# BACKGROUND ERRORS IN HIRLAM 4D-VAR


**M. Lindskog** (1), N. Gustafsson (1), S. Thorsteinsson (2), T. Landelius (1), P. Dahlgren (1) and P. Undén (1)

(1) Swedish Meteorological and Hydrological Institute, Norrköping, Sweden

(2) Icelandic Meteorological Office, Reykjavik, Iceland


## HIRLAM 4D-VAR

A limited area 4-dimensional variational data assimilation (4D-VAR) system has been developed within the international HIRLAM co-operation. The HIRLAM variational data assimilation system has an incremental formulation and the assimilation consists of finding the model state increments, δx (= [δu δv δT δq δlnps]$^T$), that minimize the following cost function:

$$J = J_b + J_o + J_c + J_k = \frac{1}{2}\delta x'^T B^{-1}\delta x' + \frac{1}{2}(Hx^b + H_{tl}\delta x' - y)^T R^{-1}(Hx^b + H_{tl}\delta x' - y) + J_c + J_k$$

Here $J_B$ measures the distance to a background model state $x^b$, which is a short-range forecast, and $J_o$ measures the distance to the vector $y$ of the observations. The non-linear observation operator H and the tangent-linear observation operator $H_{tl}$ transform the background state and assimilation increments, respectively, into the observed quantities. $B$ is the matrix containing the covariances of the background field errors, while $R$ is a matrix containing the covariances of the errors in the observations. The primes indicate that the assimilation increments may be applied at a lower horizontal resolution than the background field. Furthermore, $^T$ denotes transpose, or adjoint, operation. $J_c$ represents a weak digital filter constraint and $J_{k,}$ a large scale constraint, still under development. On-going developments include also a control of lateral boundary conditions and a representation of model errors.

## SMHI operational HIRLAM 4D-VAR

Since 30 January 2008 HIRLAM 4D-Var is running operationally at the Swedish Meteorological and Hydrological Institute (SMHI). The assimilation is run with a 6 h assimilation cycle and the length of the assimilation time window is 6 h. The horizontal resolution of the forecast model is 22 km and 40 vertical levels are used. SMHI HIRLAM forecasts are run up to 48 h, with forecasts from ECMWF (European Centre for Medium-Range Weather Forecasts) used as lateral boundaries, over the SMHI model domain (Figure 1). The horizontal resolution of the assimilation increments is 66 km.

*Fig. 1. SMHI operational HIRLAM 4D-VAR area.*

## Background error covariances

The background error covariances of the SMHI operational HIRLAM 4D-Var were obtained by calculating statistics of a large set of differences between 12 h and 36 h HIRLAM forecasts, valid at the same time (the so-called "NMC-method"). Also ensemble of 10 parallel assimilation cycles, utilising perturbed observations and perturbed lateral boundary conditions, have been used for deriving background error statistics. The derivation of the structure functions utilizes a statistical balance formulation. The horizontal variation of background error standard deviations may be represented by a horizontally varying climatological index-field. Alternatively the horizontal and temporal variation of background error standard deviatons may be based on a normalised, as well as temporally and spatially filtered, Eady baroclinicity index ($\sigma_{BI}$), calculated from the background field. The Eady index is given by:

$$\sigma_{BI} = 0.31\left|\frac{\partial V}{\partial z}\right| N^{-1}$$

where $|\partial V/\partial z|$ is the vertical wind shear and $N^{-1}$ is the bouyancy frequency. The climatological and the Eady based index field for one particular case are illustrated in Figure 2.

In close collaboration with ALADIN the introduction of a complex wavelet transform is under development to represent the background error correlations. In the future statistics of real time forecast differences calculated from an ensemble of short range forecasts could be transformed into wavelet space, to obtain synoptically varying structure functions. To demonstrate this potential, statistics of 12 h minus 36 h forecasts valid at the same time were calculated for a number of 3 day periods (resulting in a sample of 12 cases of differences for each period), characterised by different synoptic flow situations. Figure 3 illustrates the 500 hPa geopotential height in the middle of two different periods as well as the resulting wavelet representation of 500 hPa temperature derived background error correlations.

