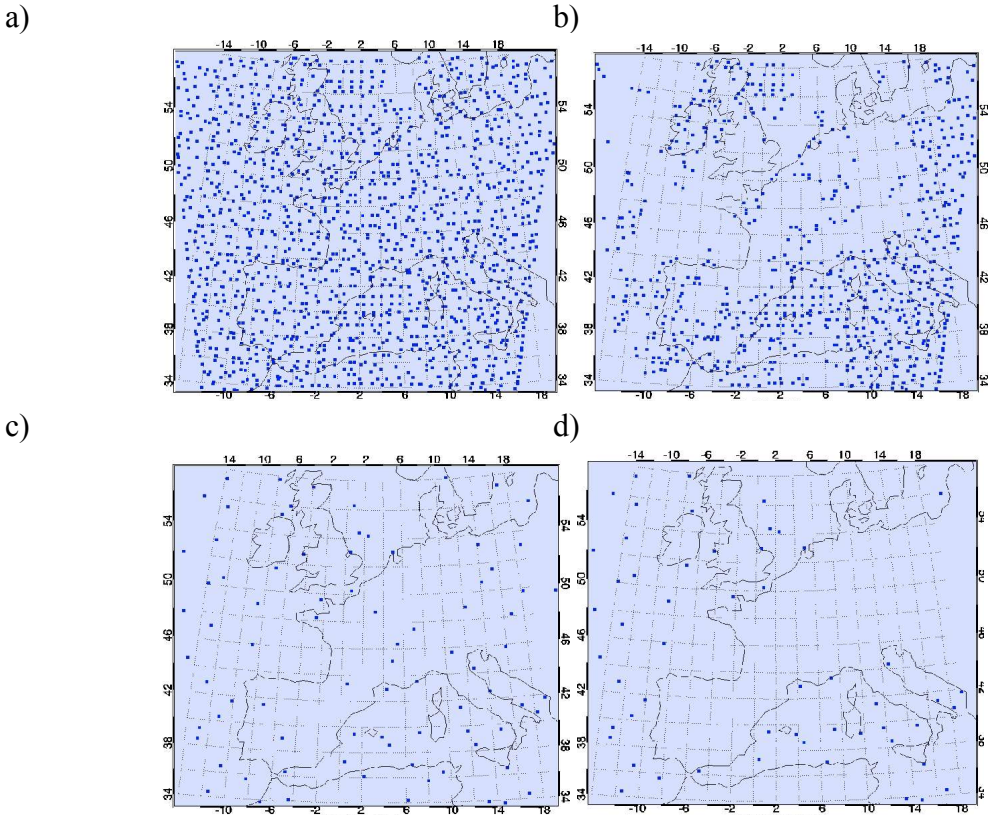


Fig. 11 : *Active data for the 18th July 2004 at 00 UTC for a) SEVIRI WV 6.2 μ, b) SEVIRI IR 10.8 μ, c) HIRS channel 12 and d) HIRS channel 4.*

- Extraction and thinning at higher horizontal resolutions
- Use of the EARS (Eumetsat ATOVS Retransmission Service) data that are already used in the operational ARPEGE suite. Their shorter reception time delay allows indeed to get more data within the +/- 3 hour assimilation window considered in the ALADIN 3DVar.

- Case of the 8th of July 2004 : The total rain forecasted between 6 and 12 hours of simulation by CNTRL and SEV compared to rain gauges and DA are displayed on Fig. 12. DA produces unrealistic large amounts (over 40 mm) of rain over NE of France contrary to the two 3DVar experiments that are more comparable to observations. CNTRL reproduces well the shape and intensity of the northern part of the N-S oriented line of heavy precipitations located in eastern France, whereas DA totally missed it. The addition of SEVIRI data allows to enhance realistically precipitations in its southern part with amounts up to 20 mm and to produce light rain over the centre of France that are observed by rain gauges.

