

# Status and performance of HIRLAM 4D-Var

Nils Gustafsson  
Swedish Meteorological and Hydrological Institute  
Norrköping, Sweden

## 1 Introduction

The development of the HIRLAM 4-dimensional variational data assimilation (4D-Var) started already 11 years ago with the coding of the tangent- linear and adjoint versions of the Eulerian spectral and adiabatic HIRLAM and the application of these models to forecast error sensitivity experiments (Gustafsson and Huang, 1996). The development of tangent-linear and adjoint versions of the HIRLAM non-linear physical parameterization schemes started in parallel (Yang, 2002). It turned out, however, that the application of these physical parameterization schemes required un-realistic large computational resources, due to the serial application of physics in HIRLAM. This implies extensive re-calculations of non-linear physics within the tangent-linear and adjoint schemes. The treatment of observations and the background error constraint in HIRLAM 4D-Var are shared with the 3-dimensional variational data assimilation (3D-Var). Once HIRLAM 3D-Var was ready for operational implementation (Gustafsson et al., 2001, Lindskog et al., 2001) it was also possible to carry out the first tests of HIRLAM 4D-Var (Huang et al., 2002).

In order to make HIRLAM 4D-Var computationally affordable, and to improve the performance as well, several new components were developed and tested over the period 2000-2004:

- The HIRLAM 4D-Var was made fully incremental, such that the minimization is carried out for the assimilation increment only at a lower resolution than in the non-linear model and, furthermore, the tangent linear model is applied for the forward propagation of the assimilation increment.
- Packages for simplified physics, including tangent-linear and adjoint versions were imported from Meteo-France (Janiskova, 1997) and ECMWF (Buizza, 1994).
- The SETTLS (Stably Extrapolating Two Time-level Semi-Lagrangian) scheme was introduced in the tangent-linear, adjoint as well as non-linear spectral HIRLAM.
- The possibility for multi-incremental minimization, with re-calculation of the non-linear model trajectory in an outer loop, was introduced. It was also made possible to optionally select the grid-point or spectral HIRLAM for the outer loop trajectory calculation.

The developments mentioned above were ready in late 2003 and extensive parallel tests and comparisons between HIRLAM 3D-Var and HIRLAM 4D-Var were carried out in 2004-2005. The results from these tests were quite disappointing, however, with a neutral or even negative impact of 4D-Var in comparison with 3D-Var. It was not until early 2006, when a serious coding error was found in the adjoint version of the SETTLS scheme was detected and corrected for, that the results of 4D-Var started to out-perform the results of 3D-Var.

This report will describe the results from some recent tests of HIRLAM 4D-Var, including results from application of the weak digital filter constraint (Gustafsson, 1992) as well as the control of lateral boundary conditions, the latter being a new component of HIRLAM 4D-Var.

## 2 Flow-dependency in HIRLAM 4D-Var

One example that illustrates the potential of implicit treatment of flow-dependency in 4D-Var, also with short assimilation windows (6 hours), will be shown. The example is taken from application of the HIRLAM 4D-Var to the meso-scale storm that hit Denmark on the 3 December 1999. A *single simulated observation experiment* was carried out. The question asked was: What would 3D-Var and 4D-Var do in case we had a single surface pressure observation available, telling us that the surface pressure in the center of the storm ought to be 5 hPa deeper?

A simulated surface pressure observation with a -5 hPa observation increment was thus inserted at 3 December 1999 12 UTC into the data assimilation in the position 57N 3E, in the center of the storm. Fig. 1, left, shows the surface pressure assimilation increment in case HIRLAM 3D-Var is applied. The assimilation increments simply reflect the homogenous and isotropic 3D-Var structure functions on a large spatial scale, describing average surface pressure forecast errors.

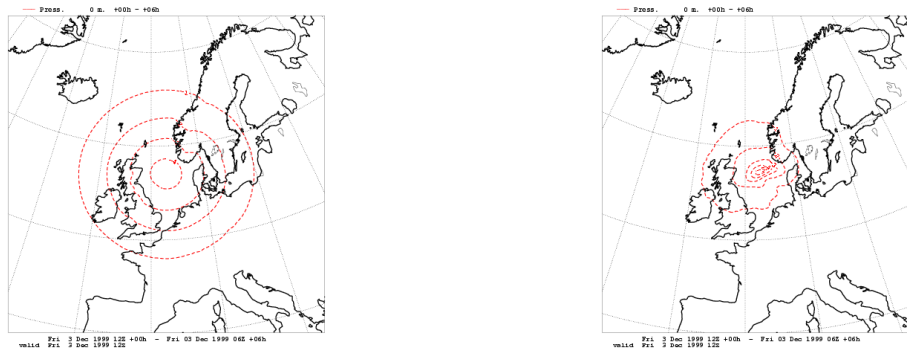


Figure 1: Mean sea level pressure assimilation increments from a single simulated surface pressure observation in the position 57N 3E at 3 December 1999 12 UTC. Left: 3D-Var. Right: 4D-Var was applied over the assimilation window 3 December 1999 06–12 UTC

