THE NUMERICAL SCHEME AND LATERAL BOUNDARY CONDITIONS FOR THE SPECTRAL HIRLAM AND ITS ADJOINT.

Nils Gustafsson

Swedish Meteorological and Hydrological Institute

1 INTRODUCTION

Haugen and Machenhauer (1993) developed a spectral integration technique for limited area models based on an extension of the integration area to obtain periodic variations in both horizontal directions, and they also applied the technique successfully to a shallow water model. The same technique was implemented for the full 3-dimensional primitive equations by Gustafsson (1991). The same basic forecast model equations, the same vertical hybrid pressure/sigma coordinate and the same physical parameterization schemes as in the gridpoint HIRLAM (Källén 1996) are utilized. All basic variables involved in the dynamical model equations are represented by the complex-valued coefficients of truncated 2-dimensional Fourier series. An elliptic truncation in wavenumber space is utilized to ensure an isotropic representation. Since calculations of non-linear terms are carried out in grid-point space, Fast Fourier Transforms (FFT) are utilized for the transformations between gridpoint space and spectral space.

The spectral HIRLAM includes a three time level Eulerian semi-implicit time integration scheme as well as a two time level semi-Lagrangian semi-implicit time integration scheme. A comparison between the gridpoint and the spectral semi-Lagrangian HIRLAM models was carried out by Gustafsson and McDonald (1996). They concluded that both model formulations had improved from the model comparison exercise, that both model versions produce equally good forecasts for the cases studied and that the computer costs for the two model formulations are almost identical. The spectral HIRLAM has been coded for distributed memory parallel computers using a message passing technique (Gustafsson and Salmond 1994). This parallel version of the spectral HIRLAM has been utilized as the model framework for development of a variational data assimilation for HIRLAM (Gustafsson and Huang 1996, Gustafsson et al., 1999).

The aim with these lecture notes is to present a fairly complete description of the Eulerian version of spectral HIRLAM and to discuss the use of this model as a framework for development of a variational data assimilation system for HIRLAM.

2 SPECTRAL FORMULATION OF HIRLAM

2.1 Area extension to obtain bi-periodic variations

The geometry of using an extension zone to obtain periodic variations in both horizontal dimensions is illustrated in Figure 1. Assume that we want to represent initial and lateral boundary fields, given in the gridpoints of the inner integration area only, by bi-periodic Fourier series. To do so, extrapolated gridpoint values are determined in the extension zone in such a way that the subsequent Fourier transforms will give a smooth representation in the inner area with preserved normal gradients along the lateral boundaries. This is achieved by an extrapolation, first along each row of gridpoint values in the x-direction and then along each column of gridpoint values in the extended area in the y-direction.
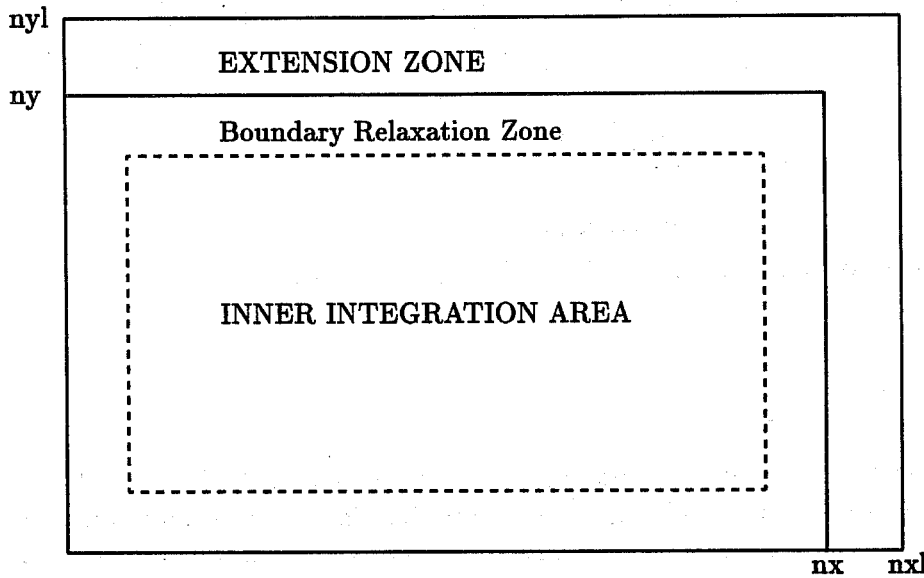
Figure 1: Geometry of the extension and boundary relaxation zones

The extrapolation is carried out with the constraint that the normal gradients and the absolute values are preserved at the boundaries of the inner integration area. To achieve best possible consistency with the spectral Fourier series representation, sine- and cosine-functions are utilized for the extrapolation. Let $f_{i_x j_y}$ denote the gridpoint values of any boundary or initial field in the inner computational area (with grid indices $1 \leq i_x \leq n_x$ and $1 \leq j_y \leq n_y$). Let the extended area have grid indices $1 \leq i_x \leq n_{xl}$ and $1 \leq j_y \leq n_{yl}$. Gridpoint values in the extension zone are first calculated for each row of gridpoints of the inner computational area (grid indices $n_x + 1 \leq i_x \leq n_{xl}$ and $1 \leq j_y \leq n_y$):

$$f_{i_x j_y} = g_{j_y 0} + g_{j_y 1} \cos(0.5 x_i) + g_{j_y 2} \sin(0.5 x_i) + g_{j_y 3} \sin(x_i) \tag{1}$$

where $x_i = \frac{2\pi(i_x - n_x)}{n_{xl} + 1 - n_x}$ and the coefficients $g_{j_y k}$ (k=0,1,2 and 3) are determined to make the extrapolation expression fit the original gridpoint-values exactly for i_x=1,2,n_x-1 and n_x. The extrapolation to obtain gridpoint values in the "upper" part of the extension zone (grid indices $1 \leq i_x \leq n_{xl}$ and $n_y + 1 \leq j_y \leq n_{yl}$) is carried out similarly:

$$f_{i_x j_y} = g_{i_x 0} + g_{i_x 1} \cos(0.5 y_j) + g_{i_x 2} \sin(0.5 y_j) + g_{i_x 3} \sin(y_j) \tag{2}$$

where $y_j = \frac{2\pi(j_y - n_y)}{n_{yl} + 1 - n_y}$ and the coefficients $g_{i_x k}$ (k=0,1,2 and 3) are determined to make the extrapolation expression fit the original and previously fitted gridpoint values exactly for j_y=1, 2, n_y-1 and n_y.

This procedure to obtain bi-periodic variations include a certain degree of arbitrariness since there is no single unique way of doing the extrapolation. Results of experiments to test the sensitivity to details in the extrapolation, have indicated that this arbitrariness only has a very minor impact on forecast results.

2.2 Fourier transforms and spectral truncation

Once the area extension algorithm has been applied, 2-dimensional forward discrete Fourier truncated transforms can be used to obtain the time-dependent spectral coefficient fields to be integrated forward in time by the spectral forecast model:

$$\hat{f}_{kl} = \frac{1}{\sqrt{n_{xl}\, n_{yl}}} \sum_{i_x=1}^{n_{xl}} \sum_{j_y=1}^{n_{yl}} f_{i_x j_y}\, e^{-2\pi i(\frac{k i_x}{n_{xl}} + \frac{l j_y}{n_{yl}})} \tag{3}$$

Similarly, needed gridpoint values of the transform grid to be used during the forecast model integrations, may be obtained by 2-dimensional inverse truncated Fourier transforms:

$$f_{i_x j_y} = \frac{1}{\sqrt{n_{xl}\, n_{yl}}} \sum_{k=-k_{max}}^{k_{max}} \sum_{l=-l_{max}}^{l_{max}} \hat{f}_{kl}\, e^{+2\pi i(\frac{k i_x}{n_{xl}} + \frac{l j_y}{n_{yl}})} \tag{4}$$

To obtain an aliasing-free representation of quadratic terms in the transform grid, this grid should include at least 3*M+1 grid-points in one direction, if M is the number of complex Fourier-components of the spectral model in that particular direction. However, in the present formulation of the spectral HIRLAM model, these quadratic terms also include map-factors. These need to be considered in the determination of the transform grid. Thus if the inverse map factors are truncated by N waves, the transform grid (covering the extended area) should contain at least 3*M+N+1 gridpoints for a completely aliasing-free representation of these quadratic terms. In addition, an "elliptic" spectral truncation defined by

$$\left(\frac{k}{k_{max}}\right)^2 + \left(\frac{l}{l_{max}}\right)^2 \leq 1 \tag{5}$$

is utilized. k and l are any wave numbers in the x- and the y-directions respectively; k_{max} and l_{max} are the maximum wavenumbers in the x- and y-directions respectively. This gives an isotropic and homogeneous resolution over the extended integration area, provided these maximum wave numbers correspond to waves of equal lengths in the two horizontal directions.

2.3   Continuous equations

The spectral HIRLAM model is based on the moist primitive equations, formulated for the pressure-based and terrain-following hybrid vertical coordinate system introduced by Simmons and Burridge (1981). This vertical coordinate $\eta(p, p_s)$ depends on pressure p and surface pressure p_s and it has the properties $\eta(0, p_s) = 0$ and $\eta(p_s, p_s) = 1$. The model equations have been written for general horizontal coordinate systems. Map-factors (h_x, h_y) have been introduced in both horizontal directions. For a distance $(\delta X, \delta Y)$ it yields that $\delta X = h_x \delta x$ and $\delta Y = h_y \delta y$. The horizontal momentum equations are

$$\frac{du}{dt} = -\frac{R_d T_v}{h_x}\frac{\partial \ln p}{\partial x} - \frac{1}{h_x}\frac{\partial \Phi}{\partial x} + fv + \frac{v}{h_x h_y}\left(v\frac{\partial h_y}{\partial x} - u\frac{\partial h_x}{\partial y}\right) + P_u + K_u \tag{6}$$

and

$$\frac{dv}{dt} = -\frac{R_d T_v}{h_y}\frac{\partial \ln p}{\partial y} - \frac{1}{h_y}\frac{\partial \Phi}{\partial y} - fu - \frac{u}{h_x h_y}\left(v\frac{\partial h_y}{\partial x} - u\frac{\partial h_x}{\partial y}\right) + P_v + K_v \tag{7}$$

where

$$\dot{\eta} = \frac{d\eta}{dt} \tag{8}$$

and

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \frac{u}{h_x}\frac{\partial}{\partial x} + \frac{v}{h_y}\frac{\partial}{\partial y} + \dot{\eta}\frac{\partial}{\partial \eta} \tag{9}$$

The thermodynamic equation is given by

$$\frac{dT}{dt} = \frac{\kappa T_v \omega}{(1 + (\delta - 1)q)p} + P_T + K_T \tag{10}$$

where $\delta = \frac{c_{pv}}{c_{pd}}$. The moisture equation is given by

$$\frac{dq}{dt} = P_q + K_q \tag{11}$$

In all the equations given above, P_x is the tendency of variable x due to physical parameterization and K_x is the tendency of variable x due horizontal diffusion. The hydrostatic equation is given by

$$\frac{\partial \Phi}{\partial \eta} = -\frac{R_d T_v}{p} \frac{\partial p}{\partial \eta} \tag{12}$$

The continuity equation is given by

$$\frac{\partial}{\partial \eta} \frac{\partial p}{\partial t} + \nabla \cdot \left(\vec{v}_h \frac{\partial p}{\partial \eta} \right) + \frac{\partial}{\partial \eta} \left(\dot{\eta} \frac{\partial p}{\partial \eta} \right) = 0 \tag{13}$$

where the divergence operator is defined by

$$\nabla \cdot \vec{v}_h = \frac{1}{h_x h_y} \left\{ \frac{\partial}{\partial x}(h_y u) + \frac{\partial}{\partial y}(h_x v) \right\} \tag{14}$$

Integrating the continuity equation, with boundary conditions $\dot{\eta} = 0$ at $\eta = 0$ and $\eta = 1$, results in the surface pressure tendency equation

$$\frac{\partial p_s}{\partial t} = -\int_0^1 \nabla \cdot \left(\vec{v}_h \frac{\partial p}{\partial \eta} \right) d\eta \tag{15}$$

and the equation for $\dot{\eta}$ will be given by

$$\dot{\eta} \frac{\partial p}{\partial \eta} = \left(1 - \frac{\partial p}{\partial p_s} \right) \frac{\partial p_s}{\partial t} + \int_\eta^1 \nabla \cdot \left(\vec{v}_h \frac{\partial p}{\partial \eta} \right) d\eta \tag{16}$$