- Case of the 22nd of July, 2004 : for that case, DA underestimates strongly the precipitations that occur over the western part of France (Fig. 13). The use of a cycled 3DVar allows to correct this failing. Shapes and intensities of the precipitating cells as forecasted by SEV seem moreover in better agreement with rain gauge observations and rain rates derived by radars over the sea (not shown) where amounts bigger than 30 mm were measured.

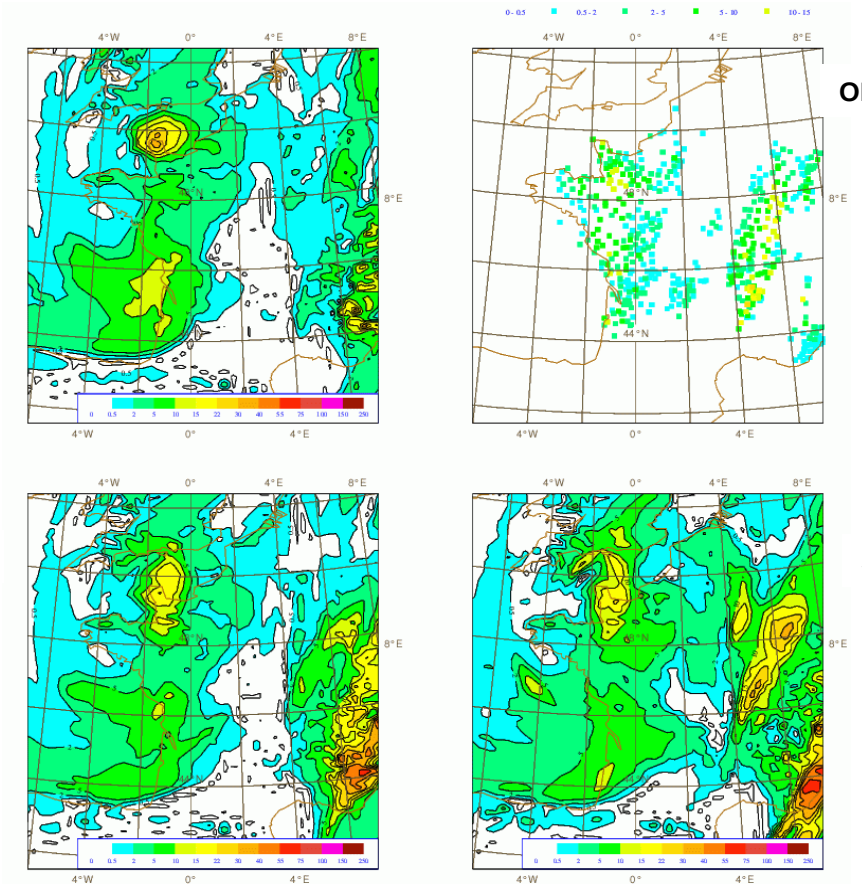


Fig. 12 : *same as Fig. 8 but for the 8th of July 2004.*

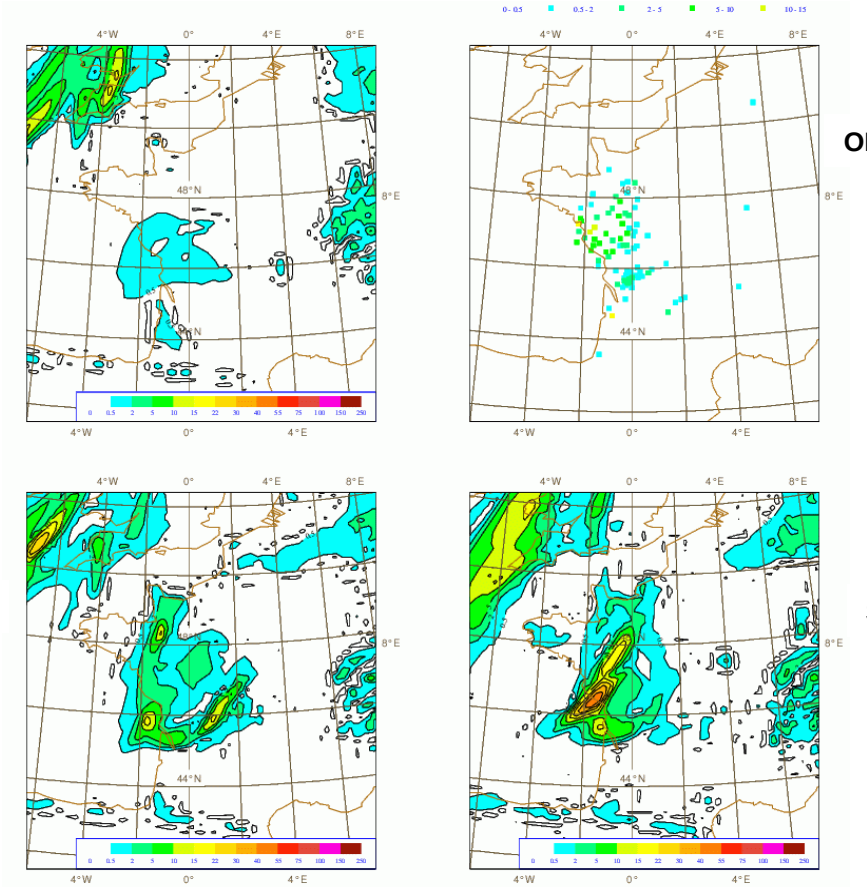


Fig. 13 : *same as Fig. 8 but for the 22nd of July 2004*

3.5.3 QPF scores

The following Quantitative Precipitation Forecast (QPF) scores were computed for the July test period for different thresholds x:

Consider the table:

	Sim<x	Sim>x
Obs<x	*A*	*B*
Obs>x	*C*	*D*

The following QPF scores are defined:

Frequency Bias: $\text{FBIAS} = \frac{B+D}{C+D}$ = (number of forecast > x)/(nb of obs. > x)

Equitable threat score: $\text{ETS} = \frac{(A+D)-(B+C)}{(B+C)(A+B+C+D)+AD-BC}$

Probability Of Detection: $\text{POD} = \frac{D}{C+D}$ = (nb of obs > x & nb of for. > x)/(nb of obs > x)

False Alarm Rate : $\text{FAR} = \frac{B}{B+D}$ = (Nb of overerestimated forecast)/(nb of forecast > x)

The observations used to compute the scores are the 6 hour total rainrates measured by rain gauges.

The two detection scores (ETS and POD) displayed in Fig. 14 are higher for the experiments that