To illustrate the effect of the implicit treatment of flow-dependency in 4D-Var, the same simulated surface pressure observation was inserted into a 4D-Var assimilation, with the assimilation window starting at 3 December 1999 06 UTC and ending at 12 UTC. Fig. 1, right, shows the 4D-Var surface pressure assimilation increments valid at 12 UTC. We can notice that the spatial scale of the surface pressure increments has shrunk significantly, roughly corresponding to the scale of the storm itself, as given by observations. Furthermore, the increments no longer have the isotropic horizontal structure of typical 3D-Var increments.

How are these flow-dependent assimilation increments achieved in 4D-Var? First of all, it can be shown that 4D-Var is equivalent to a full rank Extended Kalman Filter (EKF) over the data assimilation window, with the important limitation that the covariance structures at the start of the assimilation window are just the static (and robust) ones applied in 3D-Var. From this we can conclude that it is the application of the tangent linear model, linearized around a trajectory calculated by the full non-linear model, that provides the flow-dependency 6 hours later into the data assimilation window. The assimilation increments at the start of the assimilation window were also investigated. The 4D-Var surface pressure increments at 3 December 06 UTC turned out to be very small (not shown). Fig. 2 shows a NW-SE cross-section of upper-air temperature and wind increments, centered at 55N 0EW in an area upstream of the storm development 6 hours later. We may notice an increase of the vertical wind shear as well as a slight vertical tilt in the assimilation increments at the start of assimilation window. Thus we can simply conclude that 4D-Var manages to intensify the storm development by increasing the degree of baroclinicity in the model state at the start of the assimilation window, and this provides a faster

growth of the storm during the tangent linear propagation of the assimilation increments up to the 12 UTC, the time of the inserted simulated observation.

Considering the assimilation increments at the start of the assimilation window, for example those illustrated in Fig. 2, it needs to be mentioned again that these are heavily constrained by the applied isotropic and homogeneous structure functions.

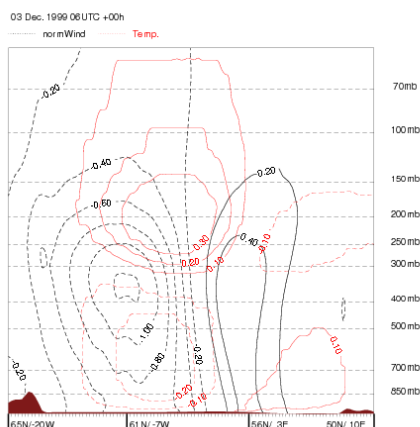


Figure 2: Vertical cross-section in the NW-SE direction, centered at 55N 0EW, showing 4D-Var upper-air temperature and wind assimilation increments valid at 3 December 1999 06 UTC. A single simulated surface pressure observation was inserted in the position 57N 3E at 3 December 1999 12 UTC. 4D-Var was applied over the assimilation window 3 December 1999 06–12 UTC

### 3 A 3 months comparison of HIRLAM 3D-Var and 4D-Var

A test to compare the performance of HIRLAM 3D-Var and HIRLAM 4D-Var was carried out for 3 months, January 2005, June 2005 and January 2006. The integration domain was one of the operational SMHI HIRLAM domains,  $306 \times 306$  horizontal grid-points in 22 km resolution and 40 vertical levels. The operational observation data files from SMHI, including AMSU-A radiances, were applied in 3D-Var as well as 4D-Var. Forecast up to +48 h were calculated using the operational ECMWF forecast data on the lateral boundaries. Some further characteristics of the 3D-Var and 4D-Var setups for the comparison runs were:

- A 6 hour assimilation cycle was used in 3D-Var as well as in 4D-Var. 3D-Var used FGAT (First Guess at Appropriate Time) with observation minus background differences calculated every hour in a  $\pm 3$  hour assimilation window centered around the nominal analysis time. 4D-Var used a 6 hour assimilation window, also centered around the nominal analysis times, and 1 hour observation windows.
- 3D-Var increments were calculated on a 44 km grid with a quadratic spectral truncation, while 4D-Var increments were calculated on a 66 km grid with a linear spectral truncation. Thus the spectral resolution was the same for 3D-Var and 4D-Var, with a shortest resolved wave-length of 132 km.
- The statistical balance formulation of the background error constraint, with statistics calculated from archives of operational SMHI HIRLAM 22 km forecast, was applied in 3D-Var as well as in 4D-Var.