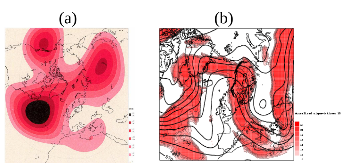


*Fig. 2. Horizontal index field (red shading) representing background error variations: based on climatology, (a) based on Eady index for 2 January 2002, 18 UTC (b). Full black lines illustrate the 500 hPa geopotential height (gpm) times 0.01 of the back-ground state.*

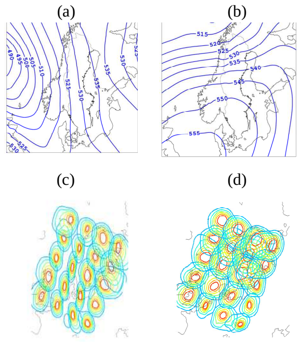


*Fig. 3. 500 hPa geopotential height (gpm) times 0.01 for 20060117 12 UTC (a) and 20060123 12 UTC (b). Wavelet representation of the horizontal corre-lations of 500 hPa temperature errors for the periods 20060116-20060118 (c) and 20060122-20060124 (d).*

## Moisture control variable

The reference HIRLAM humidity assimilation uses specific humidity increments δq in the control vector. The new version of the HIRLAM data assimilation uses a normalised relative humidity variable δRH* accounting better for the wide variation in humidity in the vertical and at small scales in the horizontal (Elías Hólm *et. al.*, 2002: Assimilation and modelling of the Hydrological cycle: ECMWF´s status and plans. *ECMWF Tech. Memo. 383*). The new control variable results in slightly flow dependent humidity structure functions (Fig. 4a and b) and in less supersaturation in full scale assimilation (Fig. 4c and d).

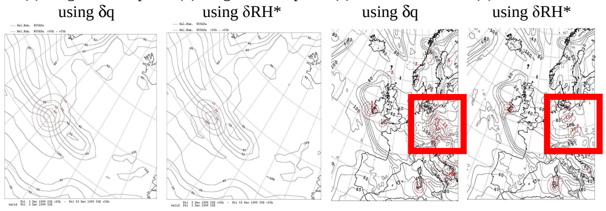


*Fig. 4. More flow dependent humidity structure function in single observation experiment (b) than (a), and less supersaturation in full scale assimilation (d) than (c). The full scale assimilation experiment is for 1 January 2002, 06 UTC. Analysis (black), increments (red). Unit: %.*

## Large scale constraint

The large scale information from the ECMWF global model that provides lateral boundaries can be used as a constraint $J_k$ in data assimilation (Guidard and Fischer, 2006: Introducing the coupling information in a limited-area variational assimilation, *QJR Meteorol. Soc.*, **132,** 1-20.)

$$J = \underbrace{J_b}_{background} + \underbrace{J_o}_{observation} + \underbrace{J_c}_{W.C-digital-filter} + \underbrace{J_k}_{large-scale}$$

The $J_k$ term measures the difference between the HIRLAM analysis and a 6 h ECMWF forecast.

$$J_k = (\mathbf{d}_k + \mathbf{H}_2\delta\mathbf{x})^T\mathbf{V}^{-1}(\mathbf{d}_k + \mathbf{H}_2\delta\mathbf{x})$$
$$\mathbf{d}_k = \mathcal{H}_2(\mathbf{x}_b) - \mathcal{H}_1(\mathbf{x}_{ls})$$

$x_{ls}$=ECMWF $+6h$ forecast
$\mathcal{H}_1$= Interpolate and truncate global field
$x_b$=HIRLAM first guess
$\mathcal{H}_2$= Truncate HIRLAM first guess
$\mathbf{H}_2$=Tangent linear of $\mathcal{H}_2$
$\mathbf{V}$=Error statistics for $\mathcal{H}_1(\mathbf{x}_{ls})$
$\delta\mathbf{x} = \mathbf{x} - \mathbf{x}_b$ increments in model space

We have as a first step calculated error statistics from ECMWF and HIRLAM ensemble assimilation experiments. The ECMWF fields were interpolated to the same grid as the HIRLAM data before the statistics was calculated. A comparison (Figure 5) between background (red) and large scale (blue) error statistics shows that the ECMWF difference field contains less energy than the HIRLAM difference field for the shorter scales. The ECMWF difference field also contains less energy than the HIRLAM difference field for larger scales - whether this is a result of larger error growth in HIRLAM, or simply due to details in perturbation techniques, needs some further attention. The horizontal resolution is 22 km for HIRLAM and 55 km (T319) for ECMWF field.

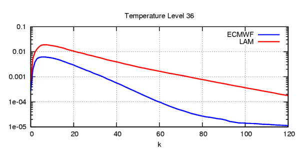


*Fig. 5. Spectral densities for (unbalanced) temperature. ECMWF and HIRLAM at model level 36 (close to 1000 hPa).*


Contact: Magnus Lindskog
SMHI
SE-601 76 Norrköping Sweden
Email: Magnus.Lindskog@smhi.se