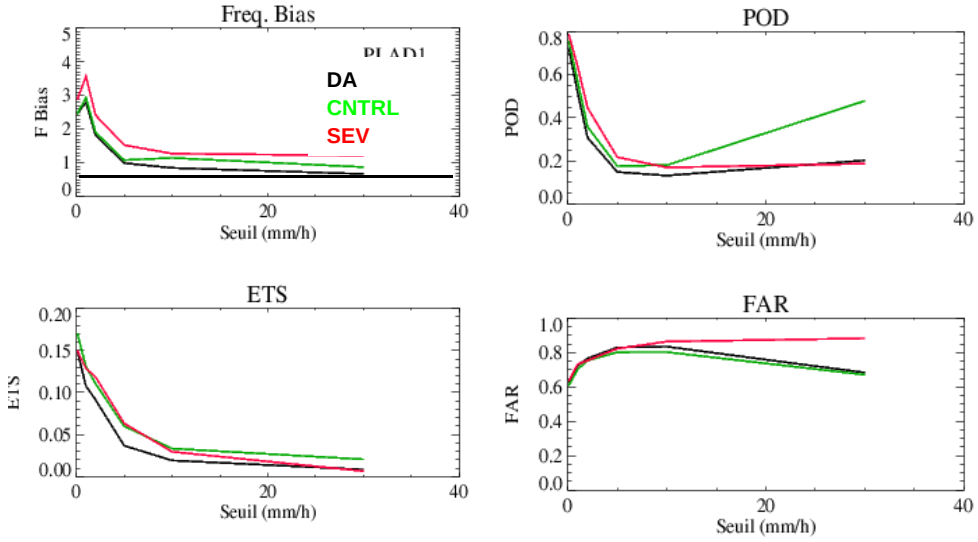


Fig. 14 : *QPF scores for DA, CNTRL and SEV computed for the whole July period for the total rain forecasted between 12 and 6 h from the 00 and the 12 UTC analysis time (see text for definition of the scores). Precipitation thresholds are 0.1, 1, 2, 5, 10 and 30 mm.*

are using an assimilation scheme. The ETS is comparable for CNTRL and SEV and shows values almost two times greater than DA for the 5 mm threshold. The addition of SEVIRI data permits to

perform a better detection of precipitating events mostly for the 2 and the 5 mm thresholds, with respective POD of 0.46 and 0.23, compared to 0.34 and 0.16 displayed by CNTRL. However, this better detection is made to the detriment of the FBIAS : SEV produces too much precipitations for all thresholds. For small thresholds, the overestimation of the number of simulated precipitating pixels shown by DA and CNTRL is accentuated for SEV. For thresholds greater than 5 mm, the FBIAS are comparable for the 3 experiments although slightly greater than 1 for SEV. Finally, the FAR is greater for SEV for the 30 mm threshold. Since a very small number of observed/simulated pixels are characterized by values greater than this threshold, QPF scores are weakly representative at this level.

Assimilating SEVIRI data using the first configuration defined in section 4.2 seems thus to produce too much precipitation spatially, particularly light rain. The amount of information given by these radiances during the assimilation step will be weakened in order to limit this drawback through the tuning of the observation error variances and/or the use of larger thinning boxes.

4. Spin-up characteristics of the proposed data assimilation cycle

Figures 15 and 16 show the time evolution of surface pressure in the first 6 hours of integration, for the situation of July 22nd (an active case from the 2 week period). Three points are displayed: one point in the lateral boundary relaxation zone (5 points from the edge), one in the gulf of Biscaya (thus, over sea), one over the Dolomits (Alps). With the chosen settings for digital filtering (a non-incremental Dolph-Tchebychev filter, with stop-band edge period 1.5h and time span 35mn), two main characteristics can be stressed:

- the model surface pressure is fairly well balanced after 1-3 hours of integration. Before this lead time, low frequency oscillations probably do exist, although their impact on the scores or quality of meteorological fields is difficult to assess. Case studies on 1-3h lead time precipitations have shown some irregularities on the precipitating patterns (too wavy structures, misplaced kernels). After 3 hours of lead time, the case studies rather confirm the spin-up evaluation displaying meteorologically acceptable fields. Note that *without* digital filtering, a significant extra noise appears in the first 3 hours and later (not shown), which forbids a total removal of any filtering for the time being.
- Over orography, a more abrupt initial adjustment occurs, with an almost 2*dt variation. This kind of reaction seems to remain local, as it was not noticed to propagate to remote areas.

A by-side conclusion of this evaluation is that we do promote the evaluation of the short-range forecast scores and case studies starting from 3 h lead time. Especially, in addition to the large-scale synoptic scores used presently in Arpège and Aladin DA, we propose to compare the forecast at a hourly or 3-hourly basis with surface mesonet observations. This would already be a valuable diagnostic as many traditional scores only are available at 12-hourly basis, which is the main frequency for TEMP data.

In future, a continuous evaluation and work will be devoted to the spin-up question. This work will include firstly fine diagnostics and a survey of the model spin-up as the data assimilation system evolves (more observations, new cost function ...), and a survey and testing of additional initialization devices (balance constraints, incremental digital filters, the impact of the coupling data choice).