- The tangent-linear and adjoint versions of the vertical diffusion and large scale condensation schemes, from the Meteo-France package for simplified physics, were applied in 4D-Var.
- The analysis at the nominal analysis time in 4D-Var was calculated by a non-linear forecast from the analysis at the start of the 4D-Var assimilation window, for example, by a non-linear forecast from 09UTC to 12UTC in case of the 12UTC assimilation cycle.
- The tangent linear version of the (non-linear) normal mode initialization was applied at the start of the tangent linear model integration and the adjoint version was applied at the end of the adjoint model integration during the 4D-Var minimizations.

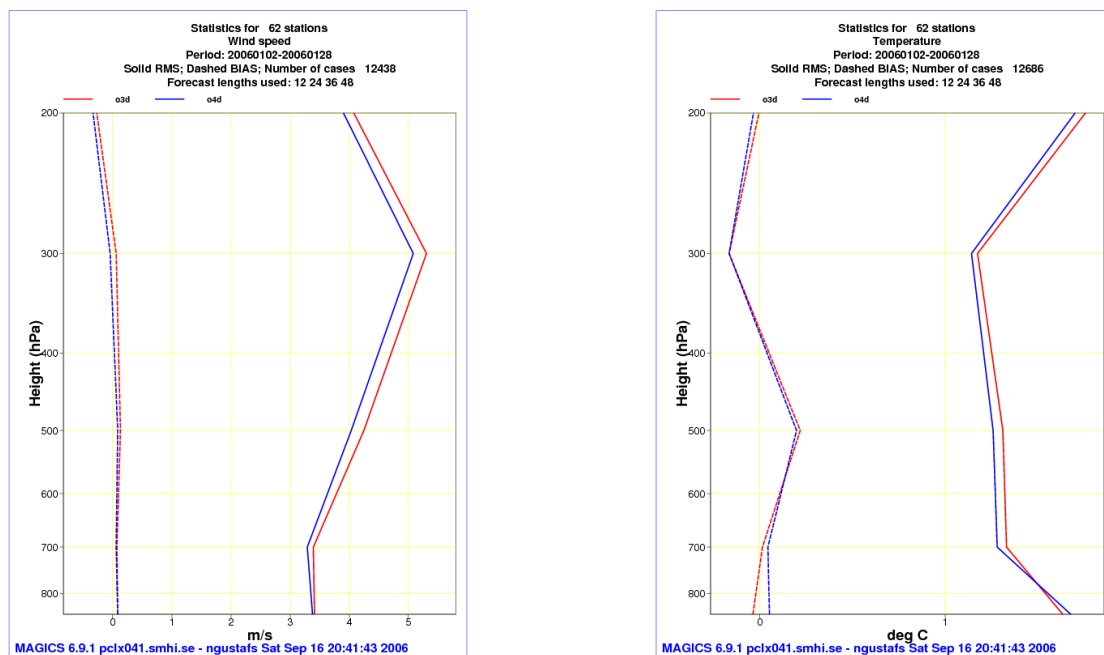


Figure 3: RMS (Root Mean Square) and Bias forecast verification scores from comparison between forecasts and radiosonde data for experiment **o3d**, with 3D-Var assimilation, and **o4d**, with 4D-Var assimilation. Average calculated for +12h, +24h, +36h and +48h forecasts 3–29 January 2006. Wind speed (left) and temperature (right).

The forecasts up to +48h were verified against radiosonde and SYNOP data from a European network of stations (the EWGLAM list of stations). The forecast scores for all three months were very similar and showed a significant positive impact of 4D-Var in comparison with 3D-Var. We will here show verification scores for January 2006 only. Fig. 3 shows vertical profiles of verification scores for wind speed (left) and temperature (right), averaged over forecast lengths +12h, +24h, +36h and +48h. The positive impact of 4D-Var in comparison with 3D-Var is clearly illustrated, in particular for the upper troposphere.

The forecast verification scores for mean sea level pressure were similarly significantly lower for the forecasts based on 4D-Var assimilation in comparison with the forecasts based on 3D-Var. Fig. 4 shows the verification scores for different forecast lengths averaged for January 2006 and Fig. 5 shows the daily verification scores for the +24h mean sea level pressure forecasts in January 2006. We may notice the tendency of 4D-Var to improve the forecast scores in cases of relatively poor forecasts based on 3D-Var.