2.4 Vertical discretization

For the discretization in the vertical derivatives, the vertical staggering and the second order vertical finite difference scheme of Simmons and Burridge (1981) is utilized. Pressure is defined at model "half" levels by

$$p_{k+1/2} = A_{k+1/2} + B_{k+1/2} p_s(x, y) \; ; \; k = 0, nlev \tag{17}$$

Where $nlev$ is the number of model "full" levels. Model variables T, q, u and v are defined at these model full levels, while vertical velocity $\dot{\eta}$ is defined at the model half levels. Pressures at model full levels are defined as averages of pressures at the model half levels

$$p_k = \frac{p_{k-1/2} + p_{k+1/2}}{2} \; ; \; k = 1, nlev \tag{18}$$

The surface pressure tendency equation is approximated by introduction of vertical finite differences and a vertical numerical integration

$$\frac{\partial p_s}{\partial t} = -\sum_{k=1}^{nlev} \nabla \cdot ((\vec{v}_h)_k \Delta p_k) = -\sum_{k=1}^{nlev} ((\vec{v}_h)_k \cdot \nabla(\Delta B_k p_s) + \Delta p_k \nabla \cdot (\vec{v}_h)_k) \tag{19}$$

where the vertical finite difference operator is defined by

$$\Delta x_k = x_{k+1/2} - x_{k-1/2} \; ; \; k = 1, nlev \tag{20}$$

Introducing

$$D_k = \nabla \cdot (\vec{v}_h)_k \tag{21}$$

and

$$S_k = ((\vec{v}_h)_k \cdot \nabla \ln p_s)\Delta B_k + \frac{\Delta p_k D_k}{p_s} \tag{22}$$

we will have

$$\frac{\partial \ln p_s}{\partial t} = -\sum_{j=1}^{nlev} S_k \tag{23}$$

It may be noticed that this discrete form of the pressure tendency equation follows the Eulerian version of the ECMWF spectral model. It deviates from the gridpoint HIRLAM formulation, designed to preserve total mass with second order finite difference approximations in the horizontal. From the continuity equation in finite difference form we have

$$\left(\dot{\eta}\frac{\partial p}{\partial \eta}\right)_{k+1/2} = p_s \left\{(1 - B_{k+1/2})\frac{\partial \ln p_s}{\partial t} + \sum_{j=k+1}^{nlev} S_j\right\} \tag{24}$$

and the vertical advection term is given by

$$\left(\dot{\eta}\frac{\partial x}{\partial \eta}\right)_k = \frac{1}{2\Delta p_k}\left\{\left(\dot{\eta}\frac{\partial p}{\partial \eta}\right)_{k+1/2}(x_{k+1} - x_k) + \left(\dot{\eta}\frac{\partial p}{\partial \eta}\right)_{k-1/2}(x_k - x_{k-1})\right\} \tag{25}$$

Geopotential gradients at model half levels are obtained by integrating the hydrostatic equation upwards

$$\nabla \Phi_{k+1/2} = \nabla \Phi_s + \sum_{j=k+1}^{nlev} R_d \nabla(T_v)_j \Delta \ln p_j \tag{26}$$

To get model full level values of geopotential gradients, the following approximation is utilized

$$\nabla \Phi_k = \nabla \Phi_{k+1/2} + \nabla(\alpha_k R_d(T_v)_k) \tag{27}$$

where

$$\alpha_k = \begin{cases} \ln 2 & k = 1 \\ 1 - \frac{p_{k-1/2}}{\Delta p_k}\Delta \ln p_k & k > 1 \end{cases} \tag{28}$$

Following Simmons and Burridge (1981), the second part of the pressure gradient term is approximated by

$$(R_d T_v \nabla \ln p)_k = \frac{R_d(T_v)_k}{\Delta p_k} \left(\Delta \ln p_k \nabla p_{k-1/2} + \alpha_k \nabla \Delta p_k \right) \tag{29}$$

with α_k given by the general expression for all k. Utilizing the mass continuity equation, the energy conversion term of the thermodynamic equation may be written

$$\left(\frac{\kappa T_v \omega}{(1 + (\delta - 1)q)p} \right)_k = \frac{\kappa(T_v)_k}{1 + (\delta - 1)q_k} \left(\frac{1}{p} \frac{dp}{dt} \right)_k =$$

$$\frac{\kappa(T_v)_k}{1 + (\delta - 1)q_k} \left\{ -\frac{1}{p} \int_0^\eta \nabla \cdot \left(\vec{v}_h \frac{\partial p}{\partial \eta} \right) d\eta + \vec{v}_h \cdot \nabla \ln p \right\}_k \tag{30}$$

For energy consistency reasons, this term is approximated by finite differences in the vertical as follows

$$\left(\frac{\kappa T_v \omega}{(1 + (\delta - 1)q)p} \right)_k =$$

$$\frac{\kappa(T_v)_k}{1 + (\delta - 1)q_k} \left\{ -\frac{p_s}{\Delta p_k} \left[\Delta \ln p_k \sum_{j=1}^{k-1} S_j + \alpha_k S_k \right] \right.$$

$$\left. + \frac{p_s}{\Delta p_k} \left[\Delta B_k + \frac{C_k \Delta \ln p_k}{\Delta p_k} \right] (\vec{v}_h)_k \cdot \nabla \ln p_s \right\} \tag{31}$$

where

$$C_k = A_{k+1/2} B_{k-1/2} - A_{k-1/2} B_{k+1/2} \tag{32}$$

2.5 Linearized model equations

We will linearize the model around a dry basic state at rest, with a constant surface pressure and with a constant temperature at each model level. In addition, the basic state Coriolis parameter and the inverse map factors are assumed constant for the linearization. The following substitutions are thus introduced

$$T_k = \overline{T}_k + (T_k)' \quad \ln p_s = \overline{\ln p_s} + (\ln p_s)' \quad f = \overline{f} + f'$$

$$h_x^{-1} = \overline{h_x^{-1}} + (h_x^{-1})' \quad h_y^{-1} = \overline{h_y^{-1}} + (h_y^{-1})' \tag{33}$$

By linearization of the pressure gradient and Coriolis terms of the horizontal momentum equations and by keeping non-linear terms and deviations form the linearized terms on the right hand side, we will have for $1 \leq k \leq nlev$

$$\frac{\partial u_k}{\partial t} - \overline{f} \, v_k + \overline{h_x^{-1}} \frac{\partial P_k}{\partial x} = N_{u_k} \tag{34}$$

$$\frac{\partial v_k}{\partial t} + \overline{f} \, u_k + \overline{h_y^{-1}} \frac{\partial P_k}{\partial y} = N_{v_k} \tag{35}$$

where N_u and N_v denotes remaining terms of the horizontal momentum equations including deviations of non-linear terms from linearized terms. P_k denotes an efficient geopotential of the linearized pressure gradient term and is given by

$$P_k = \sum_{j=1}^{nlev} G_{kj}(T_j)' + R_d \overline{T_k^*}(\ln p_s)' \tag{36}$$

where the coefficients for summation of the temperature and pressure deviations from the basic state have been determined in accordance with the vertical finite difference schemes, i.e.

$$G_{kj} = \begin{cases} 0 & j < k \\ R_d \overline{\alpha_j} & j = k \\ R_d (\Delta \ln \bar{\sigma})_j & j > k \end{cases} \tag{37}$$

where

$$\overline{\alpha_j} = \begin{cases} \ln 2 & j = 1 \\ 1 - \bar{\sigma}_{j-1/2} \dfrac{(\Delta \ln \bar{\sigma})_j}{(\Delta \bar{\sigma})_j} & j > 1 \end{cases} \tag{38}$$

$$\bar{\sigma}_{j+1/2} = \frac{A_{j+1/2}}{\overline{p_s}} + B_{j+1/2} \tag{39}$$

$$\overline{T_k^*} = \overline{T_k} \frac{B_{k+1/2}}{\bar{\sigma}_{k+1/2}} + \sum_{j=k+1}^{nlev} \overline{T_j} \left[\frac{B_{j+1/2}}{\bar{\sigma}_{j+1/2}} - \frac{B_{j-1/2}}{\bar{\sigma}_{j-1/2}} \right] \tag{40}$$

Similarly, by linearization of the divergence term of the pressure tendency equation we will have

$$\frac{\partial (\ln p_s)'}{\partial t} = \sum_{k=1}^{nlev} \Delta \bar{\sigma}_k d_k = N_{p_s} \tag{41}$$

where d_k is the linearized divergence

$$d_k = \overline{h_x^{-1}} \frac{\partial u}{\partial x} + \overline{h_y^{-1}} \frac{\partial v}{\partial y} \tag{42}$$

In the thermodynamic equation we need to linearize the term describing advection of mean state temperature and the energy conversion term. These linearizations will give

$$\frac{\partial (T_k)'}{\partial t} + \sum_{j=1}^{nlev} D_{kj} d_j = N_{T_k} \tag{43}$$

where

$$D_{kj} = \begin{cases} \left[\delta_{k1} \kappa \overline{T_k} \dfrac{(\Delta \ln \bar{\sigma})_k}{\Delta(\bar{\sigma})_k} - \dfrac{1}{2} \dfrac{\delta_{k\ nlev} \Delta \bar{T}_{k+1/2} + \delta_{k\ 1} \Delta \bar{T}_{k-1/2}}{(\Delta \bar{\sigma})_k} + b_k \right] (\Delta \bar{\sigma})_j & j < k \\ \kappa \overline{T_k} \bar{\alpha}_k - \dfrac{1}{2} \delta_{k\ nlev} \Delta \bar{T}_{k+1/2} + b_k (\Delta \bar{\sigma})_k & j = k \\ b_k (\Delta \bar{\sigma})_j & j > k \end{cases} \tag{44}$$

and

$$b_k = \frac{\delta_{k\ nlev} B_{k+1/2} \Delta \bar{T}_{k+1/2} + \delta_{k\ 1} B_{k-1/2} \Delta \bar{T}_{k-1/2}}{2 (\Delta \bar{\sigma})_k} \tag{45}$$

and

$$\delta_{kl} = \begin{cases} 1 & k = l \\ 0 & k \neq l \end{cases} \tag{46}$$

The pressure tendency equation and the thermodynamic equation may be combined to form an equation for the linearized geopotential P_k

$$\frac{\partial P_k}{\partial t} + \sum_{j=1}^{nlev} C_{kj} d_j = N_{P_k} \tag{47}$$

where

$$C_{kj} = \sum_{l=1}^{nlev} G_{kl} D_{lj} + R_d \bar{T}_k^* (\Delta \bar{\sigma})_j \tag{48}$$

2.6 Implicit horizontal diffusion

Fourth order implicit horizontal diffusion is applied, which means that the following equation is solved implicitly

$$\frac{\partial A}{\partial t} + K \nabla^4 A = 0 \tag{49}$$

This equation is solved in spectral space and to simplify the solution the map-factors are assumed to be constant. For a particular wave-number (m, n), this equation becomes

$$\frac{\partial \hat{A}_{mn}}{\partial t} + K_{mn} \hat{A}_{mn} = 0 \tag{50}$$

where

$$K_{mn} = K \left\{ \left(\frac{2\pi m}{n_{xl} \bar{\bar{h}}_x} \right)^2 + \left(\frac{2\pi n}{n_{yl} \bar{h}_y} \right)^2 \right\}^2 \tag{51}$$

After discretization in time, we will have the following expression for the modifications $\delta \hat{A}_{mn}(t + \Delta t)$ to be added to the field $\hat{A}_{mn}(t + \Delta t)$ at time $t + \Delta t$

$$\delta \hat{A}_{mn}(t + \Delta t) = -\frac{\Delta t K_{mn}}{1 + \Delta t K_{mn}} \hat{A}_{mn}(t + \Delta t) \tag{52}$$

The implicit horizontal diffusion of the wind field includes the possibility to have different horizontal diffusion coefficients for the divergent and the rotational parts of the wind field. The solution of this diffusion problem is obtained without explicit splitting of the full wind field to the divergent and rotational parts.