Fig. 15 : *Time series of surface pressure in Pa for 2 model points: lateral boundary relaxation zone (red) and Gulf of Biscaya (green). 9 model iterations represent approximately 1 hour of integration (dt=415.385 s).*

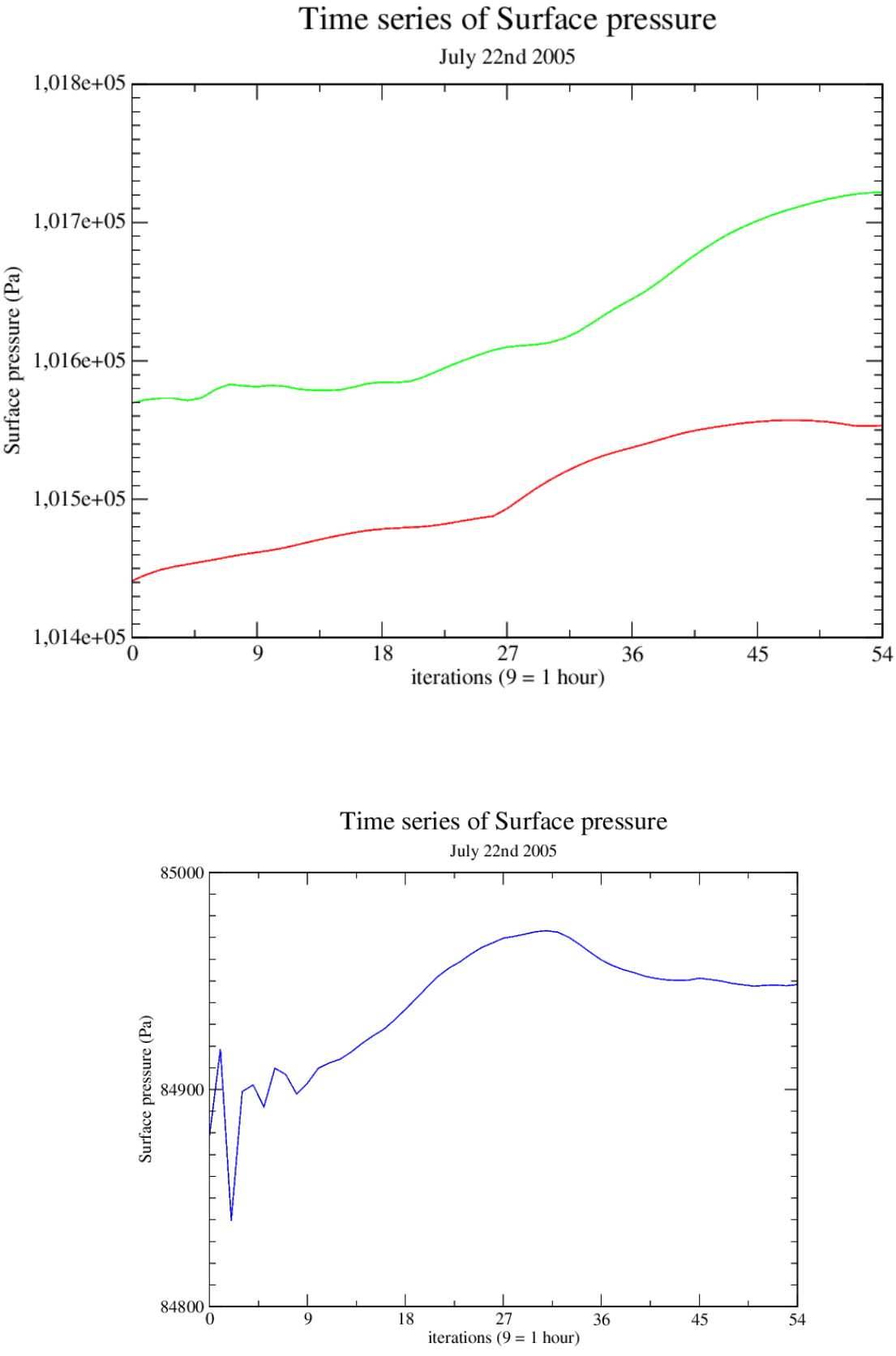


Fig. 16 : *Time series of surface pressure in Pa for 1 model point: over the Alps/Dolomit region (blue). 9 model iterations represent approximately 1 hour of integration (dt=415.385 s).*