A limitation with the Root Mean Square (RMS) verification score is that it may favor a more smooth forecast in comparison with a more detailed forecast, with the details slightly misplaced in space or time. Thus, it is necessary to complement the RMS forecast verification scores with some measure of the variability of the

forecast. In the table below we simply present the standard deviations and the mean values of mean sea level pressure forecasts, as produced from the 3D-Var and 4D-Var initial data, together with the same values given by SYNOP observations, for all three months of the experiments. It is obvious that with 4D-Var we, in addition to lower values of RMS verification scores, also provide mean sea level pressure forecasts with a variability closer to the observed one, at least as measured by the standard deviation.

Table 1: Root Mean Square errors of +24 h mean sea level pressure forecasts based on 3D-Var and 4D-Var, monthly mean values and standard deviations of mean sea level pressure forecasts, based on 3D-Var and 4D-Var, evaluated in SYNOP station positions in comparison with the same statistics for SYNOP observations.

|                              | January 2005 | June 2005 | January 2006 |
|------------------------------|--------------|-----------|--------------|
| RMS Error 3D-Var             | 1.80         | 1.17      | 1.64         |
| RMS Error 4D-Var             | 1.66         | 1.16      | 1.54         |
| Mean Observations            | 1016.9       | 1015.8    | 1025.3       |
| St.dev. Observations         | 13.9         | 6.0       | 11.3         |
| Mean 24 h forecast 3D-Var    | 1016.5       | 1015.6    | 1025.0       |
| St.dev. 24 h forecast 3D-Var | 14.2         | 6.4       | 10.8         |
| Mean 24 h forecast 4D-Var    | 1016.6       | 1015.6    | 1025.1       |
| St.dev. 24 h forecast 4D-Var | 13.8         | 6.2       | 10.9         |

## 4 A weak digital filter constraint

A weak digital filter constraint has been introduced into the HIRLAM 4D-Var. The formulation follows Gustafsson (1992) and Gauthier and Thépaut (2001). An new cost function term  $J_c$  is added to the 4D-Var cost function:

$$J_c = \frac{\gamma_{df}}{2.0} (\delta X_{N/2} - \delta X_{N/2}^{df})^T C^{-1} (\delta X_{N/2} - \delta X_{N/2}^{df}) \quad (1)$$

with

$$\delta X_{N/2} - \delta X_{N/2}^{df} = \delta X_{N/2} - \sum_{n=0}^N f_n \delta X_n = \sum_{n=0}^N h_n \delta X_n \quad (2)$$

where  $\delta X_n$  is the model state assimilation increment at time-step  $n$ , calculated from the initial state assimilation increment  $\delta X_0$  by the tangent linear model  $\delta X_n = M_n \delta X_0$ ,  $N$  is the number of timestep over the data assimilation window and  $\delta X_{N/2}^{df}$  is the digitally filtered assimilation increment at the mid-point of the data assimilation window. Note that  $N$  must be an even number. The parameter  $\gamma_{df}$  we will consider a tuning-parameter, describing the relative importance of the noise filtering constraint  $J_c$  in comparison with the observation error constraint  $J_o$  and the back-ground error constraint  $J_b$ . The diagonal matrix  $C^{-1}$  defines the relative weights given to different model variables in the weak digital filter constraint and we have defined these in accordance with the total energy norm.  $f_n$ , finally, are the digital filter weights. We have determined these in accordance with the *Dolph filter* (Lynch, 1997), which is defined by a time span (= the length of the data assimilation window = 5 hours in our case) and a stop band edge  $T_c$ .

Using the reformulated digital filter weights ( $h_n, n = 0, \dots, N$ ) the digital filter cost function  $J_c$  and its gradient with respect to the initial time assimilation increment may be calculated as follows:

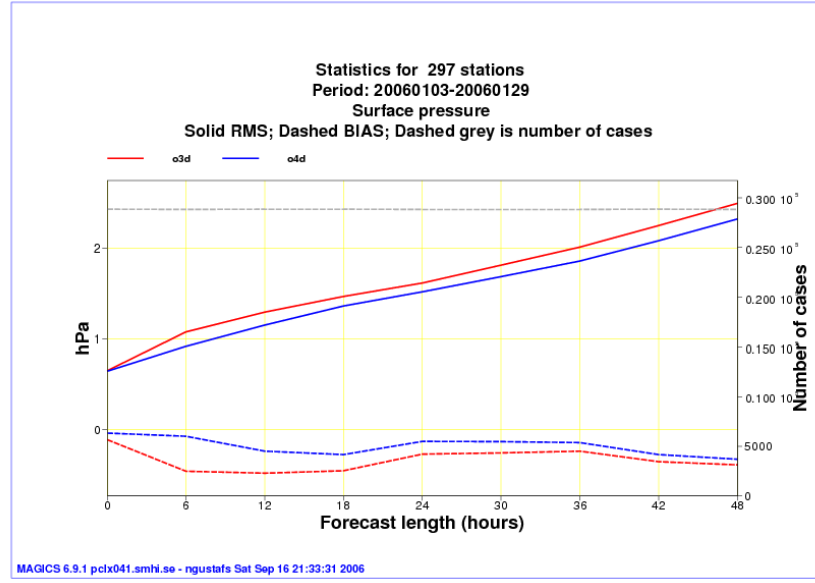


Figure 4: RMS (Root Mean Square) and Bias mean sea level pressure forecast verification scores from comparison between forecasts and SYNOP data for experiment **o3d**, with 3D-Var assimilation, and **o4d**, with 4D-Var assimilation. Average scores are calculated for forecasts from 3–29 January 2006.

$$J_c = \frac{\gamma_{df}}{2.0} \left( \sum_{n=0}^N h_n \delta X_n \right)^T C^{-1} \left( \sum_{n=0}^N h_n \delta X_n \right) \quad (3)$$

and

$$\nabla_{\delta X_0} J_c = \sum_{k=N}^0 \gamma_{df} M_k^T h_k C^{-1} \left( \sum_{n=0}^N h_n \delta X_n \right) \quad (4)$$

From the expression for the gradient  $\nabla_{\delta X_0} J_c$  we can notice that the deviations from the digitally filtered model assimilation increment will enter as a *forcing* for the adjoint model equations ( $M_k^T$ ), similarly to the way the deviations from the observations (as defined by  $J_o$ ) also enters as a forcing to the adjoint model equations.

A number of experiments to test the sensitivity of the effect of the weak digital filter constraint with regard to the general tuning parameter  $\gamma_{df}$ , with regard to the stop band edge  $T_c$  and with regard to the relative weighting  $C^{-1}$  between different model variables, have been carried out. We have also checked the behaviour of the digital filter constraint when different versions of the back-ground error constraint  $J_b$  are applied. This is motivated since the balance between different model variables is partly explicitly defined through the background error constraint and implicitly via the weak digital filter constraint and the tangent linear model equations.

All experiments were carried for the operational SMHI HIRLAM area with 22 km horizontal resolution, with  $306 \times 306$  horizontal grid-points and with 40 levels. 4D-Var is carried out over the 5 hour assimilation window 3UTC - 8UTC 1 December 1999 with a horizontal increment resolution of 44 km. The non-linear spectral HIRLAM model is used for the trajectory calculations as well as for evaluation of the noise characteristics of the initial data produced by 4D-Var.

After some initial trial-and-error experiments, the following values of the digital filter parameters turned out to provide a reasonable balance between the level of noise in forecasts based on initial data produced by 4D-Var and the level of damping of the assimilation increment amplitude needed to obtain the damping of the noise:  $\gamma_{df} = 1.0$ ,  $T_c = 3$  hours,  $C_u = C_v = (3m/s)^2$ ,  $C_T = (1K)^2$ ,  $C_{ln(p_s)} = (0.01 * \sqrt{nlev})^2$  and  $C_q = (1g/kg)^2$ .

