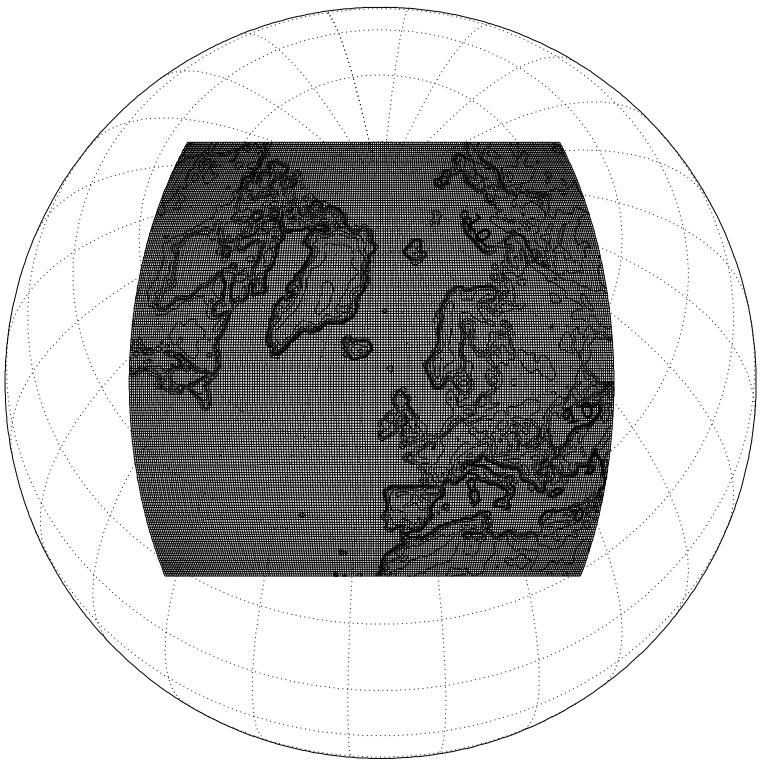
High Resolution Limited Area Model

HIRLAM-5 Scientific Documentation

December 2002



Per Undén, Laura Rontu, Heikki Järvinen, Peter Lynch, Javier Calvo, Gerard Cats, Joan Cuxart, Kalle Eerola, Carl Fortelius, Jose Antonio Garcia-Moya, Colin Jones, Geert Lenderlink, Aidan McDonald, Ray McGrath, Beatriz Navascues, Niels Woetman Nielsen, Viel Ødegaard, Ernesto Rodriguez, Markku Rummukainen,Rein Rõõm, Kai Sattler, Bent Hansen Sass, Hannu Savijärvi, Ben Wichers Schreur, Robert Sigg, Han The, Aleksander Tijm

HIRLAM-5 Project, c/o Per Undén SMHI, S-601 76 Norrköping, SWEDEN

Contents

Chapter 1

Introduction

The HIRLAM-5 project has been established in order to provide the best available operational short-range forecasting system in the member institutes. These are the National Meteorological Services in Denmark, Finland, Iceland, Ireland, Netherlands, Norway, Spain and Sweden. Météo-France has a research cooperation agreement with Hirlam. HIRLAM-5 is a continuation of the HIRLAM-4 project and its predecessors. The developments from earlier projects have been carried over and further developed together with a number of new activities.

The HIRLAM-5 members have signed a Memorandum of Understanding (MoU), which details the main objectives, targets and priorities for the project. The main objective is to provide an operational production system for short-range forecasting which generates a comprehensive set of Numerical Weather Prediction (NWP) products of highest quality.

The Hirlam system is a complete NWP system including Data Assimilation with analysis of conventional and non-conventional observations and a limited area forecasting model with a comprehensive set of physical parameterisation. This system is owned and developed jointly by the Hirlam member institutes.

The forecast model is a limited area model with a boundary relaxation scheme. The model exists both in a grid-point version and in a spectral version. Most applications use the grid-point model, except for the variational Data Assimilation, where the spectral method is used in the formulation of the background constraint and the inner loop of 4D-VAR uses the spectral model.

There is also a non-hydrostatic dynamical kernel of the grid-point model. This has been developed at the Tartu Observatory in Estonia through a research cooperation with Hirlam members.

The Hirlam physical parameterisation includes a tiled surface scheme, a TKE turbulence parameterisation, a condensation and a mass-flux type convection scheme, a radiation scheme and a gravity wave drag parameterisation.

The analysis of upper air variables is using a 3-dimensional variational method (3D-VAR) or the earlier 3-dimensional optimum interpolation (OI) method. The variational configuration also includes 4-dimensional variational assimilation (4D-VAR). Surface variables are analysed using successive correction or optimum interpolation, depending on the variables. Although there is some pre-processing of observations in Hirlam the system does not provide direct interface to e.g. WMO observations. It requires that data are extracted to WMO BUFR files beforehand. Initialisation of the analysis can be either through the implicit Normal Mode method (NMI) or Digital Filtering (DFI).

An important component for the realism and flexibility of Hirlam is to have a comprehensive climate system from which the physiographic and climatic data for the model can be generated. This is included and regularly enhanced as needs arise.

The Hirlam system includes a standard verification package against observations and analysis fields. Enhanced methods for meso-scale verification are under development. A wide range of model diagnostics can be produced, optionally, for process and circulation diagnostics.

The system is necessarily a large one and is quite complex. It is required to run on a wide range of computer hardware using Unix or Linux operating systems. It has been optimised both for vectorised system and parallel processors. Hirlam includes a great number of tasks in its setting up as well as in the running phase of an experiment or production run. The flow of jobs and task is managed through a scheduling system, developed from the one used at ECMWF.

The complete Hirlam system with its program codes and scripting system is maintained at ECMWF in the form of the so called Reference System. This is updated as development proceeds or the technical environment changes several times a year. A few times a year a number of these changes are tested together comprehensively and approved for an official release of the Reference System. For continuous validation purposes a delayed mode run (DMR) is kept running in near-real time at ECMWF and this uses the Reference System. The national operational versions normally use somewhat older versions of the Reference System. Many members also apply local modifications or adaptations of the system.

A large number of people have contributed to the development of the individual components in the Hirlam system. The list of authors of this publication is by no way complete or representing fairly the development of the system. The authors are the ones who volunteered to write or put together and edit the material required for each component. The scientific and technical contributions involve a much larger number of people in the Hirlam community.

This documentation covers in some parts alternative schemes, e.g. for analysis or convective parameterisation and this has been done for completeness, as either scheme may be used by Hirlam members.

A list of symbols can be found in Appendix 1. These variables and constants are used extensively and not explicitly defined in the text.

Chapter 2

The Model Equations and their Discretization

In this section is presented a derivation of a semi-Lagrangian discretisation of the multi-level primitive equations using the hybrid coordinate (η) in the vertical. The equations are actually written for a general map projection. However, since they are almost exclusively used on transformed latitude-longitude grids they are thus described herein.

Unless otherwise stated the notation is conventional. See appendix 1 for definitions of the symbols. As far as the vertical discretisation is concerned, the wind $\mathbf{v}_k$, temperature T_k, specific humidity q_k, and linearised geopotential height G_k are defined at the 'full levels' (k = 1,N); N is the total number of full levels in the vertical. The pressure $p_{k+1/2}$, geopotential height $\Phi_{k+1/2}$, and vertical velocities $\dot{\eta}_{k+1/2}$ and $\dot{s}_{k+1/2}$ are defined at the 'half levels'. For the horizontal discretisation, the Arakawa C-grid is used.

The hybrid-coordinate is defined in terms of the pressure as follows:

$$p_{k+1/2} = A_{k+1/2}(\eta) + B_{k+1/2}(\eta)p_s(\lambda,\theta) \tag{2.1}$$

where the choice of A and B defines the closeness of the system to the σ-coordinates ($A = 0.$) or p coordinates ($B = 0.$).

2.1 The continuity equation

The continuity equation in hybrid coordinates is

$$\left(\frac{d_H}{dt} + D\right)\frac{\partial p}{\partial \eta} + \frac{\partial}{\partial \eta}\left(\dot{\eta}\frac{\partial p}{\partial \eta}\right) = 0 \tag{2.2}$$

where $D = \nabla_H.\mathbf{v}_H$ and

$$\frac{d_H}{dt} = \frac{\partial}{\partial t} + \mathbf{v}_H.\nabla_H \tag{2.3}$$

When discretised in the vertical, (2.2) gives

$$\frac{d_H \Delta p_k}{dt_k} = -D_k \Delta p_k - \left(\dot{\eta}\frac{\partial p}{\partial \eta}\right)_{k+1/2} + \left(\dot{\eta}\frac{\partial p}{\partial \eta}\right)_{k-1/2}. \tag{2.4}$$

Substituting $\Delta p_k = \Delta A_k + \Delta B_k p_s$, and dividing by p_s yields

$$\Delta B_k \frac{d_H ln p_s}{dt_k} = -D_k \frac{\Delta p_k}{p_s} - (\dot{s}_{k+1/2} - \dot{s}_{k-1/2}) \tag{2.5}$$

where

$$\dot{s}_{k-1/2} = \frac{1}{p_s}\left(\dot{\eta}\frac{\partial p}{\partial \eta}\right)_{k-1/2} \tag{2.6}$$

This 'vertical velocity' is computed diagnostically by partially summing (2.5) in the vertical:

$$(\dot{s})_{k+1/2}^{n+1/2} = -B_{k+1/2}\left(\frac{\partial lnp_s}{\partial t}\right)^{n+1/2} - \frac{1}{p_s}\left[\sum_{j=1}^{k}\nabla_H.\{(\mathbf{v}_H)_j\Delta p_j\}\right]^{n+1/2} \tag{2.7}$$

where

$$\frac{\partial lnp_s}{\partial t} = -\frac{1}{p_s}\sum_{j=1}^{N}\nabla_H.\{(\mathbf{v}_H)_j\Delta p_j\} \tag{2.8}$$

is computed by summing (2.5) in the vertical over all levels. (The upper boundary condition is $\dot{\eta}=0$). From here on $\partial lnp_s/\partial t$ will always be defined by (2.8).

To update lnp_s we first separate out the orographic term in order to implement the 'Eulerian advection of orography' scheme of Ritchie and Tanguay (1996). Define

$$lnp_s' = lnp_s + \frac{\phi_s}{RT^r} \tag{2.9}$$

Simultaneously we substitute the diagnostically computed $\dot{s}$ from (2.7) into (2.5) The latter then becomes

$$\Delta B_k \frac{d_H}{dt_k} lnp_s' = \left(\frac{\partial lnp_s}{\partial t} + (\mathbf{v}_H)_k.\nabla_H lnp_s'\}\right)\Delta B_k \tag{2.10}$$

Now linearising in the traditional fashion yields the following discretization

$$\begin{aligned}(lnp_s')^{n+1} \quad &+ \quad \left(\frac{\Delta t_+}{2}\right)\sum_{j=1}^{N}\frac{\Delta p_j^r}{p^r}D_j^{n+1} = \sum_{k=1}^{N}\Delta B_k\left(\frac{\Delta t_+}{2}\right)(N_p)_k^{n+1/2} \\ &+ \quad \sum_{k=1}^{N}\Delta B_k\left[(lnp_s')^n + \left(\frac{\Delta t_-}{2}\right)\left\{-\sum_{j=1}^{N}\frac{\Delta p_j^r}{p^r}D_j^n + (N_p)_k^{n+1/2}\right\}\right]_{*3},\end{aligned} \tag{2.11}$$

where the subscript '$*3$' indicates a three-dimensional interpolation to evaluate the value of the field at the departure point. Δt_+ and Δt_- are defined in (2.52) and

$$(N_p)_k = \frac{\partial lnp_s}{\partial t} + (\mathbf{v}_H)_k.\nabla_H lnp_s' + \sum_{j=1}^{N}\frac{\Delta p_j^r}{p^r}D_j. \tag{2.12}$$

2.2 The temperature equation

In the hybrid coordinate system the temperature equation can be written as

$$\frac{dT_k}{dt_k} = \left[\frac{\kappa T_v}{1+(\delta-1)q}\right]_k\left(\frac{\omega}{p}\right)_k + (P_T + K_T)_k, \tag{2.13}$$

where $(T_v)_k = [1+(\frac{1}{\epsilon}-1)q_k]T_k$. The terms P_T and K_T represent the physical and diffusion processes respectively. Their discretization will be discussed separately in sections below. This section will concentrate on the discretization of the adiabatic terms.