2.7 Eulerian semi-implicit time integration scheme and lateral boundary conditions

The spectral HIRLAM model uses a semi-implicit time stepping scheme to allow for longer time-steps than those determined by fastest gravity waves. In the semi-implicit scheme for the Eulerian time-stepping scheme, the pressure gradient terms of the horizontal momentum equations and terms containing divergences in the pressure tendency equation and in the thermodynamic equation are linearized and treated implicitly. Utilizing the spectral form of the linearized equations, derived in the previous section, and introducing time averaging and time differences, these linearized equations may be written

$$\delta_t \hat{u}_{mnk} + im' \overline{h_x^{-1}} \, \overline{\hat{P}_{mnk}}^{2t} = (\frac{\partial \hat{u}_{mnk}}{\partial t})_{expl} + im' \overline{h_x^{-1}} \, \hat{P}_{mnk}(t) \tag{53}$$

$$\delta_t \hat{v}_{mnk} + in' \overline{h_y^{-1}} \, \overline{\hat{P}_{mnk}}^{2t} = (\frac{\partial \hat{v}_{mnk}}{\partial t})_{expl} + in' \overline{h_y^{-1}} \, \hat{P}_{mnk}(t) \tag{54}$$

$$\delta_t \hat{P}_{mnk} + \sum_{j=1}^{nlev} C_{kj} \overline{\hat{d}_{mnj}}^{2t} = (\frac{\partial \hat{P}_{mnk}}{\partial t})_{expl} + \sum_{j=1}^{nlev} C_{kj} \hat{d}_{mnj}(t) \tag{55}$$

$$\delta_t (\ln \hat{p}_s)_{mn} + \sum_{j=1}^{nlev} \Delta \overline{\sigma}_j \overline{\hat{d}_{mnj}}^{2t} = (\frac{\partial (\ln \hat{p}_s)_{mn}}{\partial t})_{expl} + \sum_{j=1}^{nlev} \Delta \overline{\sigma}_j \hat{d}_{mnj}(t) \tag{56}$$

$$\delta_t \hat{T}_{mnk} + \sum_{j=1}^{nlev} D_{kj} \overline{\hat{d}_{mnj}}^{2t} = (\frac{\partial \hat{T}_{mnk}}{\partial t})_{expl} + \sum_{j=1}^{nlev} D_{kj} \hat{d}_{mnj}(t) \tag{57}$$

where m and n are horizontal wave-numbers in the x- and y-directions respectively and k the number of the vertical level. Explicitly computed tendencies are denoted by $()_{expl}$. Furthermore $m' = \frac{2\pi m}{n_{xl}}$, $n' = \frac{2\pi n}{n_{yl}}$ and

$$\delta_t A = \frac{A(t + \Delta t) - A(t - \Delta t)}{2\Delta t} \tag{58}$$

and

$$\overline{(A)}^{2t} = \frac{A(t - \Delta t) + A(t + \Delta t)}{2} \tag{59}$$

The $\overline{(A)}^{2t}$-terms have been formed by averaging the linearized terms over time-steps $t + \Delta t$ and $t - \Delta t$ and by subtraction of the same terms at time t from the explicitly computed tendencies. Taking the divergence of linearized momentum equations and eliminating the linearized geopotential by the the third equation will give

$$\hat{d}_{mnk}(t + \Delta t) + (\Delta t \, \alpha_{mn})^2 \sum_{j=1}^{nlev} C_{kj} \hat{d}_{mnj}(t + \Delta t) = \hat{Q}_{mnk} \tag{60}$$

where

$$\alpha_{mn} = \sqrt{\left(m' \overline{h_x^{-1}}\right)^2 + \left(n' \overline{h_y^{-1}}\right)^2} \tag{61}$$

and

$$\hat{Q}_{mnk} = \hat{d}_{mnk}(t - \Delta t) + 2\Delta t (\frac{\partial \hat{d}_{mnk}}{\partial t})_{expl} - \Delta t \alpha_{mn}^2 \left(2\hat{P}_{mnk}(t) - \hat{P}_{mnk}(t - \Delta t)\right) +$$
$$\Delta t \alpha_{mn}^2 \left(\hat{P}_{mnk}(t - \Delta t) + 2\Delta t (\frac{\partial \hat{P}_{mnk}}{\partial t})_{expl} + \Delta t \sum_{j=1}^{nlev} C_{kj} (2\hat{d}_{mnj}(t) - \hat{d}_{mnj}(t - \Delta t))\right) \tag{62}$$

The divergence equation can be solved by the aid of the eigenvectors of the vertical structure matrix $C = \{C_{kj}\}$. Let us write the divergence equation in matrix form for each horizontal wave number couple m and n:

$$\vec{d}_{mn}(t + \Delta t) + (\Delta t\,\alpha_{mn})^2\,C\,\vec{d}_{mn}(t + \Delta t) = \vec{Q}_{mn} \tag{63}$$

Let the matrix E contain the eigen-vectors of the vertical structure matrix C, let Φ denote the vector of the corresponding eigen-values and we may transform the divergence equation in the following way:

$$E^{-1}\,\vec{d}_{mn}(t + \Delta t) + (\Delta t\,\alpha_{mn})^2\,E^{-1}CEE^{-1}\,\vec{d}_{mn}(t + \Delta t) = E^{-1}\vec{Q}_{mn} \tag{64}$$

Since $E^{-1}CE = \Phi$ is diagonal, the solution for the vector of the divergence transformed to normal mode space by the transformation matrix E^{-1} can be immediately obtained. Final model level divergences at the new time-step can then be obtained by transformation with E back to model levels.

The time history of the forecast model variables involved in the dynamical equations is given in spectral space. In addition, some forecast model variables like surface temperature, soil wetness and cloud water are defined in gridpoints of the inner integration area only. This means that model tendencies can be computed for this inner area only. To make it possible to integrate the spectral forecast model forward in time, the semi-implicit model solution is forced towards the lateral boundary conditions extrapolated to the extension zone. Following Radnoty (1995), this is achieved through a relaxation of the right hand sides of the semi-implicit equations towards the left hand sides determined from the lateral boundary conditions in the boundary relaxation zone. For example, for the u momentum equation in gridpoint space we will schematically have:

$$u(i_x, j_y, t + \Delta t) + \frac{\Delta t}{\overline{h}_x}\frac{\partial P(i_x, j_y, t + \Delta t)}{\partial x} = \alpha(i_x, j_y)u^R(i_x, j_y, t + \Delta t) +$$

$$(1 - \alpha(i_x, j_y))\left(u^B(i_x, j_y, t + \Delta t) + \frac{\Delta t}{\overline{h}_x}\frac{\partial P^B(i_x, j_y, t + \Delta t)}{\partial x}\right) \tag{65}$$

where u and P are included in the semi-implicit model solution to be determined and u^R is the right hand side for the semi-implicit u momentum equation as calculated in the inner integration area. u^B and P^B are the lateral boundary fields, also extrapolated to the extension zone.

The boundary relaxation of the right hand sides of the semi-implicit model equations is carried out by the Davies (1983) boundary relaxation scheme. The weighting factor $\alpha(i_x j_y)$ depends on the distance $r(i_x j_y)$ (in gridpoint-units) from the grid-point (i_x, j_y) to the boundary of the inner integration area

$$\alpha(i_x j_y) = 0.5\left(1 - \cos(\pi\frac{r(i_x j_y)}{n_b})\right) \tag{66}$$

where $r(i_x j_y)$ is the distance in grid-units between grid-point (i_x, j_y) and the lateral boundary and n_b is the width of the boundary relaxation zone.

It should be noted that, although the area extension algorithm (bi-periodization) involves a certain degree of arbitrariness, the algorithm is only applied once for each initial and boundary model state only. During the forecast model integrations, the right hand sides of the semi-implicit forecast model equations are obtained for the fully extended area by relaxation towards the pre-calculated boundary fields.

3 NON-LINEAR NORMAL MODE INITIALIZATION

3.1 Derivation of the normal modes

The normal modes used implicitly for the initialization of the spectral HIRLAM are eigen-functions of the linearized equations for momentum and linearized geopotential derived above. For horizontal spectral components m, n and for vertical level k these equations in spectral space may be written

$$\frac{d\hat{u}_{knm}}{dt} - \bar{f}\hat{v}_{kmn} + \frac{im'}{\bar{h}_x}\hat{P}_{kmn} = 0 \tag{67}$$

$$\frac{d\hat{v}_{knm}}{dt} + \bar{f}\hat{u}_{kmn} + \frac{in'}{\bar{h}_y}\hat{P}_{kmn} = 0 \tag{68}$$

$$\frac{d\hat{P}_{knm}}{dt} + \sum_{l=1}^{nlev} C_{kl}\left(\frac{im'}{\bar{h}_x}\hat{u}_{kmn} + \frac{in'}{\bar{h}_x}\hat{v}_{kmn}\right) = 0 \tag{69}$$

This system of linear equations is coupled in the vertical through the vertical structure matrix $C = \{C_{kj}\}$. We may de-couple the system in the vertical by projection of $\hat{u}_{kmn}$, $\hat{v}_{kmn}$ and $\hat{P}_{kmn}$ on the eigen-vectors of C. We will then obtain a system of vertically de-coupled equations for the model state variables in vertical normal mode space. For vertical mode l we will have:

$$\frac{d\tilde{u}_{lnm}}{dt} - \bar{f}\tilde{v}_{lmn} + \frac{im'}{\bar{h}_x}\tilde{P}_{lmn} = 0 \tag{70}$$

$$\frac{d\tilde{v}_{lnm}}{dt} + \bar{f}\tilde{u}_{lmn} + \frac{in'}{\bar{h}_y}\tilde{P}_{lmn} = 0 \tag{71}$$

$$\frac{d\tilde{P}_{lnm}}{dt} + \bar{\Phi}_l\left(\frac{im'}{\bar{h}_x}\hat{u}_{lmn} + \frac{in'}{\bar{h}_x}\hat{v}_{lmn}\right) = 0 \tag{72}$$

Where $\bar{\Phi}_l$ is the l'th eigenvalue (vertical equivalent depth) of the vertical structure matrix C. In order to facilitate the derivation of the horizontal eigen-values, the horizontal eigen-vectors and the non-linear normal mode initialization scheme, we will introduce the vorticity and divergence equations. If we express this system of linear equations in matrix form for the stream-function and the velocity potential, we will have

$$\frac{d}{dt}\begin{vmatrix}\tilde{\psi}_{lmn}\\\tilde{\chi}_{lmn}\\\tilde{P}_{lmn}\end{vmatrix} = \begin{vmatrix}0 & -f & 0\\+f & 0 & -1\\0 & \bar{\Phi}_l\alpha_{mn}^2 & 0\end{vmatrix}\begin{vmatrix}\tilde{\psi}_{lmn}\\\tilde{\chi}_{lmn}\\\tilde{P}_{lmn}\end{vmatrix} \tag{73}$$

where $\alpha_{mn}^2 = \frac{m'^2}{\bar{h}_x^2} + \frac{n'^2}{\bar{h}_y^2}$

For a particular vertical mode l and a particular horizontal wave-number (m,n), the eigen-values λ_k and eigen-vectors $\vec{\phi}_k$ (k=1,2,3) are given by the eigen-value problem

$$\begin{vmatrix}0 & -f & 0\\+f & 0 & -1\\0 & \bar{\Phi}_l\alpha_{mn}^2 & 0\end{vmatrix}\vec{\phi}_k = \lambda_k\vec{\phi}_k \tag{74}$$

The eigen-values will be obtained by requiring the determinant of the eigen-value problem to be zero:

$$\begin{cases}\lambda_1 &= 0\\\lambda_2 &= i\sigma_{lmn}\\\lambda_3 &= -i\sigma_{lmn}\end{cases} \tag{75}$$

where $\sigma_{lmn} = \sqrt{\bar{\Phi}_l\alpha_{mn}^2 + \bar{f}^2}$. If we utilize the following orthonormalization definition

$$\alpha_{mn}^2\bar{\Phi}_l(\phi_{k1}^*\phi_{k'1} + \phi_{k2}^*\phi_{k'2}) + \phi_{k3}^*\phi_{k'3} = \delta_{kk'} \tag{76}$$

the normalized eigen-vectors are given by

345

$$\vec{\phi}_1 = \frac{1}{\sigma_{lmn}} \begin{vmatrix} 1 \\ 0 \\ \bar{f} \end{vmatrix} \tag{77}$$

$$\vec{\phi}_2 = \frac{1}{\sqrt{2\bar{\Phi}_l \alpha_{mn}^2}\sigma_{lmn}} \begin{vmatrix} -f \\ i\sigma_{lmn} \\ \bar{\Phi}_l \alpha_{mn}^2 \end{vmatrix} \tag{78}$$

$$\vec{\phi}_3 = \frac{1}{\sqrt{2\bar{\Phi}_l \alpha_{mn}^2}\sigma_{lmn}} \begin{vmatrix} -f \\ -i\sigma_{lmn} \\ \bar{\Phi}_l \alpha_{mn}^2 \end{vmatrix} = \vec{\phi}_2^* \tag{79}$$

Note that the first eigen-mode corresponds to a stationary Rossby-mode in geostrophic balance with a constant Coriolis-parameter and that the second and third eigen-modes corresponds to west-ward and eastward travelling gravity modes.