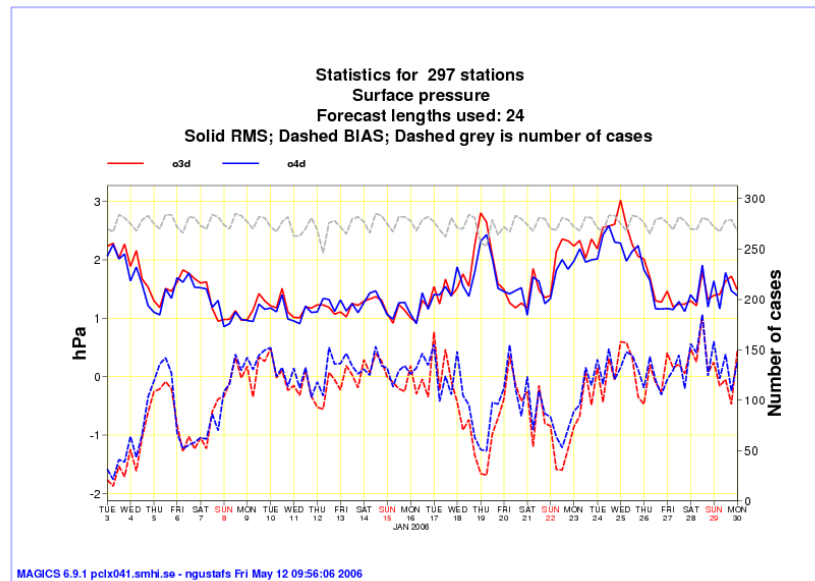


Figure 5: RMS (Root Mean Square) and Bias mean sea level pressure +24h forecast verification scores from comparison between forecasts and SYNOP data for experiment **o3d**, with 3D-Var assimilation, and **o4d**, with 4D-Var assimilation. Time series of scores for 3–29 January 2006 00UTC, 06UTC, 12UTC and 18UTC.

We compared in a first series of experiments 4D-Var with analytical structure functions and with 3 different settings for gravity wave (noise) control: (1) NOINIT: Without any gravity wave control; (2) NMI: The tangent linear version of the non-linear normal mode initialization is applied at the start of the tangent linear model integration and its adjoint is applied at the end of the adjoint model integration; (3) JCDFI: The weak digital filter constraint with the standard settings as described above is applied.

Fig.6 illustrates the evolution of the observation contribution to the total cost function during the minimization iterations for the three different 4D-Var experiments (left) and the time evolution of the spatial mean absolute surface pressure tendency in forecasts issued from the initial data obtained from the three different 4D-Var runs (right). We may notice that the evolution of the observation contribution to the cost function is quite similar for the three experiments, but with a slightly lower level of the cost function at the end of the minimization for the experiment with no gravity wave constraint. This is possible to understand since without any gravity wave constraint, 4D-Var has more freedom to utilize also higher frequency oscillations to fit the observed information. We may also notice that 4D-Var with the weak digital filter constraint requires slightly less number of iterations to obtain the minimum as defined by the minimization iteration stop criteria. From the noise characteristics of the forecasts issued from the different initial data sets, we may conclude that 4D-Var with the strong NMI constraint only marginally manages to reduce the noise compared to 4D-Var without any gravity wave control. 4D-Var with the weak digital filter constraint and the standard setting, on the other hand, obviously manages to reduce the noise in a significant way. Generally speaking, one may see from the figures that application of gravity wave (noise) constraints in 4D-Var marginally worsens the fit of the analysis to the observations and reduces the amplitude of the assimilation increments in consequence with this.

The other sensitivity experiments carried out with the weak digital filter constraint (results are not shown here) mainly supported the standard parameter settings described above. It was also shown that it was not needed to apply the weak digital filter constraint to the humidity field and that application of the weak digital filter constraint together with the statistical balance background constraint did not introduce any further problems.