$$\frac{d}{dt} = \frac{d_H}{dt} + \dot{\eta}\frac{\partial}{\partial \eta}, \tag{2.14}$$

and

$$\left(\frac{\omega}{p}\right)_k = \left[\left(\frac{d_H p}{dt} + \dot{\eta}\frac{\partial p}{\partial \eta}\right)\frac{1}{p}\right]_k. \tag{2.15}$$

The first term on the right hand side of (2.13) is estimated at the full levels k as follows.

$$\begin{aligned}\left[\frac{\kappa T_v}{1+(\delta-1)q}\left(\frac{\omega}{p}\right)\right]_k &= \frac{\kappa (T_v)_k}{[1+(\delta-1)q_k]\Delta p_k}\Big[\Delta ln p_k\Big(p_s\frac{\partial ln p_s}{\partial t} + \sum_{j=k+1}^{N}\nabla_H.\{(\mathbf{v}_H)_j\Delta p_j\}\Big) \\ &+ (\Delta ln p_k - \alpha_k)\nabla_H.\{(\mathbf{v}_H)_k\Delta p_k\}\Big],\end{aligned} \tag{2.16}$$

where $\alpha_k = 1 - (\Delta lnp/\Delta p)_k p_{k-1/2}$ when $k = 2....N$, and $\alpha_1 = ln2$. Linearising ω/p in the traditional manner,

$$\left(\frac{\omega}{p}\right)^r_k = -\left(\frac{\Delta lnp}{\Delta p}\right)^r_k \sum_{j=1}^{k-1} D_j \Delta p^r_j - \alpha^r_k D_k \tag{2.17}$$

enables us to write the following SLSI discretization of (2.16):

$$\begin{aligned}&T^{n+1}_k + \kappa T^r\left(\frac{\Delta t_+}{2}\right)\left\{\left(\frac{\Delta lnp}{\Delta p}\right)^r_k \sum_{j=1}^{k-1} D^{n+1}_j \Delta p^r_j + \alpha^r_k D^{n+1}_k\right\} = \left(\frac{\Delta t_+}{2}\right)(N_T)^{n+1/2}_k \\ &+\left[T^n_k - \kappa T^r\left(\frac{\Delta t_-}{2}\right)\left\{\left(\frac{\Delta lnp}{\Delta p}\right)^r_k \sum_{j=1}^{k-1} D^n_j \Delta p^r_j + \alpha^r_k D^n_k\right\} + \left(\frac{\Delta t_-}{2}\right)(N_T)^{n+1/2}\right]_{*3},\end{aligned} \tag{2.18}$$

where

$$(N_T)_k = \left[\frac{\kappa T_v}{1+(\delta-1)q}\left(\frac{\omega}{p}\right)\right]_k - \kappa T^r\left(\frac{\omega}{p}\right)^r_k. \tag{2.19}$$

If the variable G is defined by

$$G_k = \Phi_s + R_d \sum_{j=k+1}^{N}(T\Delta lnp^r)_j + R_d(\alpha^r T)_k + R_d T^0 ln p_s, \tag{2.20}$$

then (2.11), (2.18) and (2.20) can be combined to get an equation for G^{n+1}_k and D^{n+1}_k in the standard form, written in matrix notation as follows:

$$\mathbf{G}^{n+1} + \left(\frac{\Delta t_+}{2}\right)(\gamma\tau + R_d T^0 \nu.)\mathbf{D}^{n+1} = \mathbf{H}, \tag{2.21}$$

where

$$\mathbf{H} = \Phi_s + \gamma\mathbf{A}_T + (\gamma\tau + R_d T^0\nu.)\mathbf{A}_p; \tag{2.22}$$

$\mathbf{H}, \mathbf{G}, \mathbf{D}, \mathbf{A}_T$ and $\mathbf{A}_p$ are N -dimensional vectors and γ, τ and ν are matrix operators which are defined in appendix 2.

2.3 The Momentum Equations

In the hybrid coordinate system the momentum equations can be written as follows

$$\frac{d\mathbf{v}_k}{dt_k} = [-f\mathbf{k}\times\mathbf{v} - \nabla\Phi - R_d T_v \nabla lnp]_k + (\mathbf{P_u} + \mathbf{K_u})_k. \tag{2.23}$$

$\mathbf{P_u}$ and $\mathbf{K_u}$ represent the physical and diffusion processes respectively. Their discretization will be discussed separately below. In this section the discretization of the adiabatic terms will be described.

For consistency, Φ and ∇lnp are defined at level k as

$$\Phi_k = \Phi_s + R_d \sum_{j=k+1}^{N} (T_v \Delta lnp)_j + R_d(\alpha T_v)_k, \tag{2.24}$$

and

$$(\nabla lnp)_k = \nabla\left(\frac{p_{k+1/2}\, lnp_{k+1/2} - p_{k-1/2}\, lnp_{k-1/2}}{\Delta p_k}\right). \tag{2.25}$$

Substituting G from (2.20) yields the following set of linearised equations:

$$\frac{du_k}{dt_k} = -\frac{1}{acos\theta}\frac{\partial G_k}{\partial\lambda} + f v_k + (N_u)_k \tag{2.26}$$

$$\frac{dv_k}{dt_k} = -\frac{1}{a}\frac{\partial G_k}{\partial\theta} - f u_k + (N_v)_k \tag{2.27}$$

where the non-linear terms are defined as

$$(N_u)_k = -\frac{1}{acos\theta}\frac{\partial(\Phi - G)_k}{\partial\lambda} - \frac{R_d(T_v)_k}{acos\theta}\frac{\partial lnp}{\partial\lambda} + \left(\frac{uv}{a}\right)_k tan\theta \tag{2.28}$$

$$(N_v)_k = -\frac{1}{a}\frac{\partial(\Phi - G)_k}{\partial\theta} - \frac{R_d(T_v)_k}{a}\frac{\partial lnp}{\partial\theta} - \left(\frac{u^2}{a}\right)_k tan\theta \tag{2.29}$$

Equations (2.26) and (2.27) are discretised in time as follows:

$$\left[u + \frac{\Delta t_+}{2}\left(\frac{1}{acos\theta}\frac{\partial G}{\partial\lambda} - f v\right)\right]_k^{n+1} = (A_u)_k, \tag{2.30}$$

$$\left[v + \frac{\Delta t_+}{2}\left(\frac{1}{a}\frac{\partial G}{\partial\theta} + f u\right)\right]_k^{n+1} = (A_v)_k, \tag{2.31}$$

where

$$(A_u)_k = \left[u^n - \left(\frac{\Delta t_-}{2}\right)\left(\frac{1}{acos\theta}\frac{\partial G^n}{\partial\lambda} - f v^n - N_u^{n+1/2}\right)\right]_{*3,k} + \left(\frac{\Delta t_+}{2}\right)(N_u)_k^{n+1/2} \tag{2.32}$$

$$(A_v)_k = \left[v^n - \left(\frac{\Delta t_-}{2}\right)\left(\frac{1}{a}\frac{\partial G^n}{\partial\theta} + f u^n - N_v^{n+1/2}\right)\right]_{*3,k} + \left(\frac{\Delta t_+}{2}\right)(N_v)_k^{n+1/2}. \tag{2.33}$$

Substituting v from (2.31) in (2.30), and u from (2.30) in (2.31), and replacing G using (2.21) yields

$$\left[u - \left(\frac{\Delta t_+}{2}\right)^2\left(\frac{b}{acos\theta}\frac{\partial}{\partial\lambda} + \frac{e}{a}\frac{\partial}{\partial\theta}\right)\mathbf{MD}\right]_k^{n+1} = (Z_u)_k \tag{2.34}$$

$$\left[v - \left(\frac{\Delta t_+}{2}\right)^2\left(\frac{b}{a}\frac{\partial}{\partial\theta} - \frac{e}{acos\theta}\frac{\partial}{\partial\lambda}\right)\mathbf{MD}\right]_k^{n+1} = (Z_v)_k, \tag{2.35}$$

where

$$Z_u = b\left[A_u - \left(\frac{\Delta t_+}{2}\right)\frac{1}{acos\theta}\frac{\partial H}{\partial \lambda}\right] + e\left[A_v - \left(\frac{\Delta t_+}{2}\right)\frac{1}{a}\frac{\partial H}{\partial \theta}\right] \tag{2.36}$$

$$Z_v = b\left[A_v - \left(\frac{\Delta t_+}{2}\right)\frac{1}{a}\frac{\partial H}{\partial \theta}\right] - e\left[A_u - \left(\frac{\Delta t_+}{2}\right)\frac{1}{acos\theta}\frac{\partial H}{\partial \lambda}\right] \tag{2.37}$$

$$\mathbf{M} = (\gamma\tau + R_d T^0 \nu), \tag{2.38}$$

and H is defined by (2.22); $b = 1/(1+F_+^2)$, $e = F_+ b$, and $F_+ = f\Delta t_+/2$.

2.4 The Specific Humidity and Cloud Water Equations

The equations of motion for the specific humidity, q, and the cloud water mixing ratio, m, are

$$\left(\frac{dq}{dt}\right)_k = (P_q + K_q)_k \tag{2.39}$$

$$\left(\frac{dm}{dt}\right)_k = (P_m + K_m)_k \tag{2.40}$$

where, as before, P and K represent the physics and the diffusion contributions. Their discretization will be discussed in sections below. The adiabatic terms are discretized as

$$q_k^{n+1} = q_{*3,k}^n \tag{2.41}$$

$$m_k^{n+1} = m_{*3,k}^n. \tag{2.42}$$

2.5 Evaluating the Terms Valid at time Level $n + 1/2$

The non-linear terms, N_p, N_T, N_u, and N_v are needed at time level $n + 1/2$. To compute them u, v, T, q, and p_s are extrapolated using one of the following formulae

$$\psi^{n+1/2} = (3\psi^n - \psi^{n-1})/2 \tag{2.43}$$

$$\psi^{n+1/2} = (15\psi^n - 10\psi^{n-1} + 3\psi^{n-2})/8. \tag{2.44}$$

The latter is more accurate, in principle, but requires the storing of an additional set of fields.

2.6 The Trajectory Calculation

The model is encoded on an Arakawa C-grid, meaning that u and v are staggered in the horizontal with respect to T. The fields q and m are evaluated at the 'T-points'. The vertical velocity $\dot{s}$ is carried at the 'T-points' as far as the horizontal discretization is concerned and at the half levels in the vertical. The first step in doing the trajectory calculations is to interpolate all three velocity components to the 'T-points'. Thus u and v are interpolated in the horizontal using a cubic scheme and $\dot{s}$ is interpolated in the vertical to the full levels using a cubic scheme also.

All velocities having been defined at the same grid point, the three dimensional displacements are computed by solving the following equations iteratively, see 1999:

$$\hat{\alpha}^{(m+1)} = 1.5\alpha[\lambda_i - \hat{\alpha}^{(m)}, \theta_j - \hat{\beta}^{(m)}, \eta_k - \hat{\gamma}^{(m)}, t] \\ -0.5\alpha[\lambda_i - 2\hat{\alpha}^{(m)}, \theta_j - 2\hat{\beta}^{(m)}, \eta_k - 2\hat{\gamma}^{(m)}, t - \Delta t]$$

$$\hat{\beta}^{(m+1)} = 1.5\beta[\lambda_i - \hat{\alpha}^{(m)}, \theta_j - \hat{\beta}^{(m)}, \eta_k - \hat{\gamma}^{(m)}, t] \\ -0.5\beta[\lambda_i - 2\hat{\alpha}^{(m)}, \theta_j - 2\hat{\beta}^{(m)}, \eta_k - 2\hat{\gamma}^{(m)}, t - \Delta t] \tag{2.45}$$

$$\hat{\gamma}^{(m+1)} = 1.5\gamma[\lambda_i - \hat{\alpha}^{(m)}, \theta_j - \hat{\beta}^{(m)}, \eta_k - \hat{\gamma}^{(m)}, t] \\ -0.5\gamma[\lambda_i - 2\hat{\alpha}^{(m)}, \theta_j - 2\hat{\beta}^{(m)}, \eta_k - 2\hat{\gamma}^{(m)}, t - \Delta t],$$

where

$$(\alpha, \beta, \gamma) = \left(\frac{\Delta t}{2}\right)\left[\frac{\bar{u}}{a cos\theta}, \frac{\bar{v}}{a}, p_s \bar{\dot{s}}\left(\frac{\Delta \eta}{\Delta p}\right)\right], \tag{2.46}$$

and the bar indicates that the velocities are valid at the 'T-points'. The first guess, $\alpha^{(0)}$, is taken as $\alpha(\lambda_i, \theta_j, \eta_k, t)$; similarly for $\beta^{(0)}$ and $\gamma^{(0)}$. The displacements are then in the appropriate form to compute the departure point quantities in Eqs. (2.11), (2.18), (2.41) and (2.42). For any field ψ the departure point value is given by

$$[\psi]_{*3,k} = \psi[\lambda_i - 2\hat{\alpha}^{(m+1)}, \theta_j - 2\hat{\beta}^{(m+1)}, \eta_k - 2\hat{\gamma}^{(m+1)}], \tag{2.47}$$

To find the departure point values of the quantities in Eqs. (2.32) and (2.33) the displacements are first interpolated in the horizontal to the u- and v-points respectively using a cubic scheme; then Eq. (2.47) is used to find the departure point values.