5. Conclusions and future work

Results of studies devoted to the implementation of a pre-operational complete assimilation/forecast suite at regional scale using ALADIN and its 3DVar have been shown. A first and essential step has concerned the choice of the background error covariance matrix, the so-called **B** matrix, that allows to filter and to spread out the information brought by each type of observation during the assimilation process. Compared to the NMC-type of formulations, the **B** matrix computed from an ensemble of ARPEGE/ALADIN analyses/forecasts has shown the best compromise for data assimilation at regional scale: mesoscale correlation lengths and appropriate vertical correlations. A control experiment has been run during two test periods to test the impact of the use of a cycled assimilation scheme compared to the actual operational version of ALADIN which is simply the dynamical adaptation of the global model ARPEGE. This control run was using a 3DVar with the ensemble **B** matrix and was assimilating every 6 hours a complete set of observations within a +/- 3 hours assimilation window. The following general conclusions arised from this test experiment :

- Conventional scores match those of Arpège 4D-VAR over Western Europe. There is an impact on large scale scores ("output statistics") when the B matrix is changed ("input statistics").
- Marginal positive impacts have been shown on forecast scores against radiosoundings : compared to the dynamical adaptation, the rms errors remain slightly reduced after 12 h of forecast for specific fields and levels only. In general however, the improvement of the initial state is lost in a statistical sense, using conventional observational networks as a reference. Corresponding scores are indeed neutral which seems to indicate that after this delay the model is building solutions that are statistically comparable no matter what the initial conditions are.
- Precipitations are improved qualitatively and quantitatively for forecast lead times between +3 and +12 hours. Before (0-3 h), spin-up/spin-down processes are probably active and more investigation would be necessary. For the time being, we have decided to maintain non-incremental digital filter initialisation in both the assimilation cycle and in production forecasts. We have however decreased the strength of the filter, in a manner similar to DF blending, following the Prague experience.

SEVIRI data have then been added to this control experiment to study their relative impact. Channels 3.9 μ and 9.7 μ have been blacklisted, 1 pixel over 5 has been used, a constant bias has been applied for each channel and empirical error variances have been chosen in the first configuration. The cloud type classification computed by the CMS in the SAF/NWC framework has been used to keep data non contaminated by clouds in the variationnal process, which includes channels that peak over the cloud top. The monitoring shows stable features for all channels during the whole test period. A lot of information coming from SEVIRI radiances is taken into account in the analyses through the 3DVar, producing realistic increments. Results deduced from the 15 day test period are encouraging notably for the short term (i.e < 12h) precipitation forecast, especially when activity was under-estimated, where the addition of these kind of data allows to simulate realistic precipitation patterns in shape and intensity. Forecast scores are slightly degraded compared to the control experiment probably because of the large amount of additional data that slightly move away the analyses from radiosoundings. Moreover, QPF scores have shown that the experiment that includes SEVIRI radiances has better rain detection scores but produces spatially too much light precipitations.

One priority for the E-suite version will thus be to tune the error statistics and/or the thinning to lower the relative impact of SEVIRI in the analyses. Methods based on the use of the DFS (Degrees of Freedom for Signal) related quantities will be applied to improve covariance matrices (Desroziers and Ivanov, 2001; Chapnik etal., 2004). In parallel, the use of additional ATOVS data will be considered, using radiances coming from EARS (Eumetsat ATOVS Retransmission Sevice) and extracted with a better sampling. In particular, the impact of AMSU-B data is one major concern. Finally, the cloud top pressure product sent by the CMS and a convection detection

algorithm will be used to compute proxy humidity profiles for convective clouds for assimilation purposes.