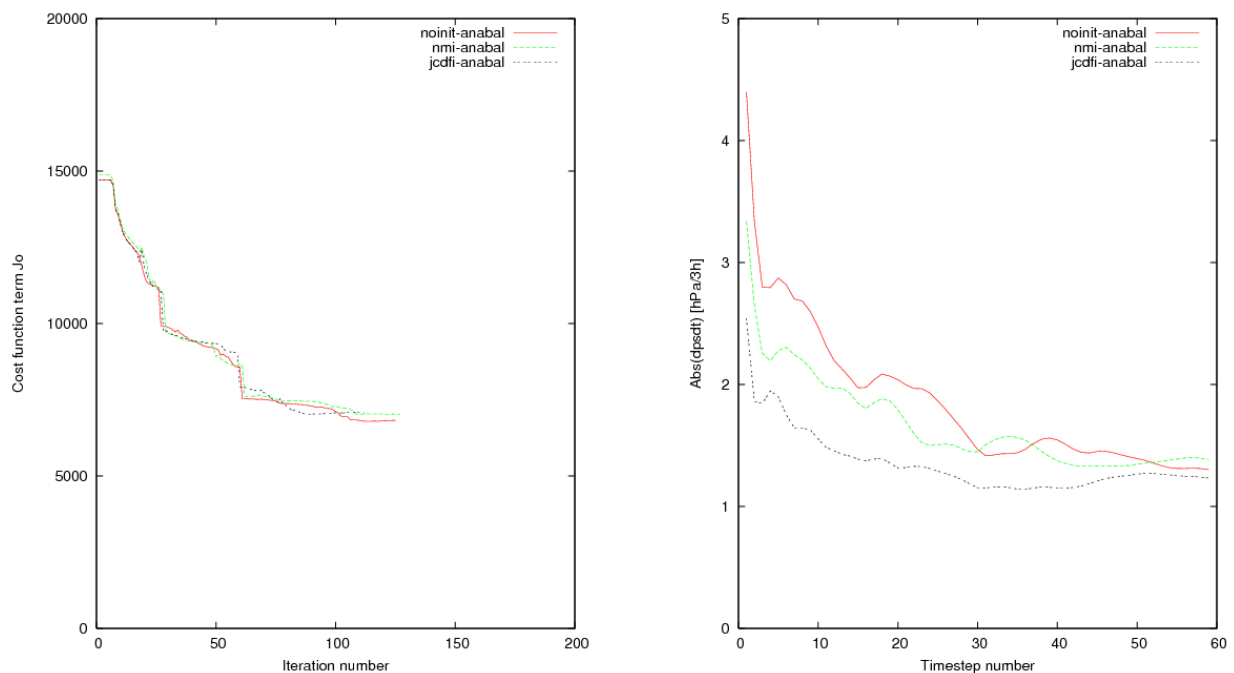


Figure 6: Evolution of the observation contribution to the total cost function with iteration number (left) and time evolution of the spatial mean of the absolute surface pressure tendency in the forecast (right) for three different 4D-Var experiments. NOINIT: without any gravity wave constraint; NMI: with a strong normal mode initialization constraint; JCDFI: with a weak digital filter constraint and the standard parameter setting as described in the text.

## 5 Control of lateral boundary conditions

With the HIRLAM 3D-Var, it is possible to obtain assimilation increments all the way out to the lateral boundaries. This was made possible through the use of full Fourier transforms and the area extension zone (see Gustafsson et al., 2001). The control of lateral boundary conditions in HIRLAM 4D-Var can be considered even more important than in 3D-Var, since inflow areas in the non-linear model integration, with regard to the lateral boundaries, become outflow areas (of error gradient information) in the adjoint model integration. Following the experience at the Japanese Meteorological Agency (JMA) to control also the lateral boundaries in a LAM 4D-Var, such a possibility has now been introduced in HIRLAM 4D-Var. The JMA formulation is very simple. A new control vector is simply introduced to represent increments to the lateral boundaries at the end of the assimilation window. The full model state over the complete model domain is used as the control vector, i.e. similar to the initial data control vector, and a lateral boundary error constraint, also similar to the back-ground error constraint, is introduced in order to make the lateral boundary perturbations smooth and balanced. With these simplifications, the modifications of the 4D-Var code for control of lateral boundary conditions become straightforward.

The control of lateral boundary in HIRLAM 4D-Var is yet not fully tested and validated. Here we will only show one example of the impact of the control of lateral boundaries on the analysis fields in HIRLAM 4D-Var. Fig. 7 shows the difference between mean sea level pressure analysis fields in the middle of the data assimilation window when control of lateral boundary conditions is applied, respective not applied, in HIRLAM 4D-Var. This example is taken from the first data assimilation cycle applied with control of lateral boundary conditions. The impact is expected to increase when more assimilation cycles with control of lateral boundary conditions

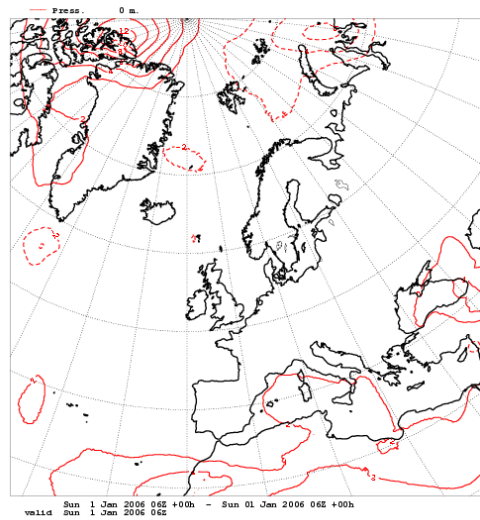


Figure 7: Differences between mean level pressure analysis fields obtained in the middle of the data assimilation window when control of lateral boundary conditions is applied, respective not applied, in HIRLAM 4D-Var.

are applied.

## 6 Computational cost of 4D-Var

Most of the tests of HIRLAM 4D-Var were carried out on the SMHI LINUX-cluster DUNDER with Dual Intel Xeon 3.4 GHz processors, 2Gb mem/node and an Infiniband interconnect. 13 nodes (26 processors) were used for the tests.