2.7 Solving the Implicit System of Equations

Let $\mathbf{E}$ be the matrix that diagonalises the matrix $\mathbf{M}$. Applying the horizontal gradient operator to Eqs. (2.34) and (2.35) and diagonalising $\mathbf{M}$ yields an elliptic equation which can be solved for D^{n+1} :

$$\left[1 - \left(\frac{\Delta t_+}{2}\right)^2 \lambda_k (\nabla^2)'\right](\mathbf{E^{-1}D})_k^{n+1} = \nabla_H.(\mathbf{E^{-1}Z})_k, \tag{2.48}$$

where $\mathbf{Z} = (Z_u, Z_v)$ and

$$(\nabla^2)'_H = \frac{1}{a cos\theta}\left[\frac{\partial}{\partial \lambda}\left(\frac{b}{a cos\theta}\frac{\partial}{\partial \lambda} + \frac{e}{a}\frac{\partial}{\partial \theta}\right) + \frac{\partial}{\partial \theta}\left(\frac{b}{a}\frac{\partial}{\partial \theta} - \frac{e}{a cos\theta}\frac{\partial}{\partial \lambda}\right)\right] \tag{2.49}$$

This is solved by the 'half implicit Coriolis scheme'. It works as follows. In the momentum equations the Coriolis term is written as $f = f_0 + f'$. The terms multiplying f_0 are then integrated implicitly. The terms multiplying f' are integrated explicitly in the same manner as N_u and N_v. This gives rise to a Helmholtz equation rather than the more complicated elliptic equation (2.48). This scheme is extremely efficient since it involves only one pass of the direct solver. In the method Eqs. (2.26) - (2.38) remain formally identical, except that f must be interpreted as f_0 and N_u and N_v must be changed as follows $N_u \rightarrow N_u + f'v$ and $N_v \rightarrow N_v - f'u$. Repeating the manipulations which gave rise to Eq. (2.48) gives rise to the Helmholtz equation

$$\left[1 - b_0\left(\frac{\Delta t_+}{2}\right)^2 \lambda_k \nabla_H^2\right](\mathbf{E^{-1}D})_k^{n+1} = \nabla_H.(\mathbf{E^{-1}Z}_0)_k, \tag{2.50}$$

where Z_0 is equal to Z but with f, N_u and N_v re-interpreted as described above and

$$\nabla_H^2 = \frac{1}{a cos\theta}\left[\frac{\partial}{\partial \lambda}\frac{1}{a cos\theta}\frac{\partial}{\partial \lambda} + \frac{\partial}{\partial \theta}\frac{cos\theta}{a}\frac{\partial}{\partial \theta}\right] \tag{2.51}$$

It should be mentioned that the fast solver employed for Eq. (2.51) uses Fourier transforms in the λ-direction. To accomodate this the divergence is put to zero on the boundary line.
Having determined D^{n+1}, u^{n+1} and v^{n+1} can be recovered from Eqs. (2.34) and (2.35); D^{n+1} from Eq. (2.18), and lnp_s^{n+1} from Eq. (2.11).

2.8 Dampers and Filters

2.8.1 'De-centering'

The de-centering scheme of 1992 has been included to damp gravity waves. Hence

$$\Delta t_{\pm} = (1 \pm \epsilon_g)\Delta t. \tag{2.52}$$

The value of $\epsilon_g = 0.1$ is used operationally.

2.8.2 Implicit Horizontal Diffusion

An implicit diffusion scheme is also included. In order to diffuse the field f we wish to solve the equation

$$\frac{\partial f}{\partial t} - K\nabla_H^2\nabla_H^2 f = 0 \tag{2.53}$$

Direct discretization of Eq. (2.53) leads to a very complicated matrix equation which may be expensive to solve. Since it is possible to invent a reasonably inexpensive method of solution, which is outlined below, the following diffusion equation is solved instead:

$$\frac{\partial f}{\partial t} + K\left(\frac{\partial^4 f}{\partial x^4} + \frac{\partial^4 f}{\partial y^4}\right) = 0, \tag{2.54}$$

where $acos\theta\Delta\lambda = \Delta x$ and $a\Delta\theta = \Delta y$. This is solved using a two-step split implicit approximation as follows:

$$\frac{f^* - f^n}{\Delta t} + K\frac{\partial^4 f^*}{\partial x^4} = 0 \tag{2.55}$$

$$\frac{f^{n+1} - f^*}{\Delta t} + K\frac{\partial^4 f^{n+1}}{\partial y^4} = 0, \tag{2.56}$$

where

$$\frac{\partial^4 f}{\partial x^4} = \frac{1}{\Delta x^4}\left(f_{i-2,j} - 4f_{i-1,j} + 6f_{i,j} - 4f_{i+1,j} + f_{i+2,j}\right) \tag{2.57}$$

$$\frac{\partial^4 f}{\partial y^4} = \frac{1}{\Delta y^4}\left(f_{i,j-2} - 4f_{i,j-1} + 6f_{i,j} - 4f_{i,j+1} + f_{i,j+2}\right) \tag{2.58}$$

Each of the Eqs. (2.57) and (2.58) involves a penta-diagonal matrix. They are solved using a method described by 1988 which works as follows. The equation

$$\mathbf{C}.\mathbf{\Psi} = \mathbf{g}, \tag{2.59}$$

where $\mathbf{C}$ is a penta-diagonal matrix and $\mathbf{\Psi}$ and $\mathbf{g}$ are column matrices, can be re-written as follows

$$\mathbf{W}.\mathbf{S}.\mathbf{\Psi} = \mathbf{g}, \tag{2.60}$$

where $\mathbf{W}$ and $\mathbf{S}$ are each nearly-diagonal triangular matrices containing at most 4 non- zero elements per column. This fact enables us to find a solution by a process akin to Gaussian elimination. For details, see the Appendix of 1988.

At the first line in from the boundary second order diffusion is applied.

This diffusion is applied, with equal coefficients, to all of the fields except liquid water. The default value is $K_2 = 3.5 \times 10^{14}$ for $\Delta x_2 = 0.5^\circ$ and $\Delta t_2 = 300$. If one is running on a different grid the coefficient is automatically scaled such that on the new grid the e-folding time for the two-grid wave will be unaltered.

The fields are diffused after the dynamics and physics updates have been completed.

2.8.3 Filtering the Non-linear Terms

Since the fields being evaluated at time level $n + 1/2$ are potential sources of instability, they are filtered as follows. Any field, call it ψ, required at time level $n + 1/2$ is computed as

$$\psi^{n+1/2} = (3\psi^n - \psi_f^{n-1})/2, \tag{2.61}$$

where

$$\psi_f^n = \psi^n + \epsilon_N[\psi^{n+1} - 2\psi^n + \psi_f^{n-1}] \tag{2.62}$$

Only the non-linear terms and centering to find the departure point are affected by this filter; the linear terms remain untouched. Thus, it should cause minimal decrease in accuracy. The value $\epsilon_N = 0.1$ is the default. This filter is the final process applied to the fields.

2.9 Boundaries

The lateral boundaries are over-specified, all variables being externally prescribed by a 'host model'. As well, a buffer zone of points surrounds the integration area. Its fields are also updated via the host model. Thus it is still possible to interpolate the fields to a departure point which is outside the area boundary. Ultimately, if the departure point lies outside at the outer line of the buffer zone the trajectory is truncated.

Because of the over-specification of the fields on the boundary line, and also because the host and guest models evolve independently, there may arise sharp differences between these imposed boundary fields and those at the adjacent points in the limited area ('guest') model. To minimise this effect the guest model fields are relaxed toward the host model fields in a 'relaxation zone' close to the boundary. See 1976.

$$\phi_i = (1 - \alpha_i)\phi_i^G + \alpha_i\phi_i^H, \tag{2.63}$$

where G and H refer to the 'guest' and 'host' model fields, $0 \leq \alpha \leq 1$, and $\alpha_1 = 1$ on the boundary, and $\alpha_i = 0$ for all lines beyond the relaxation zone, which is n lines in width ($n = 10$ is the default in the reference system). The default value of α is

$$\alpha_i = \frac{1}{2}\left\{1 + cos[\pi(i-1)/n]\right\} \tag{2.64}$$

The boundary relaxation is performed after the horizontal diffusion.

At the upper boundary, the condition $\dot{\eta} = 0$ is imposed. This has no physical basis and is a source of false reflections which can cause standing waves to form in the model atmosphere. To handle this problem the coefficient of implicit horizontal diffusion can be increased near the top of the atmosphere.

Lastly, if the departure point is outside the uppermost or lowermost boundary levels, the trajectory is truncated at that level.

2.10 Under development: nonhydrostatic HIRLAM

The nonhydrostatic model assumes that the acoustic mode is completely filtered. In physical respect this means the infinite sound speed approximation. However, in the model equations, filtration of internal acoustic mode is achieved with the help of the 3D non-divergence of the motion (i.e. incompressibility of the medium) in pressure coordinates, which in continuous pressure-coordinate presentation is expressed via the continuity equation (common with the hydrostatic model)

$$\nabla_p \cdot v + \frac{\partial \omega}{\partial p} = 0 \ .$$

The external acoustic mode is filtered, making use of the surface pressure adjustment. This means, the surface pressure is approximated in dynamics with the given model function p_0 (which in practice is taken as the "boundary field" from a larger model), whereas the divergence of the vertically integrated horizontal wind satisfies the mass balance condition

$$\nabla \cdot \int_0^{p_0} \mathbf{v} dp = -\frac{\partial p_0}{\partial t} \ .$$

In the nonhydrostatic atmosphere, the geopotential Φ consists of the hydrostatic main component and nonhydrostatic contribution. In the acoustically filtered case, it is advantageous to present Φ as a sum of the thermic geopotential φ and baric geopotential ϕ:

$$\Phi = \varphi + \phi \ ,$$

where, in the continuous pressure-coordinate presentation,

$$\varphi = gh(x, y) + \int_p^{p_0} \frac{RT}{p'} dp' \ ,$$

and the baric geopotential, which is diagnosed from the elliptic geopotential equation (which will be presented below), includes, in addition to the nonhydrostatic component, also the barotropic hydrostatic geopotential fluctuation due to the vertically integrated mass fluctuation. The actual surface pressure fluctuation is diagnosed from relationship

$$p'_s = p_0 \left(\frac{\phi}{RT} \right)_{p_0} \ .$$

In the following, the non-hydrostatic, numerical, semi-implicit, Eulerian model is presented, which makes use of the above-described acoustical adjustment assumptions. The scheme includes explicit time stepping option as a special case. The nonhydrostatic modification concerns the adiabatic part of the dynamics solely, maintaining the pre-, and post-processing, and the physical parameterization packages unchanged. Presented is a compressed overview of the implemented numerical algorithm. Extended treatment of the model foundations can be found in papers (Rõõm, 2001; Männik and Rõõm, 2001; Rõõm and Männik, 2002).

2.10.1 Discretization

The discretization takes advantage of the hydrostatic HIRLAM framework to full extent and follows in headlines the presentation of the previous HIRLAM manual (Källén, 1996). In horizontal, the classical staggered (Arakawa C) grid is used. The surfaces $i = 1/2,\ Nlon + 1/2$, and $j = 1/2,\ Nlat + 1/2$ are lateral boundaries, $k = 1/2$ corresponds to a level $\eta = 0$ (outer space) and $k = Nlev + 1/2$ corresponds to $\eta = 1$ (model surface $p = p_0$). It is convenient to consider each small cube with the center at $\{i, j, k\}$, as an elementary pseudo-particle. The facets of this particle are centered at $\{i \pm 1/2, j, k\}$, $\{i, j \pm 1/2, k\}$, and $\{i, j, k \pm 1/2\}$. The discrete scalar fields: temperature T_{ijk}, specific humidity q_{ijk}, and baric geopotential ϕ_{ijk} are located in the center of the particle, whereas the components of vectors u, v, ω (as well as $m\dot{\eta}$ and the hydrostatic vertical velocity w^s), determining in- and outflows on the particle boundaries, are located in the centers of facets, and have, consequently, discretized presentations $u_{i+1/2jk}$, $v_{ij+1/2k}$, $\omega_{ijk+1/2}$, $(m\dot{\eta})_{ijk+1/2}$, and $w_{ijk+1/2}$. The discrete pressure $p_{ijk+1/2}$ is located in the centers of the horizontal facets; consequently, the pressure difference $(\Delta_\eta p)_{ijk} = p_{ijk+1/2} - p_{ijk-1/2}$ belongs to scalars.