3.2 Transformation to normal mode space

Introduce the following vector notation for a model solution of vertical mode l and horizontal wavenumbers m and n

$$\vec{\tilde{X}}_{lmn} = \begin{vmatrix} \tilde{\psi}_{lmn} \\ \tilde{\chi}_{lmn} \\ \tilde{P}_{lmn} \end{vmatrix} \tag{80}$$

We may also express this model solution as a linear combination of the eigen-vectors derived in the previous section

$$\vec{\tilde{X}}_{lmn} = \sum_{i=1}^{3} \tilde{Y}_{lmni} \vec{\phi}_i \tag{81}$$

or in matrix notation

$$\vec{\tilde{X}}_{lmn} = \vec{\bar{\Phi}}_{lmn} \vec{\tilde{Y}}_{lmn} \tag{82}$$

where

$$\vec{\bar{\Phi}}_{lmn} = \begin{vmatrix} \vec{\phi}_1 & \vec{\phi}_2 & \vec{\phi}_3 \end{vmatrix} \tag{83}$$

This means that

$$\vec{\tilde{Y}}_{lmn} = \vec{\bar{\Phi}}_{lmn}^{-1} \vec{\tilde{X}}_{lmn} \tag{84}$$

and it may be verified that

$$\vec{\bar{\Phi}}_{lmn}^{-1} = \vec{\bar{\Phi}}_{lmn}^{*} \begin{vmatrix} \bar{\Phi}_l \alpha_{mn}^2 & 0 & 0 \\ 0 & \bar{\Phi}_l \alpha_{mn}^2 & 0 \\ 0 & 0 & 0 \end{vmatrix} \tag{85}$$

3.3 Initialization scheme

Machenhauer(1977) introduced the non-linear normal mode initialization technique for shallow water equation models. For a description of the non-linear normal mode initialization, it is convenient to express the primitive equations in their normal mode form (for simplicity of notation we have dropped the indices denoting horizontal wavenumber and vertical mode number):

$$
\begin{cases}
\frac{dY_1}{dt} = \dot{Y}_1 = \qquad\quad + \; N1(Y_1, Y_2, Y_3) \\[2mm]
\frac{dY_2}{dt} = \dot{Y}_2 = \; i\sigma Y_2 \; + \; N2(Y_1, Y_2, Y_3) \\[2mm]
\frac{dY_3}{dt} = \dot{Y}_3 = -i\sigma Y_3 \; + \; N3(Y_1, Y_2, Y_3)
\end{cases}
\tag{86}
$$

The basic idea of the normal mode initialization is that the gravity modes (Y_2 and Y_3) are slowly varying and that there is a balance between the linear term and the non-linear terms in the equations given above for these gravity modes. Machenhauer proposed an iteration scheme for obtaining this balance by calculation of a new value (iteration $\tau + 1$) of the linear terms from the non-linear terms at the present iteration (τ):

$$
\begin{cases}
0 = \; i\sigma Y_2^{\tau+1} \; + \; N2(Y_1^\tau, Y_2^\tau, Y_3^\tau) \\[2mm]
0 = -i\sigma Y_3^{\tau+1} \; + \; N3(Y_1^\tau, Y_2^\tau, Y_3^\tau)
\end{cases}
\tag{87}
$$

By subtraction of the gravity mode tendencies calculated at iteration τ from this iterative equation we obtain the following equation for modification of the gravity modes at iteration $\tau + 1$:

$$
\begin{cases}
\Delta Y_2^{\tau+1} = Y_2^{\tau+1} - Y_2^\tau = \frac{\dot{Y}_2^\tau}{i\sigma} \\[2mm]
\Delta Y_3^{\tau+1} = Y_3^{\tau+1} - Y_3^\tau = \frac{\dot{Y}_3^\tau}{-i\sigma}
\end{cases}
\tag{88}
$$

We may write this equation in matrix form

$$
\Delta \vec{Y}^{\tau+1} = \vec{Y}^{\tau+1} - \vec{Y}^\tau =
\begin{vmatrix}
0 & 0 & 0 \\
0 & \frac{1}{i\sigma} & 0 \\
0 & 0 & -\frac{1}{i\sigma}
\end{vmatrix}
\dot{\vec{Y}}^\tau = \vec{N}\dot{\vec{Y}}^\tau
\tag{89}
$$

We may now transform this equation to an equation for modification of streamfunction, velocity potential and linearized geopotential. This equation may be written in component form

$$
\Delta \psi_{lmn}^{\tau+1} = -\frac{\bar{f}}{\sigma_{lmn}^2}\dot{\chi}_{lmn}^\tau
\tag{90}
$$

$$
\Delta \chi_{lmn}^{\tau+1} = \frac{1}{\sigma_{lmn}^2}(\bar{f}\dot{\psi}_{lmn}^\tau - \dot{P}_{lmn}^\tau)
\tag{91}
$$

$$
\Delta P_{lmn}^{\tau+1} = \frac{\bar{\Phi}_l \alpha_{mn}^2}{\sigma_{lmn}^2}\dot{\chi}_{lmn}
\tag{92}
$$

By further transforming from streamfunction and velocity potential to wind components we obtain our final equation for the iterative changes in vertical normal mode space:

$$
\Delta u_{lmn}^{\tau+1} = \frac{1}{\sigma_{lmn}^2}(\bar{f}\dot{v}_{lmn}^\tau - i\frac{m'}{h_x}\dot{P}_{lmn}^\tau)
\tag{93}
$$

$$\Delta v_{lmn}^{\tau+1} = \frac{1}{\sigma_{lmn}^2}(-\bar{f}\dot{u}_{lmn}^\tau - i\frac{n'}{\bar{h}_y}\dot{P}_{lmn}^\tau)$$ (94)

$$\Delta P_{lmn}^{\tau+1} = -\frac{i\bar{\Phi}_l}{\sigma_{lmn}^2}(\frac{m'}{\bar{h}_x}\dot{u}_{lmn}^\tau + \frac{n'}{\bar{h}_y}\dot{v}_{lmn}^\tau)$$ (95)

In general, 2-3 iterations of correcting the 3-4 gravest vertical modes turns out to be an efficient initialization procedure. Note that no transformation to horizontal normal modes is needed in the initialization scheme described above. This was made possible due to the linearization around a constant Coriolis parameter and constant map factors.

4   HIRLAM VARIATIONAL DATA ASSIMILATION.

The present HIRLAM data assimilation is a forward intermittent data assimilation based on Optimum Interpolation (OI). Weaknesses of this data assimilation are the lack of a full utilization of the time dimension during the data assimilation in addition to the requirement that relations between observed quantities and model state variables must be linear. Gustafsson et al. (1997) suggested that a new data assimilation system for HIRLAM should be based on 4D-Var, 4-dimensional variational data assimilation (Le Dimet and Talagrand, 1986, Lewis and Derber, 1985).

A necessary component of any 4-dimensional data assimilation based on variational techniques is the adjoint of the tangent linear forecast model. When the adjoint of the forecast model is available, it is possible to relate forecast errors, as determined by observations at any future time within the forecast range, back to possible errors in the initial conditions for the forecast. For the development of the first version of the adjoint HIRLAM, we decided to use the spectral formulation of the model. In general, it is easier to develop adjoints of spectral models since Fourier transforms are self-adjoint and since no efforts are needed to develop adjoints of complicated finite difference operators.

A first step in the development of 4D-Var is 3D-Var, 3-dimensional variational data assimilation. An important component of 3D-Var, as well as of 4D-Var, is the background error constraint. The background error constraint involves the inversion of the background error covariance matrix, the dimension of which is the squared number of model variables. One possibility to make this problem manageable is to consider the model variable assimilation increments in spectral space. For the development of the HIRLAM 3D-Var, the same area extension idea as is used for the spectral HIRLAM model has been applied (Gustafsson et al., 1999). It should be noted, however, that this does not necessarily mean that the spectral HIRLAM has to be applied in addition. Since the spectral representation and truncation is only concerned with the 3D-Var analysis increments, these increments may as well be added to an analysis background field in gridpoint space produced by the gridpoint HIRLAM.

5   THE ADJOINT OF THE SPECTRAL HIRLAM

5.1   The adjoint of the adiabatic part

We will introduce the concepts of tangent linear and adjoint models in a symbolic form only to be able to discuss our technique for derivation of the adjoint of the spectral HIRLAM, more comprehensive introductions to adjoint models may be found elsewhere (Thepaut and Courtier, 1991). We will use the standard notations suggested by Ide et al. (1997). Consider a non-linear model M for the forecast of a model state vector $\mathbf{x}(t)$ from time $t = t_0$ until time $t = t_1$:

$$\mathbf{x}(t_1) = M(\mathbf{x}(t_0))$$ (96)

Provided we know a non-linear solution $\mathbf{x}(t)$, we may introduce a tangent-linear model $\mathbf{M}(t_0, t_1)$ for small perturbations $\delta\mathbf{x}(t)$ added to this non-linear solution:

$$\delta \mathbf{x}(t_1) = \mathbf{M}(t_0, t_1)\delta \mathbf{x}(t_0) \tag{97}$$

Assume that we have observations available at time $t = t_1$. As a measure of the forecast error at time t_1 we introduce a cost function J. For the 4-dimensional variational data assimilation problem, we need to calculate gradients of this cost function with respect to the initial conditions. Introducing a scalar product $\langle \mathbf{x}, \mathbf{y} \rangle$, first order variations δJ of this cost function with respect to a perturbation $\delta \mathbf{x}(t)$ may be expressed by the gradient of the cost function with respect to initial conditions $\mathbf{x}(t_0)$ or with respect to the forecast $\mathbf{x}(t_1)$ at the observation time t_1:

$$\delta J = \left\langle \nabla_{\mathbf{x}(t_0)} J, \delta \mathbf{x}(t_0) \right\rangle \tag{98}$$

$$\delta J = \left\langle \nabla_{\mathbf{x}(t_1)} J, \delta \mathbf{x}(t_1) \right\rangle = \left\langle \nabla_{\mathbf{x}(t_1)} J, \mathbf{M}(t_0, t_1)\delta \mathbf{x}(t_0) \right\rangle \tag{99}$$

Introducing the adjoint $\mathbf{M}^*(t_0, t_1)$ of the tangent-linear model $\mathbf{M}(t_0, t_1)$, the second expression may be modified as follows:

$$\delta J = \left\langle \mathbf{M}^*(t_0, t_1)\nabla_{\mathbf{x}(t_1)} J, \delta \mathbf{x}(t_0) \right\rangle \tag{100}$$

Identification between the two expressions for the first order variation δJ gives us the equation for calculation of the gradient of the cost function with respect to the initial conditions:

$$\nabla_{\mathbf{x}(t_0)} J = \mathbf{M}^*(t_0, t_1)\nabla_{\mathbf{x}(t_1)} J \tag{101}$$

The time integration of the tangent-linear model normally is carried out by a number of timesteps forward in time. This can be considered as a sequence of matrix multiplications, one for each time-step. Due to the matrix transpose, this means that the adjoint model runs backward in time from time t_1 to time t_0.

We used a manual coding technique to develop the first version of the adjoint of the adiabatic part of the spectral HIRLAM including horizontal diffusion and the non-linear normal mode initialization. For each subroutine containing any non-linear expressions, we first coded the corresponding tangent-linear subroutine. Then the adjoints of each tangent-linear (and originally linear) subroutine were coded in a statement-by-statement fashion. By considering each statement of the tangent-linear and linear subroutines as a complex matrix operator, the corresponding adjoint statement(s) were derived by taking the complex conjugate and transpose of this matrix operator. All tangent-linear and adjoint subroutines were tested for the scalar product identity

$$< \mathbf{T}\delta \mathbf{x}, T\delta \mathbf{x} > = < \mathbf{T}^*\mathbf{T}\delta \mathbf{x}, \delta \mathbf{x} > \tag{102}$$

where $\mathbf{T}$ is the tangent-linear operator of a particular subroutine, $\mathbf{T}^*$ the corresponding adjoint operator and $\delta \mathbf{x}$ a perturbation that is applied in different directions. The same scalar product identity test was also applied successfully for complete tangent-linear and adjoint model timesteps as well as to a sequence of such timesteps over a 6 hour period. The tangent-linear model was simply verified by comparing the evolution of a small perturbation in the non-linear and the tangent-linear models.