A representative example from 3 January 2006 12UTC gave the following wall clock timings:

- 48 h non-linear forecast (22 km, 7.5 minutes timestep): 1005 seconds
- 66 km resolution 4D-Var, 30 minutes timestep, 82 iterations (88 simulations) : 1053 seconds
- 44 km resolution 4D-Var, 15 minutes timestep, 90 iterations (97 simulations) : 3971 seconds

With 3 times the resolution (66 km) in 4D-Var as in the forecast model (22 km), 4D-Var costs computationally like a 48 h forecast.

## 7 Remaining problems and issues

There are still a number of issues to be studied for the HIRLAM 4D-Var, for example:

- Some minimization convergence problems have occurred and these need to be understood.
- We need to evaluate possible benefits of higher resolution increments. Preliminary evaluations have indicated that a 4D-Var increment resolution 3 times that of the non-linear model is sufficient, but this needs to be further studied.

- We need to design an optimal multi-incremental scheme.
- We need to evaluate the effect of controlling the lateral boundaries.
- We need to improve pre-conditioning of the minimization and the possible code speedup.
- We need to evaluate further how moist processes are handled in HIRLAM 4D-Var.
- We need to do a proper tuning of the background and observation error standard deviations for 4D-Var.
- We need to examine how 4D-Var performs also when a mixing with large scale information from a global model is applied in the data assimilation cycle.

## 8 Concluding remarks and plans

From the tests carried out so far we may conclude that HIRLAM 4D-Var now is prepared for near real time tests. Can we afford it operationally? Yes! 4D-Var provide significantly improved forecast scores compared to 3D-Var for synoptic scales and dynamical forecast variables. Furthermore, we need to look further into some minimization problems, the handling of gravity wave noise, moist processes and other issues mentioned above.

## 9 References

- Buizza, R. 1994: *Impact of simple vertical diffusion and of the optimisation time on optimal unstable structures*. ECMWF Tech. Memo. 192, 25pp. Available from the European Centre for Medium Range Weather Forecasting, Shinfield Park, Reading, Berks. RG2 9AX, UK.
- Gauthier, P. and J.-N. Thépaut, 2001: Impact of the Digital Filter as a Weak Constraint in the Preoperational 4DVAR Assimilation System of Météo-France. *Mon. Wea. Rev.*, **129**, 2089-2102.
- Gustafsson, N., 1992: Use of a digital filter as weak constraint in variational data assimilation. Workshop proceedings on *Variational assimilation, with special emphasis on three-dimensional aspects*. pp 327-338. Available from the ECMWF.
- Gustafsson, N. and Huang, X.-Y. 1996: Sensitivity experiments with the spectral HIRLAM and its adjoint. *Tellus*, **48A**, 501-517.
- Gustafsson, N., Berre, L., Hörnquist, S., Huang, X.-Y., Lindskog, M., Navascués, B., Mogensen, K.S. and S. Thorsteins-son, 2001: Three-dimensional variational data assimilation for a limited area model. Part I: general formulation and the background error constraint. *Tellus*, **53A**, 425-446.
- Huang, X.-Y., Yang, X., Gustafsson, N., Mogensen, K. and M. Lindskog, 2002: Four-dimensional variational data assimilation for a limited area. HIRLAM Technical Report, **No. 57**, pp 44, Available from SMHI.
- Janisková, M., Thépaut, J.-N. and Geleyn, J.-F., 1997: Simplified and regular physical parameterizations for incremental four-dimensional variational assimilation. *Mon. Wea. Rev.*, **127**, 26-45.
- Lindskog, M., Gustafsson, N., Navascués, B., Mogensen, K.S., Huang, X.-Y., Yang, X., Andrae, U., Berre, L., Thorsteins-son, S. and Rantakokko, J., 2001: Three-dimensional variational data assimilation for a limited area model. Part II: Observation handling and assimilation experiments. *Tellus*, **353A**, 447-468.
- Lynch, P., 1997: The Dolph-Chebyshev window: A simple optimal filter. *Mon. Wea. Rev.*, **125**, 655-660.
- Yang, X., 2002: Physical adjoint in HIRLAM 4DVAR. *Pages 50-57 of: HIRLAM Workshop on Variational Data Assimilation and Remote Sensing, 21-23 January 2002, Finnish Meteorological Institute, Helsinki, Finland. Available from HIRLAM-5, c/o Per Undén, SMHI, S-60176 Norrköping, Sweden.*