Averaging operators are

$$\overline{a_l}^{\xi} = \frac{a_{l-1/2} + a_{l+1/2}}{2} , \quad \overline{a}^{\xi}_{l+1/2} = \frac{a_l + a_{i+1}}{2} ,$$

for all directions $\xi = x$, y, or η. The horizontal differencing operators are

$$(\delta_x a)_i = \frac{a_{i+1/2} - a_{i-1/2}}{\Delta x} , \quad (\delta_x a)_{i+1/2} = \frac{a_{i+1} - a_i}{\Delta x} ,$$

$$(\delta_y b)_j = \frac{b_{j+1/2} - b_{j-1/2}}{\Delta y} , \quad (\delta_y b)_{j+1/2} = \frac{b_{j+1} - b_j}{\Delta y} .$$

The consequent vertical differencing operators are

$$(\Delta_\eta a)_{k+1/2} = a_{k+1} - a_k , \quad (\Delta_\eta b)_k = b_{k+1/2} - b_{k-1/2} ,$$

$$(\delta_p a)_{ijk+1/2} = \left(\frac{\Delta_\eta a}{\overline{\Delta_\eta p}^{\eta}} \right)_{ijk+1/2} , \qquad (\delta_p^+ b)_{ijk} = \left(\frac{\Delta_\eta b}{\Delta_\eta p} \right)_{ijk} .$$

For time–related differential operations, notation is

$$\frac{\partial a}{\partial t} = \frac{a^{t+\Delta t} - a^{t-\Delta t}}{2\Delta t}, \quad \Delta_{tt} a = \frac{1}{2} \left(a^{t+\Delta t} + a^{t-\Delta t} \right) - a^t , \qquad a^t \equiv a(t) .$$

Horizontal averaging over the kth η-level is denoted as

$$\langle a \rangle_k = \frac{1}{Nlon Nlat} \sum_{ij} a_{ijk} .$$

Also, the special notation m is introduced for the hybrid-coordinate density

$$m = \frac{\partial p}{\partial \eta} .$$

2.10.2 Diagnostics

The pressure is

$$p_{ijk+1/2} = A_{k+1/2} + B_{k+1/2} p_{0ij} \ , \tag{2.65}$$

where $p_{0ij} = p_0(\lambda_i, \theta_j, t)$ is the reference surface pressure, and

$$A_{k+1/2} = A(\eta_{k+1/2}) \ , \quad B_{k+1/2} = B(\eta_{k+1/2}) \tag{2.66}$$

$$A_{1/2} = A_{Nlev+1/2} = 0 \ , \quad B_{1/2} = 0 \ , \quad B_{Nlev+1/2} = 1 \ . \tag{2.67}$$

The model employs spherical geometry with rotated spherical coordinates λ, θ. However, in differential operators coordinates $x = a\lambda$, $y = a\theta$ are used, to which the distances $X = h_x x = \cos\theta\, x$ along a parallel, and $Y = h_y y = 1 \cdot y$ in meridional direction, correspond. Metric coefficients h_x, h_y are initially introduced for the use in optional map projection system. Their use in the particular case of spheric geometry is justified by the most symmetric appearance of difference operators. The coordinate differences are $\Delta x = a\Delta\lambda$, $\Delta y = a\Delta\theta$, $\Delta\eta$, the consequent physical differences in horizontal directions are $\Delta X_{ij} = h_{xij}\Delta x$, $\Delta Y_{ij} = h_{yij}\Delta y$, where $h_{xij} = \cos\theta_j$, $h_{yij} = 1$. In this notation, the horizontal divergence of a vector $\mathbf{a}$ is

$$(\nabla \cdot \mathbf{a})_{ijk} = \frac{1}{(h_x h_y)_{ij}} \left[\delta_x (\overline{h_y}^x a_x) + \delta_y (\overline{h_x}^y a_y) \right]_{ijk} \ , \tag{2.68}$$

whereas the horizontal gradient of a scalar b has components

$$[(\nabla b)_x]_{i+1/2jk} = \frac{1}{h_{xi+1/2j}} (\delta_x b)_{i+1/2jk} = \frac{b_{i+1jk} - b_{ijk}}{h_{xi+1/2j}\Delta x} \ , \tag{2.69}$$

$$[(\nabla b)_y]_{ij+1/2k} = \frac{1}{h_{yij+1/2}} (\delta_y b)_{ij+1/2k} = \frac{b_{ij+1k} - b_{ijk}}{h_{yij+1/2}\Delta y} \ . \tag{2.70}$$

The auxiliary vector $\mathbf{V}=(\Delta_\eta p)\mathbf{v}$ is:

$$U_{i+1/2jk} = (\overline{\Delta_\eta p}^x u)_{i+1/2jk} \ , \quad V_{ij+1/2k} = \left(\overline{\Delta_\eta p}^y u \right)_{ij+1/2k} \ . \tag{2.71}$$

The formula for the η-velocity is

$$(m\dot{\eta})_{ijk+1/2} = \sum_{k'=k+1}^{Nlev} (\nabla \cdot \mathbf{V})_{ijk'} + (1 - B_{k+1/2}) \frac{\partial p_{0ij}}{\partial t} \ , \tag{2.72}$$

from which the recurrent formula follows

$$(m\dot{\eta})_{ijk-1/2} = (m\dot{\eta})_{ijk+1/2} + (\nabla \cdot \mathbf{V})_{ijk} + \Delta B_k \frac{\partial p_{0ij}}{\partial t} \ , \tag{2.73}$$

$$(m\dot{\eta})_{ijNlev+1/2} = 0 \ . \tag{2.74}$$

The vertically integrated mass balance condition is

$$\sum_{k=1}^{Nlev} (\nabla \cdot \mathbf{V})_{ijk} = -\frac{\partial p_{0ij}}{\partial t} \ . \tag{2.75}$$

The formula for omega-velocity is

$$\omega_{ijk+1/2} = (m\dot{\eta})_{ijk+1/2} + (\overline{\mathbf{v}}^{\eta} \cdot \nabla p)_{ijk+1/2} + \frac{\partial p_{ijk+1/2}}{\partial t} \ , \tag{2.76}$$

where

$$(\overline{\mathbf{v}}^{\eta} \cdot \nabla p)_{ijk+1/2} \equiv \frac{1}{(h_x h_y)_{ij}} \left(\overline{\overline{h_y}^{x} \overline{u}^{\eta} \delta_x p}^{x} + \overline{\overline{h_x}^{y} \overline{v}^{\eta} \delta_y p}^{y} \right)_{ijk+1/2} \ . \tag{2.77}$$

Nonhydrostatic model makes use of the hybrid-coordinate presentation of the pressure-coordinate divergence $\hat{\mathbf{G}}^{+}$ and gradient $\hat{\mathbf{G}}$, the η-coordinate presentations of which are

$$(\hat{\mathbf{G}}^{+} \cdot \mathbf{v})_{ijk} = \frac{1}{(h_x h_y)_{ij}} \left[\delta_x (\overline{h_y}^{x} u) - \frac{\overline{\overline{h_y}^{x} (\Delta_\eta u) \delta_x p}^{x\eta}}{\Delta_\eta p} + \delta_y (\overline{h_x}^{y} v) - \frac{\overline{\overline{h_x}^{y} (\Delta_\eta v) \delta_y p}^{y\eta}}{\Delta_\eta p} \right]_{ijk} \tag{2.78}$$

$$(\hat{G}_x \phi)_{i+1/2jk} = \frac{1}{\overline{h_x}^{x}} \left[\delta_x \phi - \frac{\overline{(\delta_x p) \Delta_\eta \overline{\phi}^{x}}^{\eta}}{\overline{\Delta_\eta p}^{x}} \right]_{i+1/2jk} \tag{2.79}$$

$$(\hat{G}_y \phi)_{ij+1/2k} = \frac{1}{\overline{h_y}^{y}} \left[\delta_y \phi - \frac{\overline{(\delta_y p) \Delta_\eta \overline{\phi}^{y}}^{\eta}}{\overline{\Delta_\eta p}^{y}} \right]_{ij+1/2k} \tag{2.80}$$

The continuity equation is

$$\left(\hat{\mathbf{G}}^{+} \cdot \mathbf{v} + \frac{\Delta_\eta \omega}{\Delta_\eta p} \right)_{ijk} = 0 \ . \tag{2.81}$$

The thermic geopotential is

$$\varphi_{ijk-1/2} = g h_{ij} + \sum_{k'=k}^{Nlev} (RT)_{ijk'} \alpha_{ijk'} \ , \tag{2.82}$$

where $h_{ij} = h(\lambda_i, \theta_j)$ is the surface height, and

$$\alpha_{ijk} = (\Delta_\eta \ln p)_{ijk} = \ln p_{ijk+1/2} - \ln p_{ijk-1/2} \ , \ k \neq 1, \quad \text{and} \quad \alpha_{ij1} = 2 \ln 2 \ . \tag{2.83}$$

The thermic geopotential can be evaluated also from the recurrence

$$\varphi_{ijNlev+1/2} = g h_{ij} \ ,$$

$$\overline{\varphi}^{\eta}_{ijk} = \varphi_{ijk+1/2} + \frac{1}{2} (RT)_{ijk} \alpha_{ijk} \ , \tag{2.84}$$

$$\varphi_{ijk-1/2} = \overline{\varphi}^{\eta}_{ijk} + \frac{1}{2} (RT)_{ijk} \alpha_{ijk} \ .$$

Note that coefficients α_{ijk} are different from those employed in the hydrostatic case.

2.10.3 Dynamic equations

Nonhydrostatic dynamics makes use of Eulerian presentation of the equations of motion and employs the semi-implicit Eulerian time stepping scheme as the main option. It includes the explicit Eulerian integration mode as a special case.

A. Evolutional equations

Equations of motion are

$$\frac{\partial u_{i+1/2jk}}{\partial t} = F_{ui+1/2jk} - [\widehat{G}_x(\phi + \Delta_{tt}\overline{\varphi}^{\eta})]_{i+1/2jk} \ , \tag{2.85}$$

$$\frac{\partial v_{ij+1/2k}}{\partial t} = F_{vij+1/2k} - [(\widehat{G}_y(\phi + \Delta_{tt}\overline{\varphi}^{\eta})]_{i+1/2jk} \ , \tag{2.86}$$

$$\frac{\partial T_{ijk}}{\partial t} = F_{Tijk} + (S\Delta_{tt}\overline{\omega}^{\eta})_{ijk} \ , \tag{2.87}$$

where F_u , F_v, F_T are the the explicit hydrostatic tendencies (common with the hydrostatic model), $\hat{\mathbf{G}}\phi$ is the baric tendency (includes the nonhydrostatic acceleration), and terms with $\Delta_{tt}\overline{\varphi}^{\eta}$, $\Delta_{tt}\overline{\omega}^{\eta}$ are responsible for the implicit adjustment.