¿From a technical point of view, it is straight-forward to extend the idea of the adjoint model to the treatment of the lateral boundaries (Errico et al., 1993). The boundary relaxation is a simple linear weighting, and thus the tangent linear equations for relaxation of lateral boundary perturbations are identical to forward full model boundary relaxation. Consider as an example, the boundary relaxation expression on the right hand side ($=RHS$) of (65):

$$RHS = \alpha(i_x, j_y)u^R(i_x, j_y, t + \Delta t)+$$

$$(1 - \alpha(i_x, j_y))\left(u^B(i_x, j_y, t + \Delta t) + \frac{\Delta t}{h_x}\frac{\partial P^B(i_x, j_y, t + \Delta t)}{\partial x}\right) \tag{103}$$

The corresponding adjoint boundary relaxation is easily derived:

$$\delta(u^R(i_x, j_y, t + \Delta t))^{AD} = \alpha(i_x, j_y)(RHS)^{AD} \tag{104}$$

$$\delta(u^B(i_x, j_y, t + \Delta t))^{AD} = (1 - \alpha(i_x, j_y))(RHS)^{AD} \tag{105}$$

$$\delta\left(\frac{\partial P^B(i_x, j_y, t + \Delta t)}{\partial x}\right)^{AD} = (1 - \alpha(i_x, j_y))\frac{\Delta t}{h_x}(RHS)^{AD} \tag{106}$$

It is not obvious how to proceed with the practical implementation of the adjoint of the lateral boundary relaxation scheme. In principle, the equation above will give adjoint boundary perturbations, representing a gradient of the error norm with respect to the lateral boundary condition, for each time-step of the model integration. A technical drawback of this solution is that it requires storing of the adjoint boundary values from all time-steps in the computer. An alternative is therefore to include also the adjoint of the linear time-interpolation of the boundary values.

5.2 Adjoints of HIRLAM physical parameterization schemes

Two strategies have been adopted for the development of tangent linear and adjoints of physical parameterization schemes for the HIRLAM 4D-Var. The first strategy is to use the automatic code generator TAMC of Ralf Giering for derivation of tangent linear and adjoints of various versions of the reference HIRLAM physics. TAMC is not yet perfect, which means that additional manual coding is needed for most occasions. The second strategy is to use some form of simplified physics for HIRLAM 4D-Var. A package of simplified physics is being developed by the ARPEGE- and ALADIN-groups (M. Janiskova, personal communication) and it is the intention to apply this package to HIRLAM 4D-Var.

Physical parameterization schemes with strong non-linear behavior have been a long lasting problems for HIRLAM. Recently a new version (4.2) of the HIRLAM forecast system was introduced. This new version include a new condensation, precipitation and convection scheme (STRACO), developed by Bent Hanssen Sass on the basis of the Sundquist scheme. Much effort has been spent on making the behavior of the new scheme more smooth and regular both in time and in space. In addition to the new condensation scheme, also a new non-local vertical diffusion scheme (Holtslag) has been introduced to HIRLAM 4.2. The effects of these modifications were clearly reflected in RMS-differences between un-perturbed and perturbed non-linear model runs with the two different versions of the HIRLAM model. The new physics, HIRLAM 4.2, turned out to have a much more regular behavior with regard to the sensitivity to very small initial perturbations. Tangent linear and adjoint of the STRACO, the Holtslag scheme and the other components of the HIRLAM 4.2 physics have been developed. Validation will continue with sensitivity experiments.

6 SENSITIVITY EXPERIMENTS

6.1 Design of the experiments

The adjoint model technique is a powerful tool to relate the origin of numerical forecast errors to errors in the initial data (Rabier et al., 1996, Gustafsson and Huang, 1996). The idea is to identify a

case with large forecast errors and then to trace the errors back to the initial state by calculating the gradient of the forecast error norm with respect to the initial conditions. For this purpose, the adjoint model is utilized to project the gradient of the forecast error norm from the forecast time back to the initial data time. Subtracting a fraction of the gradient of the forecast error norm from the initial condition and performing a new (sensitivity) forecast run with changed initial conditions would lead to an improved forecast.

For limited area models, poor lateral boundary conditions is another important error source distinct from the initial state errors. In addition to the direct negative effect of poor lateral boundary data during the forecast model integration, there may also be an accumulated negative effect during data assimilation forecast cycles, in particular if the lateral boundaries are situated upstream of data-sparse areas within the model integration area (Gustafsson, 1990). To distinguish between errors caused by erroneous boundary conditions and inaccuracies in the initial state, we will carry out sensitivity experiments using the adjoint of the HIRLAM forecast model. We will use the same experimentation technique as described by Gustafsson and Huang (1996), extended to take possible errors of the lateral boundary conditions also into account (Gustafsson et al., 1998). Thus, the adjoint backward integration provides us with an estimate of the gradient of a quadratic cost function of the forecast errors with respect to the initial conditions as well as to the lateral boundary conditions. Sensitivity forecast experiments are carried out by perturbing the initial as well as lateral boundary conditions with fractions of the respective cost function gradients. A fraction α_i of the cost function gradient with respect the initial conditions was subtracted from the original initial data in the sensitivity forecast, while the scaling factor α_b was similarly applied to the lateral boundary conditions.

Some preliminary sensitivity experiments with regard to errors in the lateral boundary conditions indicated that it is advantageous to apply individual lateral boundary perturbations for each timestep of the forward model integration. The spatial structures introduced by the lateral boundary relaxation appeared to become more accurate with perturbations available for every timestep, and this can be understood from the non-linear time variations of the lateral boundary perturbations over a 6 hour period. On the other hand, it is also known that the boundary relaxation scheme creates artificial divergent winds and related gravity wave oscillations in the boundary relaxation zone. In order to avoid enhanced gravity wave oscillations in the sensitivity forecast runs, a simple time filter was applied to the lateral boundary perturbations obtained by the adjoint model integration. Experimentation indicated that time averaging over ± 1 hour gave satisfactory results.

We carried out several experiments to study the sensitivity of forecast errors to the specification of the initial as well as the lateral boundary data. The experiments were run with the spectral version of HIRLAM, with a shortest resolved wave-length of approximately 180 km and with 16 vertical hybrid levels. The experiments were started from operational SMHI HIRLAM analysis fields, valid 12 hours before the verification time. For the lateral boundary conditions, ECMWF forecast fields based on initial data 24 hours before the verification time, were utilized. The frequency for updating of the lateral boundary conditions was once every 6 hours. Linear interpolation was used for the time-interpolation of the lateral boundary data. For calculation of the quadratic forecast error norm and its gradient at the verification time, operational SMHI HIRLAM analysis fields were used.

6.2   Results

A case occurring on 16 February 1995 was chosen. This case turned out to be particularly sensitive to lateral boundary conditions over eastern Canada, i.e. in an area with good data coverage. The 300 hPa height errors (verification analysis - forecast from the operational SMHI HIRLAM) starting from initial data valid at 0000 UTC on 16 February 1995, have been plotted in Figs. 2a–b for forecast lengths of 12 and 24 h. Note that the operational forecast grossly underestimates the forecast depth of a low propagating through and from the west side boundary over Newfoundland and later over the NW Atlantic. This indicates that the forecast error is likely to be related to poorly described structures in the lateral boundary conditions provided by the coarser resolution global model.
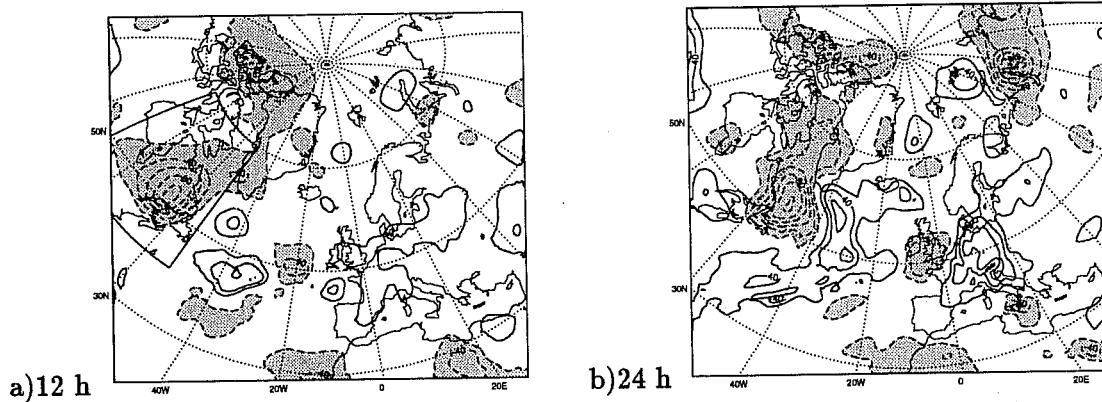
351

a)12 h   b)24 h

Figure 2: 300 hPa height error fields (verification analysis - forecasts from operational SMHI HIRLAM) for forecast lengths of 12 h and 24 h from 0000 UTC on 16 February 1995. Positive (negative) values are represented by solid (dashed) contours. Contours every 20 m. The shaded areas correspond to errors $\leq$ -20 m. The sector-formed area in a) is used for calculation of the area-concentrated 12 h forecast error norm.

Sensitivity experiments were carried out for the period between 0000 and 1200 UTC on 16 February 1995. Since the purpose of this investigation is to detect the origin of the forecast error over New-foundland, the area-concentrated forecast error norm at +12 h, indicated in Fig. 2a by heavy lines, was used. The first sensitivity experiment was carried out using only the initial data perturbation. The results are shown in Figs. 3a–b for 300 hPa height and wind differences between the sensitivity and reference +0 h and +12 h forecasts. Comparing Figs. 2a and 3b, it is obvious that the +12 h forecast error is only partly retrieved by the initial data sensitivity experiment. Considering that the initial perturbation, introduced by the aid of the adjoint model integration, at the initial time of the forecast is placed very close to the lateral boundaries, we may presume that errors in the lateral boundary conditions are also needed for a full explanation of the +12 h forecast errors.
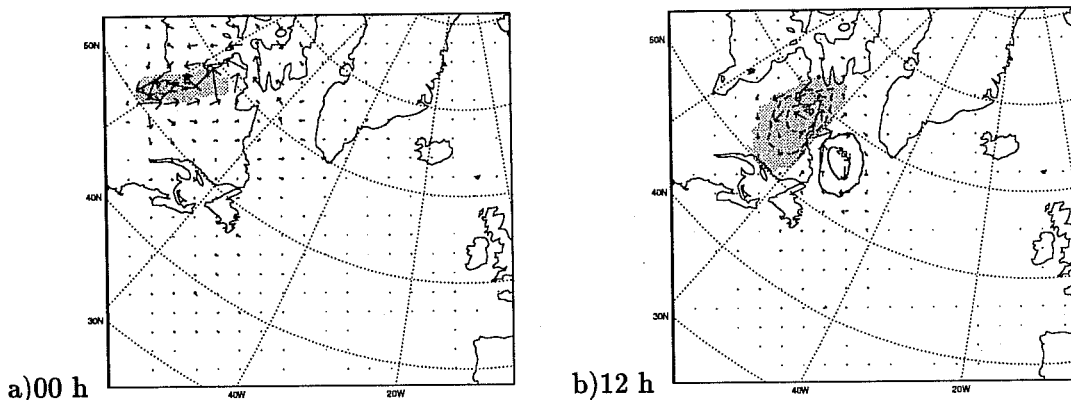

a)00 h   b)12 h

Figure 3: 300 hPa height and wind differences between the sensitivity and reference 0 h and 12 h forecasts from 0000 UTC on 16 February 1995 using initial perturbations only with $\alpha_i = 0.1$ and the area-concentrated 12 h forecast error norm. Positive (negative) values are represented by solid (dashed) contours. Contours every 10 m. The shaded areas correspond to retrieved forecast errors $\leq$ -5 m. The western boundary of the model integration area coincides with the left side of the maps.

The second sensitivity experiment was carried out using only the lateral boundary perturbations for the sensitivity forecast run. It turned out to be necessary to utilize a scaling coefficient $\alpha_b = 5.0$ for the lateral boundary perturbations in order to achieve the successful results, as illustrated in Fig. 4b for the +12 h forecasts. The needed magnitude of this scaling coefficient may partly be justified by the shorter time-period over which the lateral boundary perturbations in the average are allowed to grow, as compared to the initial data perturbations. More important, however, is that the adjoint

boundary relaxation expressions also efficiently act as a down-scaling of the gradient of the forecast error norm with respect to the lateral boundary conditions.
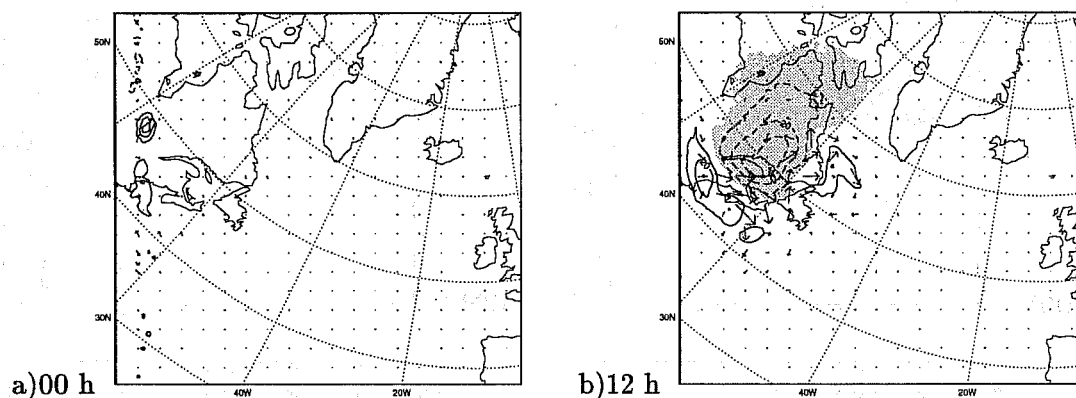

a)00 h   b)12 h

Figure 4: As in Fig. 3, but using only the lateral boundary perturbations for the sensitivity forecast run with $\alpha_b = 5.0$ and with a time-averaging period of ± 1 h: a) 0 h (boundary perturbation), b) 12 h. (a) Contour interval of 2 m, (b) contour interval of 10 m.