The first configuration of ALADIN 3DVar (including SEVIRI radiances) is scheduled to go into E-suite hopefully around March 2005.

The further improvements for 2005/6 include the variationally formulated coupling with the Arpège analysis (so-called Jk), the FGAT computation of innovations (at appropriate time), the testing of a new humidity control variable (with more Gaussian error statistics), a more refined balance constraint (beta-plane, non-linear and omega equations), the inclusion of mesonet surface data and the testing of a more frequent analysis frequency (3 hours). The system will of course continue to benefit from the ongoing increase of Arpège data, possibly with specific thinning parameters for the regional scale. In parallel, the algorithm will need to be further adpated and shaped toward the requirements of the Arome model, including the cycling of specific surface fields, the proper treatment of extra dynamical variables and, at a much later stage, the inclusion of microphysical fields into the Arome control variable.

6. Appendix: evolution of E-suites in spring 2005

We list here the main characteristics of the experimental suites run in spring and summer 2005, with a quick scientific overview:

- Version 1 (V1): CY29T1_bf.03 with reduced DFI (TAUS=5400 s, NSTDFI=5). All sets of Arpège observations (Ps synop, buoys, aircraft airep/amdar/acars, paob/AMV, radiosonde data, pilot soundings, cleared radiances RAD1C) but not yet QuikSCAT wind data. In addition, Météosat-8 SEVIRI radiances also are assimilated, with a uniform and constant (simple) bias correction. Version 1 has run from March 23rd 0 UTC until May 22nd 18UTC. Main results from the V1 suite are:
    - an increased bias and RMS error on mean sea level pressure, by about 0.2 hPa. This error affects all forecast ranges from 3 hours onwards.
    - Too strong precipitations (factor 2 to 3 !) at 6 and 12 hours.
    - Too wet analyses (relative humidity at 850 hPa or 500 hPa)
    - the occurrence of more « Aladinades » in the assimilation suite has *not* been noticed (a good news, given the previously mentioned shortcomings).
- Version 2 (V2): CY29T1_op1, with code corrections for surface data thinning and SEVIRI bias correction and blacklisting. DFI are set back to dynamical adaptation (TAUS=10800 s, NSTDFI=9). All sets of Arpège data are used, plus 2m temperature and relative humidity from synop. SEVIRI radiance biases are now « dynamically » corrected using 4 predictors (2 geopotential depths of atmospheric layers, surface temperature, total water vapour content; the learning regression has been run over a sequence of specific screenings, through the V1 E-suite period). Additionally, all infrared channels over land have been blacklisted, as their assimilation was too much depending on the (poor) quality of the surface temperature of the model. The relative weight of the observational weak constraint (Jo), compared to the background constraint (Jb), has been reduced: REDNMC=1.8 in V1, decreased to 1.5 in V2, which amounts to a changing of the background error variances from 3.24 to 2.25. Furthermore, the number of iterations in the minimizer has been decreased from 70 (V1) to 50 (V2), which allows a correct convergence with a reduction of the norm of the gradient by about 100 to 1000, but not a fully horizontal curve on the Jo(iter) graphic. SEVIRI data are proven to considerably speed down convergence. Compared to V1, the new bias correction for SEVIRI diminishes drastically the assimilation of systematic biases from these data. The inclusion of 2m data dries off the PBL and lower troposphere: this systematic, biased, behaviour is accepted as it counter-balances the wet bias of SEVIRI data in the middle troposphere. DFI have been set back to « old » values as part of the MSLP problem most likely is a result of remaining imbalances, that may be of small intensity and fairly short period (1-2 h), but that probably were not efficiently enough filtered in V1 (DFI are not perfect step functions ...) and therefore survived and prospered. The Jb weighting reduction is performed to decrease the amplitude of analysis increments (in order to diminish their possible imbalance effects and the observational bias problems). When tested individually, these changes had the following impacts:
    - DFI: half less MSLP bias and RMS increase, compared to dynamical adaptation
    - SEVIRI bias correction: far less precipitation spin-down at 6 and 12 hours, better humidity analysis compared with V1
    - 2m T and RH: dryer low levels, less precipitations globally over the Aladin domain with more dry spots (less « drizzle »)
    - REDNMC and NITER/NSIMU: almost neutral