B. Explicit hydrostatic tendencies

The explicit (given on the time level t) hydrostatic tendency components of velocity are:

$$\begin{aligned} F_{ui+1/2jk} &= -\left[\frac{1}{\overline{\overline{h_x}^{x}}}\left(-\overline{Z\overline{\overline{h_x}^{y}V}^{x}}^{y} + \delta_x E + \frac{\overline{\overline{h_x m\dot{\eta}}^{x}\Delta_\eta u}^{\eta}}{\overline{\Delta_\eta p}^{x}}\right)\right]_{i+1/2jk} \\ &\quad -(\widetilde{G}_x\overline{\varphi}^{\eta})_{i+1/2jk} + (P_u + K_u)_{i+1/2jk} \ , \end{aligned} \tag{2.88}$$

$$\begin{aligned} F_{vij+1/2k} &= -\left[\frac{1}{\overline{\overline{h_y}^{y}}}\left(\overline{Z\overline{\overline{h_y}^{x}U}^{y}}^{x} + \delta_y E + \frac{\overline{\overline{h_y m\dot{\eta}}^{y}\Delta_\eta v}^{\eta}}{\overline{\Delta_\eta p}^{y}}\right)\right]_{ij+1/2k} \\ &\quad -(\widetilde{G}_y\overline{\varphi}^{\eta})_{ij+1/2k} + (P_v + K_v)_{ij+1/2k} \ , \end{aligned} \tag{2.89}$$

where

$$Z_{i+1/2j+1/2k} = \frac{\left[f\overline{h_x h_y}^{xy} + \delta_x(\overline{h_x}^{y}v) - \delta_y(\overline{h_y}^{x}u)\right]_{i+1/2j+1/2k}}{(\overline{h_x h_y \Delta_\eta p}^{xy})_{i+1/2j+1/2k}} \tag{2.90}$$

$$E_{ijk} = \frac{1}{2}\left(\frac{1}{h_y}\overline{\overline{h_y}^{x}u^2}^{x} + \frac{1}{h_x}\overline{\overline{h_x}^{y}v^2}^{y}\right)_{ijk} \ , \tag{2.91}$$

and

$$(\widetilde{G}_x\overline{\varphi}^{\eta})_{i+1/2jk} = \frac{1}{\overline{h_x}^{x}}\left[\delta_x\overline{\varphi}^{\eta} - \frac{\delta_x\overline{p}^{\eta}\Delta_\eta\overline{\varphi}^{x}}{\overline{\Delta_\eta p}^{x}}\right]_{i+1/2jk} \ . \tag{2.92}$$

$$(\widetilde{G}_x\overline{\varphi}^{\eta})_{ij+1/2k} = \frac{1}{\overline{\overline{h_y}^{y}}}\left[\delta_y\overline{\varphi}^{\eta} - \frac{\delta_y\overline{p}^{\eta}\Delta_{\eta}\overline{\varphi}^{y}}{\overline{\Delta_{\eta}p}^{y}}\right]_{ij+1/2k} . \tag{2.93}$$

The explicit temperature tendency is

$$F_{Tijk} = -(\hat{a}T)_{ijk} + \left(\frac{\kappa T\alpha}{\Delta_{\eta}p}\Omega\right)_{ijk} + (P_T + K_T)_{ijk} , \tag{2.94}$$

where

$$(\hat{a}T)_{ijk} = \frac{[\overline{h_yU\delta_xT}^{x} + \overline{h_xV\delta_yT}^{y} + h_xh_y\overline{(m\dot{\eta})\Delta_{\eta}T}^{\eta}]_{ijk}}{(h_xh_y\Delta_{\eta}p)_{ijk}} , \tag{2.95}$$

and

$$\Omega_{ijk} = \overline{\omega}^{\eta}_{ijk} - \left[\frac{\overline{\overline{h_y}^{x}\Delta_{\eta}\left(\delta_xp\Delta_{\eta}u\right)}^{x} + \overline{\overline{h_x}^{y}\Delta_{\eta}\left(\delta_yp\Delta_{\eta}v\right)}^{y}}{4h_yh_x}\right]_{ijk} . \tag{2.96}$$

Terms P_{xijk}, K_{xijk} are tendencies from physical parametrization and horizontal diffusion, respectively.

C. Implicit adjustment terms

The implicit adjustment terms $\Delta_{tt}\overline{\varphi}^{\eta}$ and $\Delta_{tt}\overline{\omega}^{\eta}$ make use of the time step Δt and non-dimensional parameter

$$\nu = \Delta tN^* , \tag{2.97}$$

where constant N^* is the Brunt-Väisälä frequency of the reference state. It is reasonable to choose N^* to exceed the maximum of the actual Brunt-Väisälä frequency in the modeled airmass. The implicit adjustment scheme is not too sensitive to the small variation of N^*. In application, the value $N^* = 0.025\ s^{-1}$ is applied.

The stratification term S in the temperature equation (2.87) is

$$S_{ijk} = \left(\frac{\kappa\alpha T - \overline{\Delta_{\eta}T}^{\eta}}{\Delta_{\eta}p}\right)_{ijk} . \tag{2.98}$$

The implicit adjustment terms $\Delta_{tt}\overline{\varphi}^{\eta}$ and $\Delta_{tt}\overline{\omega}^{\eta}$ are expressed via baric geopotential ϕ, fields $T^{t-\Delta t}$, T^t, $\omega^{t-\Delta t}$, ω^t, and their explicit tendencies, F_T and F_ω, as follows

$$\Delta_{tt}\overline{\varphi}^{\eta}_{ijk} = \nu^2(\phi_{ijk} - \phi_{ijNlev}) + Q^0_{ijk}, \tag{2.99}$$

$$(\overline{\Delta_{tt}\omega}^{\eta})_{ijk} = -\Delta t\left[\overline{\left(\frac{p}{\overline{H}^{\eta}}\right)^2\delta_p\phi}^{\eta}\right]_{ijk} + (\overline{\mu_\omega}^{\eta})_{ijk} , \tag{2.100}$$

$$[Q^0]_{ijk} = \sum_{k'=k+1}^{Nlev}\left[\alpha R\left(\mu_T + \Delta tS\overline{\mu_\omega}^{\eta}\right)\right]_{ijk'} + \frac{1}{2}\left[\alpha R\left(\mu_T + \Delta tS\overline{\mu_\omega}^{\eta}\right)\right]_{ijk} , \tag{2.101}$$

$$\mu_\omega = \Delta tF_\omega - \omega^t + \omega^{t-\Delta t} , \quad \mu_T = \Delta tF_T - T^t + T^{t-\Delta t} , \tag{2.102}$$

where $H = RT/g$ represents the scale height. Note, that (2.100) is the nonhydrostatic evolutional equation for omega-velocity, written here in a different way as a relationship for the diagnosis of implicit adjustment term $\Delta_{tt}\overline{\omega}^{\eta}$ (the omega–velocity itself is a diagnostic field like in the nonhydrostatic case, diagnosed from (2.76)). The explicit tendency term for ω is

$$F_{\omega ijk+1/2} = \left(\frac{p}{\overline{\overline{H}^{\eta}}}\hat{a}w^s\right)_{ijk+1/2} + \left(\frac{\omega}{\overline{\overline{T}^{\eta}}}\overline{F_T}^{\eta}\right)_{ijk+1/2} , \tag{2.103}$$

$w^s_{ijk+1/2}$ is the hydrostatic vertical velocity,

$$w^s_{ijk+1/2} = -\left(\frac{\omega\overline{H}\eta}{p}\right)_{ijk+1/2} , \tag{2.104}$$

and $\hat{a}w^s$ presents material transport of w^s:

$$(\hat{a}w^s)_{ijk+1/2} = \frac{[\overline{h_y\overline{U}^{\eta}\delta_x w^s}^{x} + \overline{h_x\overline{V}^{\eta}\delta_y w^s}^{y} + h_x h_y\overline{m\dot{\eta}^{\eta}\Delta_{\eta}w^s}^{\eta}]_{ijk+1/2}}{(h_x h_y\overline{\Delta_{\eta}p}^{\eta})_{ijk+1/2}} . \tag{2.105}$$

D. Elliptic equation for baric geopotential

For nonhydrostatic updating of evolutional equations and for determination of the implicit adjustment terms, it is necessary to specify the baric geopotential ϕ.

The baric geopotential is diagnosed from elliptic equation

$$(1+\nu^2)(\hat{\mathbf{G}}^{+}\cdot\hat{\mathbf{G}}\phi)_{ijk} + (\hat{\mathbf{L}}\phi)_{ijk} = \mathcal{A}_{ijk} + (\hat{\mathbf{G}}^{+}\cdot\hat{\mathbf{G}})_{ijk}\rho_{ij} + \gamma_{ij}\delta_{kNlev} , \tag{2.106}$$

where two-dimensional fields γ and ρ have to be specified from conditions

$$\{\widehat{\mathcal{B}}[(1+\nu^2)\phi - \rho]\}_{ij} = b_{ij} - (\widehat{\mathcal{B}}Q^0)_{ij} , \tag{2.107}$$

$$\rho_{ij} = \nu^2\phi_{ijNlev} . \tag{2.108}$$

Together, (2.106) - (2.108) perform elliptic system for ϕ, γ, and ρ. In these equations, operator $\hat{\mathbf{L}}$ is

$$\hat{\mathbf{L}}\phi = \delta_p^{+}\left(\frac{p^2}{(\overline{\overline{H}^{\eta}})^2}\delta_p\phi\right) , \tag{2.109}$$

the source function A is

$$\mathcal{A} = (\hat{\mathbf{G}}^{+}\cdot\mathbf{F}_{\mathbf{v}}) + \delta_p^{+}F_{\omega} - \hat{\mathbf{G}}^{+}\cdot\hat{\mathbf{G}}Q^0 , \tag{2.110}$$

operator $\mathcal{B}$ is

$$(\widehat{\mathcal{B}}\phi)_{ij} = \sum_{k=1}^{Nlev}\left[\delta_x\left(\overline{h_y}^{x}\overline{\Delta_{\eta}p}^{x}\hat{G}_x\phi\right) + \delta_y\left(\overline{h_x}^{y}\overline{\Delta_{\eta}p}^{y}\hat{G}_y\phi\right)\right]_{ijk} , \tag{2.111}$$

and the source term b in (2.107) is

$$b_{ij} = \sum_{k=1}^{Nlev}\left[\delta_x\left(\overline{h_y}^{x}\overline{\Delta_{\eta}p}^{x}\hat{F}_u\right) + \delta_y\left(\overline{h_x}^{y}\overline{\Delta_{\eta}p}^{y}\hat{F}_v\right)\right]_{ijk}$$
$$+(h_x h_y)_{ij}\left[\frac{\partial^2\overline{p}_0}{\partial t^2} + \nabla\cdot\left(\frac{\partial\overline{p}_0}{\partial t}\sum_{k=1}^{Nlev}(\Delta_{\eta}B)_k\mathbf{v}_k\right)\right]_{ij} . \tag{2.112}$$

The equation (2.107) is vertically integrated mass balance condition for the baric geopotential. The new field ρ_{ij} is introduced for solution convenience, it may be eliminates with the help of relationship (2.108).

In principal, the elliptic equation in the bounded domain requires boundary conditions also on the lateral walls. Such conditions can be easily handled, introducing the singular, surface-located sources on the walls. These singular sources would be analogous to the bottom source $\gamma_{ij}\delta_{kNlev}$ in (2.106), the aim of which is to satisfy integral condition (2.107). Details of boundary source technique in lateral condition treatment is in detail described in paper by 2001. However, in the case of the Davies' boundary relaxation zone, which is (like in the nonhydrostatic model) the choice of nonhydrostatic scheme, the valid boundary condition at lateral walls is the homogeneous Neumann condition (*ibid*). This condition is satisfied automatically with the application of the cosine-Fourier basis in horizontal direction, which also is applied in the present nonhydrostatic model (details are presented further). Thus, we have no complication due to lateral boundaries, like there will be no missing components of solution due to lateral boundaries, so far the Davies' boundary relaxation zone is employed, and the cosine-Fourier basis is used at solution of elliptic equation for ϕ.

2.10.4 Solver for the elliptic system

The solution algorithm for equations (2.106) - (2.108) is performed in three steps. First, the original system is reduced to an iterative algorithm. Secondly, at each iteration, the main part of the elliptic operator is inverted explicitly for optional ρ, γ, in the orthogonal basis. At last, ρ and γ are specified.

A. Iterative elliptic system

Operators in (2.106) - (2.107) are presented as the sums of horizontally homogeneous main parts and non-homogeneous perturbations

$$\hat{\mathbf{G}}^{+}\cdot\hat{\mathbf{G}} = (\mathcal{L}_x+\mathcal{L}_y) + [\hat{\mathbf{G}}^{+}\cdot\hat{\mathbf{G}}]', \qquad \mathbf{L} = \mathcal{L}_\eta + \mathbf{L}' \ , \qquad \widehat{B} = \widehat{\mathcal{B}}^0 + \widehat{\mathcal{B}}' \ , \tag{2.113}$$

where the main parts $\mathcal{L}_x$, $\mathcal{L}_y$, $\mathcal{L}_\eta$, and $\widehat{\mathcal{B}}^0$ are the horizontal averages of the consequent full operators:

$$(\mathcal{L}_x\phi) = \left(\frac{1}{\langle h_x\rangle}\delta_x\phi\right)^2, \ (\mathcal{L}_x\phi) = \left(\frac{1}{\langle h_y\rangle}\delta_y\phi\right)^2, \ (\mathcal{L}_\eta\phi) = \left[\frac{1}{\langle\Delta_\eta p\rangle}\Delta_\eta\left(\frac{\langle p\rangle^2}{\langle\overline{H}^\eta\rangle^2}\frac{\Delta_\eta\phi}{\langle\overline{\Delta_\eta p}^\eta\rangle}\right)\right], \tag{2.114}$$

$$(\hat{\mathcal{B}}^0\phi) = \langle h_x\rangle\langle h_y\rangle\left[(\mathcal{L}_x+\mathcal{L}_y)\sum_{k=1}^{Nlev}\langle\Delta_\eta p\rangle_k\phi_k\right] \ . \tag{2.115}$$

Operation $<\cdot>_k$ means horizontal averaging over the kth η-level. Perturbation operators, denoted with prime, are computed as differences of exact and mean operators.