Since the initial data perturbation is zero for this particular forecast sensitivity run, we have included the lateral boundary perturbation, i.e. the scaled gradient of the forecast error norm with respect to the lateral boundary conditions, at time +0 h in Fig. 4a. We may notice that non-zero values of the lateral boundary perturbations are restricted to the interior of the boundary relaxation zone, as explained by a simple analysis to be presented below. We may also notice a significant boundary perturbation in the area of strong inflow upstream of the +12 h forecast error area of interest for this case study. It should be pointed out, however, that the whole time-series of boundary perturbations between +0 h and +12 h influence the sensitivity forecast run in this particular forecast model setup.

Comparing the sensitivity difference fields of the initial data and the lateral boundary data sensitivity experiments (Figs. 3b and 4b) with the real forecast error field (Fig. 2a) at +12 h, we may conclude that the sensitivity difference field in the lateral boundary data experiment has a significantly better position than the sensitivity difference field in the initial data experiment. At the same time the results from the initial data experiment include sensitivity difference field structures which are not recovered by the boundary data experiment. We may thus conclude that the forecast errors for this particular case are likely to be related to initial as well as lateral boundary data errors. This is confirmed by running a sensitivity experiment with regard to both initial and lateral boundary data errors (see results for +1 h, +6 h, +12 h and +24 h in Fig. 5). This experiment was carried out with a time interval of ± 1 h for the averaging of the lateral boundary perturbations, with a scaling coefficient of $\alpha_b = 5.0$ for the lateral boundary perturbations and with a scaling coefficient of $\alpha_i = 0.1$ for the initial data perturbations. The sensitivity difference field pattern in Fig. 5c at +12 h results in a sharper upper air trough over eastern Canada, which agrees very well with the forecast error pattern in Fig. 2a. We may also notice that the retrieved forecast error pattern at +1 h reflects the initial data perturbation (see Fig. 3a) as well as the lateral boundary data perturbation at the initial time (Fig. 4a).

The sensitivity forecast experiment utilizing initial data perturbations and lateral boundary data perturbations during the first 12 h of the model integration was continued up to +24 h, using the original lateral boundary conditions between +12 h and +24 h. The results of this experiment are included for +24 h in Figs. 5d. It is obvious that also the +24 h forecast is significantly improved in the critical area south of Greenland due to the initial and boundary data perturbations provided by the +12 h sensitivity experiment.

The successful application of the adjoint of the lateral boundary treatment for forecast sensitivity experiments, described here, can simply be interpreted to mean that we are using observed information inside the model integration area to improve poor lateral boundary conditions. This result may be
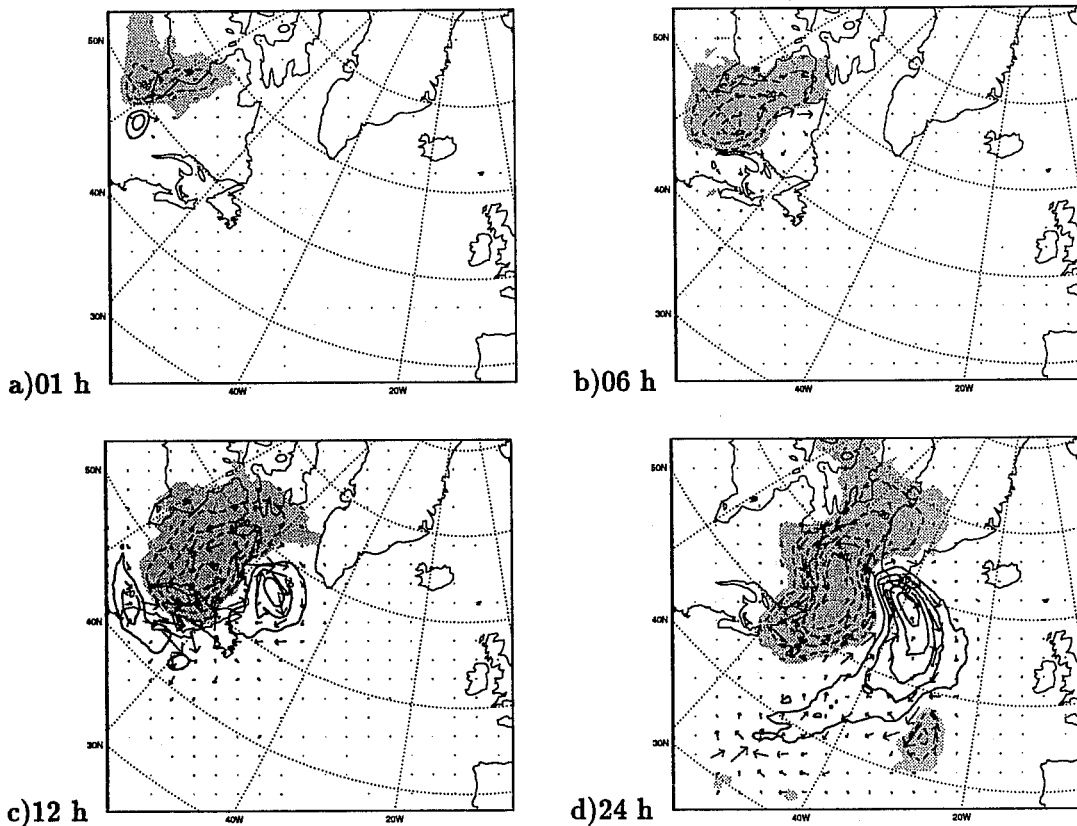
a)01 h b)06 h c)12 h d)24 h

Figure 5: As in Fig. 3, but using initial as well as lateral boundary perturbations for the sensitivity forecast run with $\alpha_i = 0.1$ and $\alpha_b = 5.0$ and with a time-averaging period of ± 1 h: a) 1 h, b) 6 h, c) 12 h, d) 24 h.

directly applied in four dimensional variational data assimilation (4D-Var); during the assimilation period the boundary values can be modified together with modifications in the initial state. 4D-Var assimilation of rainfall observations, including a similar control of lateral boundary conditions, has been tested by Zou and Kuo (1996).

6.3 Limitations of present lateral boundary conditions.

A basic weakness of the lateral boundary relaxation technique was revealed during the experimentation with the adjoint of the scheme. The strongest sensitivity of the forecast errors to the lateral boundary conditions occurred in a region of strong physical inflow with a low pressure system passing the lateral boundaries during the time period of the sensitivity experiment. During the forward model integration, the lateral boundary relaxation scheme managed quite well in this inflow region to introduce the time variation of the meteorological fields as given by the coarser resolution model providing the lateral boundary conditions. During the adjoint model backward integration, however, this region of physical inflow becomes an area of outflow of forecast error gradient information. In the present construction of the HIRLAM boundary relaxation scheme both inflow and outflow points are handled identically. Davies (1983) has shown that inflow of information is handled reasonably well at the same time as outflow information is reasonably well absorbed by the boundary zone damping.

To demonstrate the propagation properties of the boundary zone scheme we have found that the following very simple model system can be used. Consider a linear advection equation where $A(x, t)$ is linearly advected with a $U > 0$. Assume a semi-Lagrangian time stepping scheme with a time step such that the CFL number is equal to one, $\Delta t = \Delta x / U$. This implies that

$$A(x, t + \Delta t) = A(x - \Delta x, t) \tag{107}$$

354

If we now apply a relaxation boundary zone to this problem we may write

$$A(x, t + \Delta t) = g(x) A_b(x, t + \Delta t) + (1 - g(x)) A_i(x, t + \Delta t) \qquad (108)$$

where A_b is the externally prescribed boundary forcing field and A_i is the interior domain field. The boundary zone relaxation function is given by $g(x)$ and we assume the outermost boundary point to be located at $x = 0$. Let us now determine the relative weights of the different external boundary point values at the first fully interior point of the computational domain, i.e. the first point where $g(x) = 0$. Let us call this point x_k. A successive insertion of (107) into (108) gives

$$A(x_k, k\Delta t) = A(x_k - \Delta x, (k-1)\Delta t) =$$
$$g(x_k - \Delta x)A_b(x_k - \Delta x, (k-1)\Delta t) + (1 - g(x_k - \Delta x))A_i(x_k - \Delta x, (k-1)\Delta t) =$$
$$g(x_k - \Delta x)A_b(x_k - \Delta x, (k-1)\Delta t) + (1 - g(x_k - \Delta x)) A(x_k - 2\Delta x, (k-2)\Delta t) \qquad (109)$$

We may continue with such insertions and we will find that the relative weight of each externally prescribed boundary point depends on the form of $g(x)$. Let us assume a four point boundary zone where $g(0) = 1.$; $g(\Delta x) = 0.9$; $g(2\Delta x) = 0.5$; and $g(3\Delta x) = 0.1$. We will then obtain

$$A(4\Delta x, 4\Delta t) = 0.1A_b(3\Delta x, 3\Delta t) + 0.45A_b(2\Delta x, 2\Delta t) + 0.405A_b(\Delta x, \Delta t) + 0.045A_b(0, 0) \qquad (110)$$

thus showing that at the time when the information from the outermost boundary point has reached the interior of the domain very little of the original amplitude remains. Most of the information at the first fully interior point comes from boundary value information at the second and third boundary zone points. This statement may be generalized to other boundary zone functions. The weight of the outer-most boundary point is always a product of several weighting factors significantly less than one and thus will be a small number, while the information coming from the boundary zone point adjacent to the first fully interior point must also be small as the weighting factor must be near zero at this point. Larger values occur at intermediate boundary zone points.

The boundary zone relaxation function used in this study is cosine-shaped and the width of the boundary relaxation zone is 8 gridpoints. A similar advection scheme as in the demonstration above, extended to CFL numbers also smaller than one, was applied to obtain the weights given to the externally prescribed boundary points in the solution just inside the boundary relaxation zone. These sensitivity weights obtained for CFL numbers 1.0 and 0.5 are illustrated in Fig. 6. The general conclusion reached with the help of equation (110) still holds, most of the information comes from the middle of the boundary relaxation zone. It can also be seen that the area of sensitivity is shifted towards the inner integration area for smaller CFL numbers.

In our sensitivity experiments we see clearly that the gradient of the forecast error norm with respect to the boundary conditions will be near zero at the outermost boundary zone points. On the other hand, immediately inside the lateral boundary, within the boundary relaxation zone, the calculated boundary perturbation may well achieve a significant magnitude. Such discontinuities of a perturbation introduced in the lateral boundary relaxation zone will contribute to an enhanced generation of gravity wave noise. Thus we may find that boundary zone perturbations give rise not only to geostrophically well balanced perturbations which propagate into the interior domain, but also to undesired gravity waves which are quickly damped out by horizontal diffusion and the semi-implicit time stepping scheme. This result shows that the lateral boundary relaxation scheme needs to be modified to take the occurrence of inflow and outflow regions into account (Orlanski, 1976).
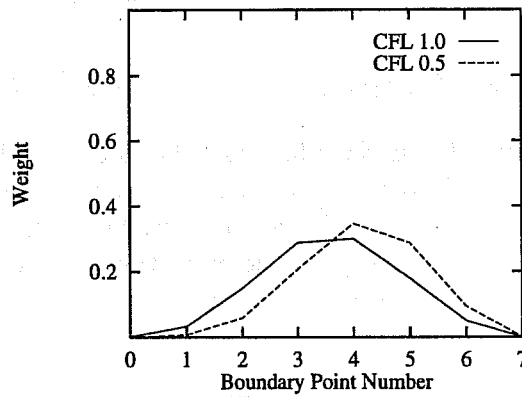
Figure 6: Efficient weights of lateral boundary values as determined by linear advection over the entire boundary relaxation zone to form the model value just inside this zone. Fast advection (CFL=1, full line) and slow advection (CFL=0.5, dashed line).