  Version V2 has run in E-suite from June 2nd, 0 UTC until July 25th. V2 has become operational on July 25th .
- Version 3 (V3): porting to cycle CY29T2. Additional changes are:
    - new climatological files (improved input data from ECOCLIMAP, computation performed with CY29T2/3 (?)).

- Lopez microphysics scheme for large scale precipitations switched on, which imposes the treatment of three extra fields in the assimilation cycle: cloud liquid water, ice, rain. These fields are cycled, but not analyzed.
- ECMWF radiation scheme RRTM switched on for long wave radiation, but the solar radiation scheme remains an older version of ECMWF's.
- New bias correction files for radiosoundings (same as in Arpège)
- miscellaneous technical adaptations to Arpège observation dataflow, as required by the change from OBSOUL to BUFR format.

Acknowledgements : The assimilation of SEVIRI into 3DVar is based on the outcome from EUMETSAT, previous studies being funded by Alcatel Space

References:

Berre, L, 2000: Estimation of synoptic and mesoscale forecast error covariances in a limited area model. *Mon. Wea. Rev.*, **128**, 644-667.

Chapnik, B., G. Desroziers, F. Rabier, and O. Talagrand. 2004. Diagnosis and tuning of observational error statistics in a quasi operational data assimilation setting. *Q. J. R. Meteo. Soc.*, Accepted for publication.

Courtier, P., Thépaut, J.-N. and Hollingsworth, A, 1994: A strategy for operational implementation of 4D-Var using an incremental approach. *Q. J. R. Meteo. Soc.*, **120**, 1367-1387.

Derber, J.C. and Bouttier, F., 1999: A reformulation of the background error covariance in the ECMWF global data assimilation system. *Tellus*, **51A**, 195-221.

Desroziers, G. and S. Ivanov, 2001: Diagnosis and adaptative tuning of observation error parameters in a variational assimilation. *Q. J. R. Meteo. Soc.*, **127**, 1433-1452.

Montmerle, T., 2004: Assimilation of satellite data in a regional mesoscale model. *EUMETSAT research fellowship 1st semester report*. 20 pp.

Parrish, D.F. and J.C. Derber, 1992 : The National Meteorological Center's spectral statistical interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747-1763.

Parrish, D.F, Derber, J.C. , Purser, R.J. , Wu, W.-S, Pu, Z.-X, 1997: The NCEP global analysis system: recent improvements and future plans. *JMSJ*, **75**, 359-365.

Saunders, R.W., M. Matricardi, and P. Brunel, 1999: An improved fast radiative model for assimilation of satellite radiance observation. *Q. J. R. Meteo. Soc.*, **125**, 1407-1425.

Široka, M., C. Fischer, V. Cassé, R. Brožkova and J.-F. Geleyn, 2002: The definition of mesoscale selective forecast error covariances for a limited area variational analysis. *MAP*, 1-18.

Stefanescu, S., L. Berre and M. Belo-Pereira, 2005: The evolution of dispersion spectra and the evaluation of model differences in an ensemble estimation of error statistics for a limited area analysis. *Submitted to Mon. Wea. Rev*.

Berre, L., S.E. Stefanescu and M. Belo-Pereira, 2005: A formal and experimental comparison between an ensemble estimation of limited area model error statistics and two other error simulation methods. *Submitted to Tellus*