Using these operator separations, system (2.106) - (2.108) is replaced with the iterative set of equations (horizontal indexes i, j are omitted in following)

$$(1+\nu^2)(\mathcal{L}_x+\mathcal{L}_y)\phi_k^{(l)} + (\mathcal{L}_\eta\phi^{(l)})_k = \mathcal{A}_k^{(l)} + (\mathcal{L}_x+\mathcal{L}_y)\rho^{(l)} + \gamma\delta_{kNlev} \tag{2.116}$$

$$\widehat{\mathcal{B}}^0[(1+\nu^2)\phi^{(l)} - \rho^{(l)}] = b^{(l)} \ , \tag{2.117}$$

$$\rho^{(l)} = \nu^2\phi_{Nlev}^{(l)} \ , \tag{2.118}$$

where $\{\phi^{(l)},\ \rho^{(l)},\ \gamma^{(l)}\}$ represent the lth iteration to the exact solution $\{\phi,\ \rho,\ \gamma\}$ for iterated sources

$$\mathcal{A}_k^{(l)} = \mathcal{A}_k - \{[(1+\nu^2)(\hat{\mathbf{G}}^+ \cdot \hat{\mathbf{G}})' + \hat{\mathbf{L}}']\phi^{(l-1)}\}_k + [\hat{\mathbf{G}}^+ \cdot \hat{\mathbf{G}}]'_k \rho^{(l-1)} \ , \tag{2.119}$$

$$b^{(l)} = b - \hat{\mathcal{B}}Q^0 - (1+\nu^2)\hat{\mathcal{B}}'\phi^{(l-1)} + \hat{\mathcal{B}}'\rho^{(l-1)} \ . \tag{2.120}$$

B. Solver for the main part of elliptic operator

The iterative set of equations (2.116)–(2.118) is solved with optional γ, ρ, using the three-dimensional orthogonal basis $X \otimes Y \otimes E$, where the one-dimensional sub-bases in x, y and η directions are

$$X = \{X_q, q = 1, .., Nlon\} = \{\{X_{iq}, i = 1..., Nlon\}, q = 1, ..., Nlon\}$$

$$Y = \{Y_r, r = 1, .., Nlat\} Y_r = \{\{Y_{jr}, j = 1..., Nlat\}, r = 1, ..., Nlat\}$$

$$E = \{E_s, s = 1, .., Nlev\} E_s = \{\{E_{ks}, k = 1..., Nlev\}, s = 1, ..., Nlev\}$$

They are chosen as the eigenvectors of the one-dimensional Laplacians:

$$(\mathcal{L}_x X_q)_i = -\lambda_q^x X_{iq} \ , \quad (\mathcal{L}_y Y_r)_j = -\lambda_r^y Y_{jr} \ , \quad (\mathcal{L}_\eta E_s)_k = -\lambda_s^\eta E_{ks} \ , \tag{2.121}$$

where λ_q^x, λ_r^y, λ_s^η are the corresponding eigenvalues.
For X and Y the discrete normalized cosine bases are employed, which yield eigenvalues

$$\lambda_q^x = \frac{4}{\langle h_x \rangle^2 \Delta x^2} \sin^2\left(\frac{\pi}{2}\frac{q-1}{Nlon-1}\right) \quad \lambda_r^y = \frac{4}{\langle h_y \rangle^2 \Delta y^2} \sin^2\left(\frac{\pi}{2}\frac{r-1}{Nlat-1}\right),$$

whereas the basis E, and eigenvalues λ^η (they depend on the vertical temperature distribution in the atmosphere) are specified, numerically solving the vertical eigenvalue problem.
Using notation

$$\tilde{\phi}_{qrs} = \sum_{ijk} X_{qi} Y_{rj} E^{-1}_{sk} \phi^{(l)}_{ijk} \ , \quad \tilde{\mathcal{A}}_{qrs} = \sum_{ijk} X_{qi} Y_{rj} E^{-1}_{sk} \mathcal{A}^{(l)}_{ijk} \ , \tag{2.122}$$

$$\tilde{\gamma}_{qr} = \sum_{ij} X_{qi} Y_{rj} \gamma^{(l)}_{ij} \ , \quad \hat{\rho}_{qr} = \sum_{ij} X_{qi} Y_{rj} \rho^{(l)}_{ij} \ , \quad \tilde{b}_{qr} = \sum_{ij} X_{qi} Y_{rj} b^{(l)}_{ij} \ , \tag{2.123}$$

$$c_s = \sum_k \langle \Delta p \rangle_k E_{ks} \ , \quad d_s = \sum_k E^{-1}_{sk} \ , \tag{2.124}$$

the solution of (2.116) in the basis is

$$\tilde{\phi}_{qrs} = \frac{(\Lambda_q^x + \Lambda_r^y) d_s \tilde{\rho}_{qr} - \tilde{\gamma}_{qr} E^{-1}_{sNlev} - \tilde{\mathcal{A}}_{qrs}}{(1+\nu^2)(\Lambda_q^x + \Lambda_r^y) + \Lambda_s^\eta} \ . \tag{2.125}$$

C. Determination of coefficients γ and ρ

Solution (2.125) still includes the unspecified coefficients γ and ρ. These coefficients can be determined, substituting (2.125) into equations (2.117) and (2.118). This results in the two-dimensional set of linear algebraic equations for each pair of coefficients $\tilde{\gamma}_{qr}$, $\tilde{\rho}_{qr}$ (repetitive indexes q, r are omitted in following)

$$m_{11}\tilde{\rho} + m_{12}\tilde{\gamma} = n_1 \ , \tag{2.126}$$

$$m_{21}\tilde{\rho} + m_{22}\tilde{\gamma} = n_2 \ , \tag{2.127}$$

where

$$m_{11} = (\lambda^x + \lambda^y)\sum_s \frac{c_s d_s}{(1+\nu^2)(\lambda^x+\lambda^y)+\lambda^\eta} - \frac{\nu^2\langle p\rangle_{Nlev+1/2}}{(1+\nu^2)} \ , \tag{2.128}$$

$$m_{12} = -\sum_s \frac{c_s E^{-1}_{s,Nlev}}{(1+\nu^2)(\lambda^x+\lambda^y)+\lambda^\eta} \ , \tag{2.129}$$

$$m_{21} = 1 - (\lambda^x + \lambda^y)\sum_s \frac{E_{Nlev,s} d_s}{(1+\nu^2)(\lambda^x+\lambda^y)+\lambda^\eta} \ , \tag{2.130}$$

$$m_{22} = \sum_s \frac{E_{Nlev,s}E^{-1}_{s,Nlev}}{(1+\nu^2)(\lambda^x+\lambda^y)+\lambda^\eta} \ , \tag{2.131}$$

$$n_1 = \sum_s \frac{c_s \tilde{\mathcal{A}}_s}{(1+\nu^2)(\lambda^x+\lambda^y)+\lambda^\eta} - \frac{\tilde{b}}{(1+\nu^2)\langle h_x\rangle\langle h_y\rangle(\lambda^x+\lambda^y)} \ , \tag{2.132}$$

$$n_2 = -\sum_s \frac{E_{Nlev,s}\tilde{\mathcal{A}}_s}{(1+\nu^2)(\lambda^x+\lambda^y)+\lambda^\eta} \ . \tag{2.133}$$

After the system (2.126) - (2.127) is solved, $\tilde{\gamma}$ and $\tilde{\rho}$ are substituted into (2.125) and the resulting iterative solution is inverted back into the coordinate presentation:

$$\phi^{(l)}_{ijk} = \sum_{qrs} X_{iq}Y_{jr}E_{ks}\tilde{\phi}_{qrs} \ . \tag{2.134}$$

The iterative process is stopped at l, for which:

$$\langle|\phi^{(l)} - \phi^{(l-1)}|\rangle < \varepsilon\langle|\phi^{(l)}|\rangle \ , \tag{2.135}$$

where ε depends on the required precision. The typical value in application is $\varepsilon \sim 10^{-3} - 10^{-4}$, and the typical number of iterations is 2 – 3.

2.10.5 Surface pressure diagnostics

In the acoustically relaxed model, the actual surface pressure is a diagnostic field. It can be computed from the baric geopotential distribution as

$$p_s = p_0\left[1 + \left(\frac{\phi}{RT - p\Delta_\eta\phi/\Delta_\eta p}\right)_{p_0}\right] \ .$$

Usually, the second term in the denominator is a small quantity in comparison with the first term, and the formula can be simplified to

$$p_s = p_0\left[1 + \left(\frac{\phi}{RT}\right)_{p_0}\right] \ . \tag{2.136}$$

2.10.6 Case of the explicit time-stepping

The explicit-Eulerian integration scheme follows automatically in the limit $\Delta t \rightarrow 0$. For finite Δt, the explicit scheme can be obtained, nullifying $\Delta_{tt}\overline{\varphi}^{\eta}$, $\Delta_{tt}\overline{\omega}^{\eta}$ in equations (2.85), (2.86), and ν^2, ρ, Q^0 in (2.106) - (2.108), and(2.110). Obtained in this way model will coincide with the explicit scheme, described in papers (Rõõm, 2001; Männik and Rõõm, 2001). That means, the numerical algorithm for explicit and implicit scheme is, in essence, the same. Especially important is, that the elliptic system is the same and there is no need for double (i.e., separate for explicit time stepping and for implicit adjustment) integration of it in the in semi-implicit version. At 10 to 5 km horizontal resolutions, this means approximately double efficiency of semi-implicit scheme in comparison with the explicit case due to larger time-step. For instance, at the maximum wind speed 50 m/s, and 5.5 km (0.05 deg) horizontal resolution, the maximum time-steps of the explicit and semi-implicit models are 40 and 70 s, consequently. Toward the shorter scales, the time-step difference diminishes, and becomes negligible at approximately 0.5 km resolution. That means, semi-implicit scheme has there, in comparison with the explicit scheme, no significant advantage in computational speed. However, the preference should be at all instances given to the semi-implicit scheme. It has the same accuracy at equal time step as the explicit model does. At the same time, in the equal conditions, it is more stable (less 'capricious' numerically). For instance, in the 'normal' mid-latitude atmospheric conditions, as well as in the model atmosphere with relatively high stability ($N \geq 0.01\ s^{-1}$), it does not require any sponge layer near the top of the atmosphere, whereas the explicit scheme would without sponge layer fail due to strong, spurious reflection of buoyancy waves.

Chapter 3

Physics

3.1 Physical parameterization at high resolution

The role of physical parameterizations in an atmospheric model is to describe diabatic effects. In addition, the physics are necessary in order to obtain predictions of weather parameters such as rain and clouds. The physics comprise the processes of radiation and subgrid scale transports of dependent variables such as momentum, heat and moisture down to small scales associated with turbulence. In addition, the thermodynamics associated with latent heat release (e.g., condensation, evaporation, sublimation and precipitation) must be described. The boundary conditions at the ground need also to be taken into account.

Parameterization of subgrid scale phenomena depends on the resolution of the model, because resolution defines an ability to resolve a given phenomenon. A simple example is the parameterization of cloud cover. In large scale models a stratiform cloud cover is often parameterized as a function of relative humidity, starting to become nonzero at a threshold relative humidity much below 100 percent, in order to account for subgrid scale variability of moisture. As the model resolution increases it is obviously not realistic to retain the cloud cover threshold at a fixed value, since the subgrid scale variability must decrease as the resolution becomes increasingly high.

It is therefore concluded that physical parameterizations should depend on the resolution of the atmospheric model. As the model mesh size used in HIRLAM covers a range from that of a large scale model (around 50 km) to a size of about 5 km, it is to be expected that some adjustment of parameterizations with resolution is important. Two strategies are possible here.

One may choose to do some tuning which is often rather 'ad hoc' for a modified resolution. This practice has been widely used in atmospheric modelling due to the lack of precise knowledge about a correct resolution dependence of physics.

Another approach is to realize that physics should be scale dependent and to formulate mathematical expressions for such dependencies, based on reasonable estimates for high ad low resolutions. The latter approach has to some extent been addressed in the formulation of condensation processes in the HIRLAM model (see subsection on clouds and condensation).