7 LIMITED AREA 3D-Var.

The variational data assimilation problem may be formulated as the minimization of a cost-function J with respect to the assimilation increments $\delta \mathbf{x} = (\mathbf{x} - \mathbf{x}^b)$, where $\mathbf{x}$ denotes the model state analysis vector and $\mathbf{x}^b$ a model state background vector, normally a short range forecast. In the HIRLAM 3D-Var formulation, the cost function consists of one term J_b, the background error constraint, that measures the distance between the analysis and a background field, and J_o, the observational error constraint that measures the distance between the analysis and the observations:

$$J = J_b + J_o = \frac{1}{2}(\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^b) + \frac{1}{2}(\mathbf{y} - H\mathbf{x})^T \mathbf{R}^{-1}(\mathbf{y} - H\mathbf{x}) \tag{111}$$

Here $\mathbf{y}$ denotes the vector of observed values, H the transformation from model state variables to the observed quantities, $\mathbf{B}$ a matrix with the covariances of background errors and $\mathbf{R}$ a matrix with covariances of observational and representativity errors. The observation operator H may in general be non-linear.

7.1 Analytical balance structure functions.

The spectral HIRLAM provides a good framework for construction of a first version of HIRLAM 3D-Var. The introduction of statistical assumptions like homogeneity and isotropy with regard to the spatial correlations included in $\mathbf{B}$ takes a very simple form in spectral space. The following ideas were thus applied for the first version of the HIRLAM 3D-Var background error constraint J_b:

- By 2-dimensional horizontal Fourier transforms to control variables in spectral space, it is possible to assume that spectral components representing different horizontal wave-numbers are statistically independent.

- For the HIRLAM 3D-Var we have chosen to apply an area-extension to obtain periodic variations in both horizontal dimensions.

- By subtraction of the geostrophic analysis increments, as calculated from the mass field analysis increments, it is possible to assume that the mass field and the ageostrophic wind field analysis increments are statistically independent.

- By projection on eigen-vectors of vertical background error correlation matrices, it is possible to assume that coefficients of different vertical eigen-vectors are statistically independent.

Through such transformations of a model state increment vector δx, it is possible to obtain an assimilation control vector χ for which the error covariance matrix is a unit matrix. This also gives an efficient pre-conditioning of the minimization problem. Non-separable analysis structure functions, based on the model assumptions described above, have been derived by the NMC-technique from the time history of +24h and +48h HIRLAM forecasts (Berre, 1997). The basic idea of the NMC-technique is that spatial structures of differences between e.g. +24h and +48 h forecasts valid at the same time are similar to the spatial structures of +6h forecast errors. The derived analysis structure functions indicated a clear vertical-horizontal non-separability, for example with regard to horizontal length-scales of different variables at different levels.

7.2 Empirical balance structure functions.

A more general J_b based on empirical balance relationships (Parrish et al., 1997) has been implemented for the HIRLAM 3D-Var (Berre, 1998). The balance relationships take the following general form:

$$
\begin{aligned}
\zeta &= \zeta \\
\eta &= \mathcal{MH}\zeta + \eta_u \\
(T, p_s) &= \mathcal{NH}\zeta + \mathcal{P}\eta_u + (T, p_s)_u \\
q &= \mathcal{QH}\zeta + \mathcal{R}\eta_u + \mathcal{S}(T, p_s)_u + q_u
\end{aligned}
$$

where $(\zeta, \eta, (T, p_s), q)$ are the forecast errors of vorticity, divergence, temperature, surface pressure and specific humidity and where $(\eta_u, (T, p_s)_u, q_u)$ are the un-balanced parts of η, (T, p_s) and q. $\mathcal{H}$ is a horizontal balance operator that relates vorticity to a balanced geopotential P_b. $\mathcal{M}$, $\mathcal{N}$, $\mathcal{P}$, $\mathcal{Q}$, $\mathcal{R}$ and $\mathcal{S}$ are all vertical balance operators relating the spectral coefficients of the predictors with the spectral coefficients of the predictands. A novel feature of the scheme is the possibility to utilize fully multivariate relationships for the assimilation of moisture together with the other model variables.

The NMC method was applied to derive coefficients in the empirical model for forecast errors described above. It was first applied to +12h and +36h forecasts from the Aladin limited area model (Berre, 1998) and later on to +24h and +48h HIRLAM model forecasts (Berre, personal communication). Results from the two trial applications are consistent. Horizontally averaged covariances between balanced geopotential and moisture forecast errors are plotted as a function of level number pair in Figure 7. We can notice a strong negative covariance between low level balanced geopotential and low level moisture forecast errors. Efficiently this means that under-prediction of a low pressure development is associated with under-prediction of moistening in the lower troposphere. The full multivariate relationships include also other significant contributions. The total average variance of moisture forecast errors explained statistically by other forecast variable errors is of the order of 30%. Assimilation experiments to investigate the impact of these multivariate relationships are carried out at present.

7.3 Single observation impact studies.

A first validation of the HIRLAM 3D-Var structure function formulation is to study the impact of single observations. The analysis increments resulting from a single 500 hPa temperature observation increment of +1 K in a position close to the right lateral boundary have been plotted for temperature increments on model level 8 in Figure 8a and for wind increments on model level 8 in Figure 8b. The analytical balance formulation for 16 vertical levels was used in this particular study with model level 8 being close to the observation level 500 hPa. We may notice the rather strict isotropic influence of the temperature observation, the corresponding wind increments in near geostrophic balance and the lateral boundary formulation causing no obvious detrimental effects. The width of the boundary extension zone is an important parameter in this respect, it was chosen to be 18 grid-lengths or 900 km in the x-direction in this case, and this seems to be sufficient to avoid any influence of an observation close to the right boundary on analysis increments close to the left boundary of the area.
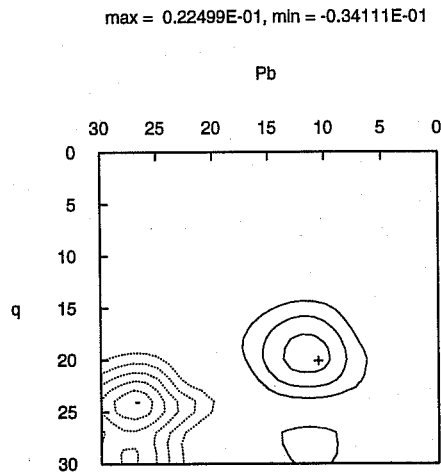
max = 0.22499E-01, min = -0.34111E-01

Pb

Figure 7: Spectral averages of vertical cross-covariances between balanced geopotential (derived from vorticity) and specific humidity.
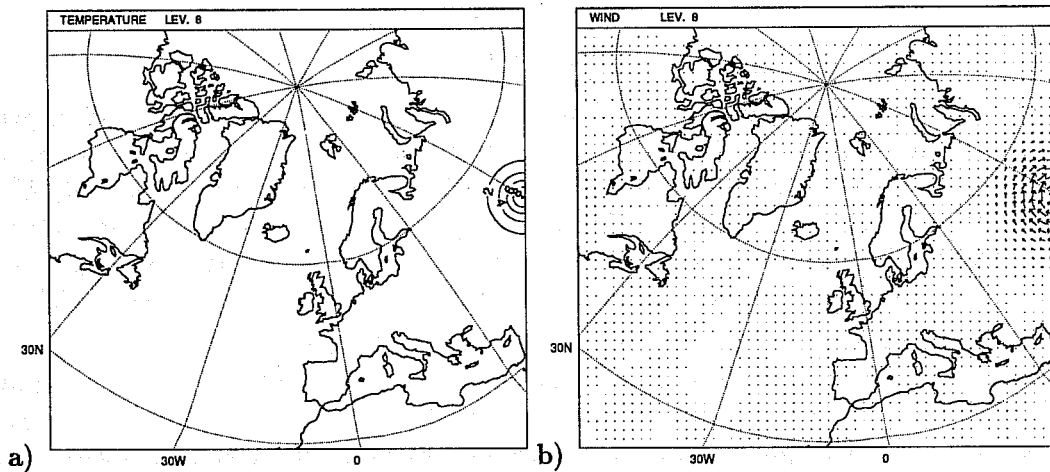

Figure 8: Impact of a single 500 hPa temperature observation on a) model level 8 temperature analysis increments, b)model level 8 wind analysis increments. Wind arrows have been plotted in every 2nd gridpoint. Isoline spacing is 0.2 K.

7.4 Short parallel run to compare OI and 3D-Var.

As a first test for HIRLAM 3D-Var, we selected a case with a forecast failure experienced with the operational HIRLAM at INM in Madrid. Forecasts valid at 00 UTC on the 6th of November 1997 for an explosive cyclogenesis over Spain were rather poor for forecast lengths longer than +12 h. For the data assimilation experiment with the HIRLAM 3D-Var, we used the SMHI 55 km forecast area. Furthermore, the non-separable analysis structure functions derived from operational forecasts were used. In order to apply background error standard deviations representative of 6 h forecast errors, the standard deviations derived by the NMC method were scaled by a factor of 0.8 for the temperature, surface pressure and specific humidity and with a factor of 0.5 for the ageostrophic wind components. TEMP, PILOT, AIREP, SYNOP, SHIP and DRIBU observations were utilized. Data assimilation started from 3 November 1997 00 UTC with a 6 h data assimilation cycle. Data assimilation was carried out until 7 November 12 UTC. The spectral HIRLAM with HIRLAM 2.4 physics was used for assimilation as well as forecast cycles. ECMWF analyses every 6th hour were used as lateral boundary conditions.

A HIRLAM reference system experiment was carried out in addition. This experiment differed from the 3D-Var experiment in the following aspects: (1) The HIRLAM reference system Optimum Interpolation (OI) analysis was used. (2) The grid point HIRLAM with HIRLAM 2.4 physics was used for assimilation as well as forecast cycles.

There are several important differences between the two experiments. The forecast model in 3D-Var experiment is the spectral version of HIRLAM, while the grid point HIRLAM is used for the OI experiment. The 3D-Var has not yet been adapted to the grid point HIRLAM, and we considered it more important to carry out a first *demonstration* of 3D-Var than to wait for the grid point model adaptation of 3D-Var. In order to put some light on the degree of impact from the forecast model formulation, we have carried out some spectral model integrations from the OI initial data in addition. Thus, the following combinations of experiments will be briefly validated below:

A : OI analysis. Grid point model for assimilation cycles as well as for the forecasts. Verification against its own analysis.

B : 3D-Var analysis. Spectral model for assimilation cycles as well as for the forecasts. Verification against its own analysis.

C : Spectral model forecasts with initial analyses taken from **A**. Verification against **A** analyses.

D : Spectral model forecasts with initial analyses taken from **A**. Verification against **B** analyses.

Forecasts up to +48 h were carried out for 5 different cases - 3 November 1997 12 UTC, 4 - 5 November 1997 00 and 12 UTC. These forecasts were verified against analysis fields over a European area. RMS verification scores are presented for +24 h and +48 h forecasts in Table 1 below.

Table 1: Verification against analyses over a European area. Root Mean Square (RMS) verification scores for +24 h and + 48 h :

+24 h forecast	A	B	C	D
Mean sea level pressure (hPa)	2.0	2.0	2.0	2.0
500 hPa height (m)	19.	16.	18.	18.
300 hPa height (m)	27.	24.	27.	26.
925 hPa temperature (K)	1.7	1.2	1.7	1.4
700 hPa temperature (K)	1.1	1.1	1.2	1.2
850 hPa relative humidity	14%	14%	16%	16%
+48 h forecast	A	B	C	D
Mean sea level pressure (hPa)	3.7	3.2	3.7	3.6
500 hPa height (m)	32.	27.	31.	31.
300 hPa height (m)	45.	40.	44.	44.
925 hPa temperature (K)	2.3	1.8	2.4	2.0
700 hPa temperature (K)	1.7	1.6	1.7	1.7
850 hPa relative humidity	20%	19%	20%	19%

¿From these forecast verification scores, we can first notice that the RMS values for experiment B (3D-Var + spectral model) are always smaller than or equal to the RMS values for experiment A (OI + grid point model). Let us first consider the differences in verification scores for mean sea level pressure, 500 hPa height and 300 hPa height. These may be taken as representative of forecasting ability for synoptic scale developments. It is encouraging that these scores are better for the 3D-Var based forecasts than for the OI based forecasts. Considering also the verification scores for the +48 h spectral model forecasts based on the OI analyses (experiments C and D), it appears reasonable to conclude that this positive impact originates from assimilation technique and not from the model formulation. Certainly, it is necessary to be careful and not to make too firm conclusions, since the forecast sample is small.