Each process has its own problems to address in high resolution. For radiation processes a challenging situation occurs at very high resolution since the transmission of radiation from neighbouring grid boxes can contribute to the local heating rate. In addition, the local cloud cover in a vertical column will in some situations not provide accurate information when computing solar radiation on the ground. This is because slant beams may pass through neighbouring grid boxes with differing cloud amounts. Also the slope of the ground exposed to direct solar radiation becomes locally significant at high resolution, implying increased spatial variability of

solar energy flux to the ground as resolution increases. These phenomena which become of some importance at resolutions below 10 km have not yet been addressed in the HIRLAM physics. The emphasis has been on describing radiative processes with reasonable accuracy in the troposphere. A fast radiation scheme (Savijärvi, 1990) has been further developed for HIRLAM. The scheme is normally called every time step of a forecast. This allows for describing radiation associated with short time fluctuating cloud cover.

The atmospheric subgrid scale transports of heat, moisture and momentum at traditional model resolutions are dominated by the vertical transports. These take place over a large range of spatial scales, from large convective scales to small scales associated with turbulence. The goal is to describe accurately the effect of all subgrid scales on the resolved scales of motion. Traditionally these transports are separated into two scales, namely the small subgrid scales expressed by turbulence, and lager scales described by a convective parameterization. Due to the complexity of atmospheric states it is a challenge to make a coherent formulation of turbulence and convection parameterization. In addition, the transports of moisture also has to describe phase changes due to condensation and evaporation or sublimation of condensate.

At increased model resolution more of the transports will be done by the resolved scale dynamics exposed to larger vertical velocities as a consequence of higher amplitude small scale convergence in the atmospheric flow. Hence the convection parameterization should accomplish less vertical moisture transports as the model resolution increases. A special problem associated with convection schemes is the break down at high resolution of the traditional assumption that convective cloud samples exist in balance with the synoptic forcing acting on the column of air above each grid square. The current HIRLAM model attempts in simple ways to account for some of the scale dependent features of the vertical transports.

The interaction of the model atmosphere with the ground takes place via complex processes of heat, moisture and momentum transfer. The surface processes are the boundary conditions for the vertical subgrid scale transports.

A special effect is created by subgrid scale orography exerting a gravity wave drag on the atmosphere. Parametrization of this effect, which is considered important in synoptic scale models, is currently being implemented into the HIRLAM system. With increased model resolution this model limitation gradually diminishes. Currently the surface fluxes are described in terms of efficient roughness lengths that may differ substantially for momentum, heat and moisture. The turbulent transports are communicated to the atmosphere by a turbulence scheme based on turbulent kinetic energy (Cuxart et al., 2000). The inclusion of this additional dependent variable provides a better potential to describe the effect of turbulent processes as compared to simple first order schemes. One reason is that the 'memory' of turbulence may be advected by the resolved flow. This represents an additional way for physics and dynamics to interact.

The extremely inhomogeneous and complex lower boundary condition with huge spatial variations of the physical properties of the surface in some regions makes it practically impossible in a large scale model to parameterize accurately the averaged surface fluxes of momentum, heat and moisture. The HIRLAM model has made steps to improve on this situation. Global physiographic data bases, e.g., from the U.S. Geological Survey (USGS, 1997, 1998) are utilized to distinguish and describe five parts (forest, low vegetation, bare ground, sea/lake and ice) of a grid square. For each part a separate flux computation is made and combined using the so called 'mosaic' approach. The surface scheme used is the ISBA scheme (Noilhan and Planton, 1989). A unique feature of the HIRLAM model is that this scheme is used for each subsurface of the grid square.

The individual physical parameterizations of HIRLAM are described in some detail below.

3.2 Radiation Scheme

The purpose of a radiation scheme is to provide the surface net radiative fluxes plus the temperature tendency of air resulting from terrestrial (long-wave) and solar (short-wave) radiation. The HIRLAM radiation scheme, based on Savijärvi (1990), is documented in Sass et al. (1994) and Wyser et al. (1999); its main features are summarized here. The temperature tendency is proportional to the net radiative flux divergence according to

$$\frac{\partial T}{\partial t} = -\frac{g}{c_p}\frac{\partial F_{net}}{\partial p}. \tag{3.1}$$

The influence of the surface net radiation flux in the temperature and humidity is calculated by solving the surface energy balance equation.

The scheme was designed to be fast, so only one vertical loop is allowed in both the solar (short-wave, SW) and the thermal (long-wave, LW) part. Only these two spectral bands are considered in the scheme. The SW clear-air global flux is obtained by reducing the top-of-the-atmosphere (TOA) horizontal flux by broadband average ozone absorption (350 DU), water vapour absorption (depending on the scaled precipitable water content u), and Rayleigh scattering in the column. Average aerosol, CO_2, and O_2 effects are also included. In cloudy air the SW flux is reduced by the total cloud transmissivity $\hat{T}$, also taking into account multiple reflections between cloudbase and the surface. In a partly cloudy column the clear-sky and cloudy sky results are linearly combined.

The cloud SW transmissivity and absorptivity functions are fits to a two-stream 5-band radiative transfer model derived from Savijärvi et al. (1997); Hu and Stamnes (1993). The functions depend on the solar zenith angle θ and the modified cloud condensate amount $\hat{M}$ (in $\mathrm{g\,m^{-2}}$), which is the vertical integral above the level under consideration of cloud condensate content CCC, multiplied by the relation of the cloud cover C to the maximum cloud cover of the whole column C_{max},

$$\hat{M}(z) = C_{max}^{-1}\int_z^{\mathrm{TOA}} CCC(z')C(z')\,dz'. \tag{3.2}$$

The absorptivity is

$$\hat{A} = b_{10}(b_{11} + \cos\theta)\log(1 + b_{12}\hat{M}). \tag{3.3}$$

and the transmissivity is given by

$$\hat{T} = \hat{T}_1 / \left(\hat{T}_1 + \hat{M}\right), \tag{3.4}$$

where

$$\hat{T}_1 = b_{13}(b_{14} + \cos\theta). \tag{3.5}$$

Parameters b_{10} and b_{13} are functions of effective droplet radius r_{eff} diagnosed from cloud condensate content according to empirical formulae, based on Martin et al. (1994) for water clouds and Ou and Liou (1995) for ice clouds. Effective radius for the SW calculations is determined as a weighted mean of values above the level under consideration. Values of all b-parameters are given in Wyser et al. (1999).

Clear-air solar heating due to water vapour absorption, $a(u)$, is obtained by vertical flux convergence, fitting $\partial a/\partial u$ from the line-by-line-based broadband $a(u)$ curves of Chou (1986). In

and below clouds the clear-air values are reduced by the cloud transmittance $\hat{T}$. In clouds there is also extra heating due to cloud drop absorption, represented by the flux convergence of the absorptivity, $\partial\hat{A}/\partial p$.

The LW part uses a broadband emissivity scheme in a local isothermal approximation. The water vapour line emissivity is a cubic function of $\log u$; continuum, CO_2, and O_3 effects are added as extra terms. The computation of gaseous longwave emissivities has been modified as proposed by Räisänen et al. (2000). The modified scheme computes explicitly the emissivity for CO_2 and improves the treatment of H_2O continuum aborption.

There can be clouds both above and below the layer in question. Cloud mass absorption coefficient $k_{a,x}$, where the index x stands for water or ice, depends on the effective radius of water droplets and ice crystals, according to an empirical formula from ECHAM 4, c.f. Savijärvi and Räisänen (1998). The effective emissivity ϵ of a cloudy layer is

$$\epsilon = C\{1 - \exp\left[-(k_{a,water}m_w + k_{a,ice}m_i)\right]\}, \tag{3.6}$$

where m_w and m_i are the cloud water and cloud ice amounts (in gm^{-2}) of the layer. The longwave radiative flux below the cloud base is calculated as a combination of contributions from clear-sky and cloud-covered parts.

When calculating the radiative fluxes at surface, grid-square averages of SW albedo (depending on surface type, snow cover and solar zenith angle), LW emissivity (different for ice, snow, water and land surfaces) and skin temperature ($T_{skin} = \sqrt[4]{\sum_{j=1}^{N} T_{s,j}}$, where the index j denotes the N different surface types within the grid square and T_s is the surface temperature) are used.

3.3 Clouds and Condensation

The HIRLAM Project assumed the challenge of achieving a satisfactory performance of the cloud and condensation schemes at all horizontal resolutions down to a grid size of about 5 km. The description of cloud and condensation processes is particularly difficult for several reasons: first, the parameterization of convection depends critically on the resolution of the model; second, the vertical subgrid scale transports conducted by the turbulence and convection schemes should pay attention to their inter-dependence; third, the micro-physical aspects, usually involving additional variables for different hydrometeors, need to be considered. Furthermore, it is often difficult to validate which formulations should be preferred by means of solid observational evidence. In that sense, recent comparisons, as, e.g., the GEWEX (Global Energy and Water Cycle Experiment) Cloud System Study (GCSS), have been useful as a guidance.

Two condensation schemes are currently available in the HIRLAM code. Both schemes make use of the 'cloud condensate' as a prognostic variable which is advected by the model dynamics. The condensation scheme currently used as default (named STRACO) is based on a Kuo type convection scheme with a moisture convergence closure. It puts emphasis on gradual transitions between convective and stratiform regimes. The micro-physics treatment of condensation and precipitation processes follows closely (Sundqvist, 1993). The scheme has been recently upgraded to simulate the entrainment at the cloud top, which compensates the lack of an appropriate treatment of turbulent process in a cloudy environment. The other alternative scheme is based on the (Rasch and Kristjánsson, 1998) scheme for large scale condensation, which is a further development of the Sundqvist scheme, and on the (Kain and Fritsch, 1993) scheme for convection, which was originally designed for mesoscale. The Kain-Fritsch approach is based on an entraining-detraining plume model with CAPE closure and moist downdrafts.

3.3.1 The STRACO scheme for large scale and convective condensation

STRACO (Soft TRAnsition COndensation) parameterizes both large scale and convective condensation and puts special emphasis in achieving gradual transitions between both regimes. It can be considered as a modified version of the 1974 and 1989 schemes. Despite the theoretical shortcomings of the Kuo type convective schemes, STRACO has been proved to be very competitive and it is used in most operational implementations of the HIRLAM community. The microphysics related to the condensation and precipitation processes follow rather closely the comprehensive treatment by 1993.

Convective condensation

The moisture convergence closure includes the effect of surface evaporation. A formulation of the vertical redistribution of cloud condensate is included which is an extension of the original formulations. Also convection can start from any level in the atmosphere whereas several convection schemes treat only deep convection originating from the lowest model layer. The cloud fraction is computed in a semi-prognostic way.

The vertical extent of convection is determined by adiabatic cloud parcel lifting including latent heat release, starting with a small temperature excess as a 'trigger' for convection. Entrainment and detrainment are not explicitly taken into account during parcel ascent, but a simple constraint inhibits the vertical extent of convective clouds in case of a weak net convergence of moisture. Otherwise the Equilibrium Temperature Level (ETL) determines the top of the convective layers forming a convective entity. Formally there is no limit on the possible number of convective entities in the vertical air column. The relevant equations are :

$$\left(\frac{\partial T}{\partial t}\right) = A_T + \frac{L}{c_p}\left(\widehat{Q}_a(1-\beta)\frac{F_h}{\widehat{F}_h}\delta_* - E\right) - \frac{L}{c_p}E_p \tag{3.7}$$

$$\left(\frac{\partial q}{\partial t}\right) = A_q(1-\delta_*) + \widehat{Q}_a\beta\frac{F_q}{\widehat{F}_q}\delta_* + E + E_p \tag{3.8}$$

$$\left(\frac{\partial q_c}{\partial t}\right) = A_{q_c} + \widehat{Q}_a(1-\beta)\frac{F_c}{\widehat{F}_c}\delta_* - E - G_p \tag{3.9}$$

The left hand sides of these equations express the combined effect of both dynamical advection, turbulence and convection. A represent the tendency excluding convection. $\widehat{Q}_a$ is the total moisture accession per unit mass and time in the convective cloud.

$F_h = T_{vc} - T_v + \epsilon_T$ is a function describing the vertical variation of convective heating, $F_q = q_{sc} - q + \epsilon_q$ is a function describing the vertical variation of convective moistening and $F_c = q_{cc} + \epsilon_c$ is a function describing the vertical variation of convective condensate supply. In the above equations for F_h, F_q and F_c index v stands for 'virtual'. $\widehat{F}$ stands for a vertical average value over the convective cloud. Finally, c-index means a value applicable to cloud and s signifies a saturation value. The constants $\epsilon_T, \epsilon_q, \epsilon_c$ are currently set to zero.