It is also interesting to note the lower RMS verification scores for the 3D-Var based forecasts of low level temperatures. One may speculate whether this is a result of real forecast improvements, for example, due to the use of significant level data in 3D-Var. It may be, however, that this verification result is simply caused by the 3D-Var verification analyses being too biased towards the analysis background field (a 6 h forecast). Again considering the scores for the spectral model forecasts based on OI initial data, we see that RMS scores for low level temperature forecasts are smaller when 3D-Var analyses are used for the verification than if OI-analyses are used. Thus, we cannot exclude that the improved verification scores for 3D-Var-based lower level temperature forecasts are just artifacts of biased verification analysis fields.

8 COMPUTATIONAL EFFICIENCY AND PARALLELIZATION.

Gustafsson and McDonald (1996) compared the computational costs of the gridpoint and spectral semi-Lagrangian HIRLAM models. Table 2 below includes measurements of CPU-time in seconds for an average time-step and the average time spent in various parts of the codes on a CONVEX C-3810 vector processor.

Table 2: CPU-time in seconds, measured on a single processor CONVEX C-3810 computer, for the various types of calculations during one time-step of the gridpoint and spectral HIRLAM models.

TYPE OF CALCULATION	GRIDPOINT MODEL	SPECTRAL MODEL
Fourier transforms		4.2
Non-linear dynamics	1.2	1.0
Semi-Lagrangian interpolations	7.6	7.3
Physics	6.6	6.5
Semi-implicit part	1.8	0.5
Implicit horizontal diffusion	2.7	0.1
Boundary relaxation	0.2	0.2
Total per time-step	20.2	19.8

We may notice that the total time for a time-step is almost identical for the two model versions. Time spent in Fourier transforms by the spectral version is balanced by larger time spent in the semi-implicit solver and in implicit horizontal diffusion by the gridpoint version. We may conclude that there is no difference between the two model versions from a computational economy point of view.

The spectral version of the HIRLAM model was implemented on distributed memory parallel computers by Gustafsson and Salmond (1994). The basic parallelization strategy, utilized for forecast model integrations as well as for the calculation of the spatial transforms in the 3D-Var is briefly illustrated in Figure 9. The calculations include three different types of transforms that have to be considered in the parallelization strategy: (1) FFTs in the x-direction, (2) FTTs in the y-direction and (3) vertical transforms. The basic data transposition ideas applied are that each one-dimensional transform should be calculated on a single processor only, that the number of transforms calculated on each processor should be as near equal as possible to balance the work load, and that data should be re-distributed among the processors between the transforms in the different spatial directions. The data transposition scheme is illustrated for a four (4) processor parallel computer in Figure 9.
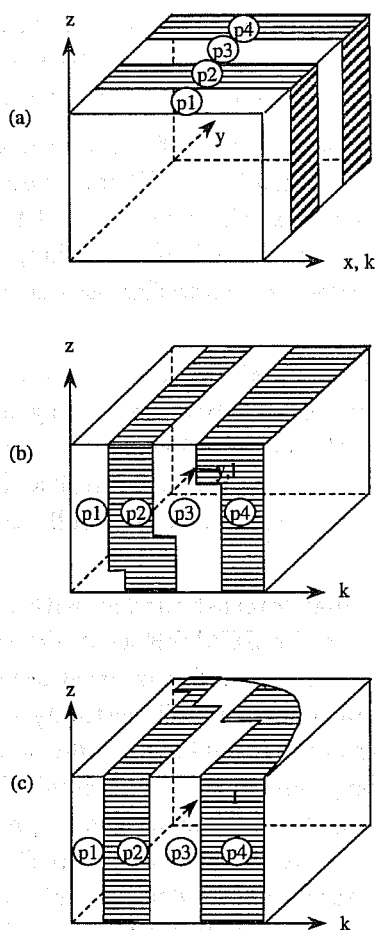
Figure 9: Data distribution over four processors p1, p2, p3 and p4 during transform calculations in the spectral HIRLAM and in HIRLAM variational data assimilation: (a) x-direction Fourier transforms, (b) y-direction Fourier transforms and (c) vertical transforms. The three spatial dimensions of the HIRLAM model domain are indicated by (x, y, z), the wave number in the x-directions by k and the wave number in the y-direction by l.

The spectral parallel HIRLAM scales well with the number computational processors as is illustrated in Table 3 for test runs on a Cray T3D computer carried out by Gustafsson and Salmond (1994).

Table 3: Performance of the parallel spectral HIRLAM with 256 x 256 horizontal gridpoints and 16 levels on the CRAY T3D.

NPE's	Typical Step Time (Seconds)	Speed-up over 16 PE's
16	10.219	1.00
32	5.265	1.94
64	2.645	3.86
128	1.344	7.60
256	0.720	14.20

9   CONCLUDING REMARKS

A spectral version of the HIRLAM forecast model has been developed. This model uses the same continuous model equations, the same vertical discretization and the same physical parameterizations as the grid-point version of HIRLAM. Gustafsson and McDonald (1996) compared the semi-Lagrangian versions of the spectral and grid-point HIRLAM in a number of model test integrations. They concluded that the competition between the two model versions had been very productive and ended in a

draw. By a systematic comparisons of results from the two different models, both models may be tuned and improved to provide a similar forecast quality on the basis of an almost identical computational cost. The strategy of the HIRLAM Project is therefore to keep both model versions up-to-date.

The spectral HIRLAM is being used as a framework for development of a variational data assimilation system. The tangent linear and the adjoint of the adiabatic version of the model have been used in sensitivity experiments. Experiments to test the sensitivity of forecast errors with respect to lateral boundary conditions have indicated the potential of controlling lateral boundary conditions during limited area 4D-Var but have also revealed serious weaknesses in the present mathematical formulation of the lateral boundary conditions.

The development of a 3-dimensional variational data assimilation (3D-Var) has almost been completed. Non-separable analysis structure functions have been derived by the NMC-technique. The first version is based on analytical expressions for the balance between the mass field and the wind field analysis increments while further developments are based on empirical balance conditions and multivariate relationships including also the moisture field. The first full scale tests with the HIRLAM 3D-Var indicates highly promising results.

The application of high resolution global forecast models with sophisticated 4-dimensional data assimilation schemes at forecasting centers like ECMWF naturally raises several questions related to the future of limited area modelling activities at smaller weather services. High resolution modelling, i.e. mesoscale modelling with non-hydrostatic models, will certainly be meaningful, but how do we acquire fresh and high quality lateral boundaries for these high resolution models? Is limited area 4D-Var, and in particular control of lateral boundary conditions, meaningful? Would it not be more meaningful for the smaller weather services to apply global models in a joint project to produce best possible initial and lateral boundary conditions for the mesoscale model runs? The answers to these questions are not obvious. However, international numerical modelling centers like ECMWF should also take into account that they are highly dependent on NWP activities in the smaller weather services, when it comes to, for example, recruitment of competent staff.

References

Berre, L., 1997: Non-separable structure functions for the HIRLAM 3DVAR. *HIRLAM Technical report*, **30**, November 1997, 40 pp.

Berre, L., 1998: Estimation of synoptic and meso scale forecast error covariances in a limited area model. *Mon. Wea. Rev.*, accepted.

Davies, H. C., 1983: Limitations of some common lateral boundary schemes used in NWP models. *Mon. Wea. Rev.*, **111**, 1002–1012.

Errico, R. M., Vukićević, T. and Raeder, K., 1993: Comparison of initial and lateral boundary condition sensitivity for a limited-area model. *Tellus*, **45A**, 539–557.

Gustafsson, N., 1990: Sensitivity of limited area model data assimilation to lateral boundary condition fields. *Tellus*, **42A**, 109–115.

Gustafsson, N., 1991: The HIRLAM model. *Proceedings of Seminar on Numerical Methods in Atmospheric Models*. ECMWF, Reading, UK, 9-13 September 1991, Volume 2, 115–146.

Gustafsson, N. and Salmond, D., 1994: A parallel spectral HIRLAM with the data parallel programming model and with message passing. A comparison. In *Coming of Age*, Proceedings of the 6th ECMWF Workshop on the use of Parallel processors in Meteorology. Reading, UK, 21-25 November, 1994. World Scientific Publishing Co., 32–48.

Gustafsson, N. and Huang, X.-Y., 1996: Sensitivity experiments with the spectral HIRLAM and its adjoint. *Tellus*, **48A**, 501–517.

Gustafsson, N. and A. McDonald, 1996: A comparison of the HIRLAM gridpoint and spectral semi-

Lagrangian models. *Mon. Wea. Rev.*, **124**, 2008–2022.

Gustafsson, N., Lönnberg, P. and Pailleux, J., 1997: Data assimilation for high resolution limited area models. *J. of Met. Soc. of Japan*, **75**, 367–382.

Gustafsson, N., Källén, E. and Thorsteinsson, S., 1998. Sensitivity of forecast errors to initial and lateral boundary conditions. *Tellus*, **50A**,167–185.

Gustafsson, H., Hörnquist, S., Lindskog, M., Berre, L., Navascues, B., Thorsteinsson, S., Huang, X.-Y., Mogensen, K.S. and Rantakokko, J., 1999: Three-dimensional variational data assimilation for a high-resolution limited area model (HIRLAM). *HIRLAM technical Report*, No. 40, 72 pp.

Haugen, J.-E. and B. Machenhauer, 1993: A Spectral Limited-Area Model Formulation with time-dependent boundary conditions applied to the shallow-water equations. *Mon. Wea. Rev.*, **121**, 2631–2636.

Huang, X.-Y., Gustafsson, N. and Källén, E., 1997: Using an adjoint model to improve an optimum interpolation based data assimilation system. *Tellus*, **49A**, 161–176.

Källén, E., 1996: HIRLAM Documentation Manual, System 2.5, Available from SMHI, S-60176 Norrköping, Sweden.

Ide, K., Bennet, A.F., Courtier, P., Ghil, M. and Lorenc, A.C., 1997: Unified notations for data assimilation: Operational, sequential and variational. *J. Meteorol. Soc. Japan*, **75**, 181–189.

Le Dimet, F.X. and Talagrand, O., 1986: Variational Algorithms for Analysis and Assimilation of Meteorological Observations. Theoretical Aspects. *Tellus*, **37A**, 309–322.

Lewis, J.M. and Derber, J.C., 1985: The Use of Adjoint Equations to Solve a Variational Adjustment Problem with Advective Constraints. *Tellus*, **37A**, 309–322.

Machenhauer, B., 1977: On the dynamics of gravity oscillations in a shallow water model with applications to norlam mode initialization. *Beitr. Atmos. Phys.*, **50**, 253–271.

McDonald, A. and J.-E. Haugen, 1993: A two time-level, three-dimensional, semi-Lagrangian, semi-implicit limited area gridpoint model of the primitive equations. Part II: Extension to hybrid vertical coordinates. *Mon. Wea. Rev.*, **121**, No. 7, July 1993.

Orlanski, I., 1976: A simple boundary condition for unbounded hyperbolic flows. *J. Comp. Phys.*, **21**, 251–269.

Parrish, D.F., Derber, J.C., Purser, R.J., Wu, W.-S. and Pu, Z.-X., 1997: The NCEP global analysis system: recent improvements and future plans. *J. of Met Soc. of Japan*, **75**, No. 1B, 359–365.

Rabier, F., Klinker, E., Courtier, P. and Hollingsworth, A., 1996: Sensitivity of two-day forecast errors over the Northern Hemisphere to initial conditions. *Q. J. R. Meteorol. Soc.*, **121**, 121–150.

Radnoti, G., 1995: Comments on "A Spectral Limited-Area Formulation with Time-dependent Boundary Conditions Applied to the Shallow-Water Equations". *Mon. Wea. Rev.*, **123**, 3122–3123.

Ritchie, H., A. Simmons, M. Hortal, T. Davies, D. Dent and M. Hamrud, 1995: Implementation of the semi-Lagrangian method is a high resolution version of the ECMWF forecast model. *Mon. Wea. Rev.*, **123**, 489–514.

Simmons, A.J. and D.M. Burridge, 1981: An energy and angular momentum conserving vertical finite difference scheme and hybrid vertical coordinates. *Mon. Wea. Rev.*, **109**, 758–766.

Thepaut, J.-N. and Courtier, P., 1991: Four-dimensional variational data assimilation using the adjoint of a multilevel primitive-equation model. *Q.J.R. Meteorol. Soc.*, **117**, 1225–1254.

Zou, X. and Kuo, Y.-H., 1996: Rainfall assimilation through an optimal control of initial and boundary conditions in a limited-area mesoscale model. *Mon. Wea. Rev.*, **124**, 2859–2882.