The parameter β is a moistening parameter (Kuo, 1974). In the present scheme, contrary to models without prognostic cloud condensate, moistening can take place also from evaporation of cloud condensate.

$$\beta = \left(1 - \frac{\sum_{j=jbot}^{jtop}\frac{q}{q_s}\Delta p}{p_{jbot} - p_{jtop}}\right)^{\gamma_0} \tag{3.10}$$

Here p represents 'pressure', and *jbot* and *jtop* are the model level numbers for the bottom and top of convection, respectively. Currently γ_0 is set to a value of 2. The parameter δ_* determines a link between convective moisture transports on one hand and turbulence plus dynamics effects on the other.

$$\delta_* = B \left(\frac{\Delta p_c}{p_{00}}\right)^{\gamma_1} \left(\frac{\Delta\theta}{\theta_{00}}\right)^{\gamma_2} \tag{3.11}$$

A resolution dependence has been introduced through the scaled latitude increment $\Delta\theta$ for the given resolution. The term involving Δp_c and p_{00} is a scaled cloud depth. Currently $\theta_{00} = 0.15^\circ$ and $p_{00} = 10^4$ Pa. γ_1 and γ_2 are equal to 1. Both factors are constrained to be no larger than 1. The effect of this formulation is that δ_* goes to zero for extremely shallow phenomena and for very high horizontal resolution. A zero δ_* means that the convection scheme is decoupled as it is reasonable in the limit of very shallow phenomena and high resolution where dynamics and turbulence should suffice. The factor B is a dimensionless buoyancy which is 0 for a vanishing mean buoyancy in the convective cloud and becomes 1 when cloud buoyancy exceeds a small threshold. It serves to prevent an abrupt switch from stratiform to convective regime.

The term $E = K_c q_c (q_s - q)$ represents the evaporation/sublimation of cloud condensate that is proportional to the saturation deficit of the environmental air (K_c is a constant). A similar formulation, using cloud cover as the leading variable instead of q_c, has been used by others, e.g., (Tiedtke, 1993) in the ECMWF cloud scheme. Finally, the terms involving G_p and E_p concern generation and evaporation of convective precipitation, respectively.

The equations above are applied to the layers of the convective entities while the stratiform condensation apply to the remaining parts of the atmosphere.

Cloud fraction and cloud condensate

A statistical scheme is chosen using total water content, $q_t = q + q_c$, as prognostic variable. Within a model grid box q_t is assumed to be a random variable with subgridscale variability around the mean, $\overline{q_t}$, described by a Probability Distribution Function (PDF) $P(q_t)$ such that

$$\overline{q_t} = \int_0^{q_{max}} q_t P(q_t) dq_t \tag{3.12}$$

$$\int_0^{q_{max}} P(q_t) dq_t = 1 \tag{3.13}$$

the cloud fraction f and the grid box average of cloud water content are given by

$$f = \int_{q_s}^{q_{max}} P(q_t) dq_t \tag{3.14}$$

$$\overline{q_c} = \int_{q_s}^{q_{max}} (q_t - q_s(T_c)) P(q_t) dq_t \tag{3.15}$$

Using a rectangular asymmetric PDF for convection and defining supersaturation with respect to a characteristic air parcel temperature T_c

$$f_{cv} = \begin{cases} \frac{1}{1+\sqrt{\frac{q_s(T_c)-\overline{q}}{\overline{q}_c}}} & \text{if } q_s(T_c) \geq \overline{q}_{tot} \\ \frac{1}{2} + \frac{(\overline{q}_{tot} - q_s(T_c))}{2A_{cv}\overline{q}_{tot}} & \text{if } \overline{q}_{tot}\cdot(1-A_{cv}) \leq q_s(T_c) \text{ and } q_s(T_c) < \overline{q}_{tot} \\ 1 & \text{if } q_s(T_c) < \overline{q}_{tot}(1-A_{cv}) \end{cases} \tag{3.16}$$

where A_{cv} is a dimensionless threshold ~ 0.1 defining the width of the lower part of $P(q_t)$.

In the stratiform case a symmetric rectangular density function is used and super-saturation is defined with respect to the grid box saturation specific humidity $q_s(\overline{T})$.

$$f_{st} = \frac{1 + A_{st} - \frac{q_s(\overline{T})}{\overline{q}_{tot}}}{2A_{st}} \tag{3.17}$$

The formula above is only used in the range where it provides a cloud cover between zero and one. Outside this range the cloud cover is either 0 or 1, respectively. The probability function for q_{tot} also defines the equilibrium cloud condensate. It is assumed that also the specific humidity at sub-saturation is distributed equally as a rectangular probability function around the grid box average value $\overline{q}$. The amplitude of the fluctuating q_{tot}, as defined by A_{st} is constrained not to violate a rectangular distribution of q.

A similar approach for describing subgrid scale condensation has been used by (Redelsperger and Sommeria, 1986). They emphasize the virtues of having a subgrid scale parameterization of condensation even at horizontal resolutions of a few kilometers.

Effect of resolution. The amplitude A_{st} of the PDF in the stratiform case should ideally depend on model resolution. Redelsperger and Sommeria (1986) estimated that the subgrid scale effects should go to zero quite slowly at high resolution. This, combined with typical values used in coarse resolution models have been used as a guidance for constructing (3.18) below. The effect of resolution is described by the first factor in (3.18). Formulations of A_{st} depending on other factors such as the height above ground have been suggested by Sundqvist et al. (1989) who assume a decrease of the subgrid scale variability close to the ground. A similar idea has been adopted in the STRACO scheme introducing the term in the second bracket of (3.18).

$$A_{st} = A_{st1}\left(1 - \exp\left(-A_{st2}\sqrt{D_{gr}}\right)\right)\left(\frac{A_{st3} + A_{st4}(1 - \sigma_{st}^3)}{A_{st3} + A_{st4}}\right) \tag{3.18}$$

where $A_{st1} = 0.30$, $A_{st2} = 0.003$, $A_{st3} = 0.03$, $A_{st4} = 0.02$. $\sigma_{st} = \min(\frac{p}{p_{st}}, 1)$. $P_{st} = 1013$ hPa. D_{gr} is the grid box length in metres.

'Equilibrium' cloud fraction and condensate. A relaxation towards the equilibrium is applied to define stratiform condensation. However, the constraint is always imposed that gridbox saturation is not exceeded. Condensation associated with equilibrium is determined by a first order adjustment.

$$\Delta q_{st} = \frac{(q_c - q_{ceq})}{1 + f_{st}\frac{L}{c_p}\left(\frac{\partial q_s}{\partial t}\right)} \tag{3.19}$$

Due to the changing conditions as regards q, q_c and T in a model layer, the 'equilibrium' distribution is disturbed. This may, for example, happen with the onset of stratiform conditions after convection or vice-versa. To describe the actual transitions the cloud cover f is made time dependent by relaxing towards the equilibrium cloud cover f_{eq} which may be either a stratiform (f_{st}) or a convective equilibrium (f_{cv}).

$$\frac{\partial f}{\partial t} = -K_f(f - f_{eq}) \tag{3.20}$$

Currently $K_f^{-1} = 900$ s.

Revised convection scheme to enhance fluxes across the cloud top

A revised formulation has been introduce to improve the entrainment on clouds capped by a stable layer. The work has been partially motivated by a deficiency in the precipitation forecasts with the HIRLAM models which tends to predict small precipitation amounts too often. It seems that the underestimation of the entrainment at the top of the low tropospheric clouds could explain part of the problem. The revised formulation increases the heat and moisture exchange between the cloudy layer and the overlying atmosphere expressing the effect of shallow convection partially penetrating the stable layer. The modifications have a significant impact on the simulation of stratocumulus clouds but are also able to describe the effect of overshooting in deep or mid-level convection.

The extension stays within the framework of separating vertical subgrid transports and condensation computations in three steps: first, the vertical transports of potential temperature, specific humidity and cloud condensate are computed by the turbulence scheme; second, adjustments to vertical transports of these variables are performed by the convection and finally, by the stratiform condensation.

Provided that the convection scheme describes the effect of larger turbulent eddies in an environment where condensation takes place, the computations make use of the cloud parcel ascent computation of the convection scheme. Physically we may think of the heat- and moisture transports as accomplished by mainly the larger eddies penetrating through the stable layer on top of a cloud layer. The penetration of these eddies into the stable layer can be estimated from the cloud parcel ascent. We denote by w_b a characteristic vertical velocity of convective motions in the cloud right below the cloud top and will estimate a distance D of penetration into the stable layer. The maximum fluctuation s' of the variable s possible at the interface between cloud and stable layer is estimated to be approximately equal to the increase above the value s_- at the interface on the cloudy side up to the value s_D at depth D into the stable layer. The value s_D is estimated from the mean gradient of the variable at cloud top. More specifically, it is assumed that the flux at the cloud top of the scalar s is a factor $\tilde{w}$ times s'. It is reasonable to assume that the velocity scale $\tilde{w}$ is closely linked to the subgrid scale kinetic energy E already computed in the model: $\tilde{w} = K_* \cdot \sqrt{E}$. $K_* = 0.25$ has been determined after numerical experimentation. The depth of penetration D into the stable layer is estimated as follows: from dimensional analysis it has been argued that the vertical velocity w_r of an idealized thermal depends on its size r, the dimensionless buoyancy B of the 'bubble' and the acceleration of gravity $g(\mathrm{ms}^{-2})$ according the following combination (Rogers and Yau, 1989): $w_r = c\sqrt{gBr}$. Where B is the virtual temperature difference between the cloud parcel and environment, divided by the environmental temperature, $c = 1.2$. Choosing $r = 50$ m as representing the dimension of convective eddies near cloud top we get: $w_b = w_0 \cdot \sqrt{B}$, $w_0 \approx 27\mathrm{ms}^{-1}$. B is computed in the cloud ascent of the convection scheme. It is noted that the turbulent kinetic energy E appearing in the expression of the velocity scale $\tilde{w}$ is assumed to be no less than $1.0B\mathrm{m}^2\mathrm{s}^{-2}$ in order to avoid a risk of very low values with the present turbulence scheme in the case of small wind shear.

The maximum penetration depth D is estimated from the deceleration in the stable layer. Utilizing the start velocity of w_b for the deceleration we get

$$D = w_o \sqrt{\frac{B \cdot T}{g|\gamma_c - \gamma|}} \tag{3.21}$$

where D is in meters, T is temperature (K), γ_c in Km^{-1} is the lapse rate associated with moist adiabatic ascent ($\gamma_c > 0$) and γ in Km^{-1} is the ambient lapse rate in the stable layer. It is demanded that $\gamma < \gamma_c$. A numerical security computation has been implemented to avoid extreme behavior if the two lapse rates become almost equal, and the penetration is not allowed

to exceed a depth corresponding to 25 hPa. The sensible heat flux $F_H(\mathrm{Jm^{-2}s^{-1}})$ at the level of transition between cloud and the stable layer is computed according to

$$F_H = \rho C_p \cdot \tilde{w} \cdot D \cdot \frac{\partial \theta}{\partial z} \tag{3.22}$$

Similarly we get for the moisture flux of total specific humidity $F_{q_t}(\mathrm{kg_{water}m^{-2}s^{-1}})$ at the transition level

$$F_{q_t} = \rho \cdot \tilde{w} \cdot D \cdot \frac{\partial q_t}{\partial z} \tag{3.23}$$

We assume that the fluxes of heat and moisture determined from the above formulas are distributed linearly with height in the convective cloud depth D_- and in a stable layer D_+. The latter should approximately be equal to D apart from the constraints set by vertical resolution. If D is larger than the depth of one model layer above cloud a sufficient number of levels are included to exceed D. Currently specific humidity q and cloud condensate q_l are processed independently according to the method described above, but the flux of the moist conserved variable of 'total specific humidity' is then also linear, which preserves moisture structures in a well mixed cloud.

The present formulation of shallow convection has been tuned to compensate the limitation of the current turbulence scheme, which is not designed to describe turbulence adequately inside clouds. Any modification of the turbulence scheme towards a moist formulation of the turbulent processes would require adjustments of the present formulation of shallow convection.

3.3.2 The Rasch-Kristjánsson scheme for stratiform regime

This parameterization (hereafter RK) is based on the work of Sundqvist (1988) and was originally developed by Rasch and Kristjánsson (1998) for the National Center for Atmospheric Research (NCAR) Community Climate Model (CCM3). The treatment of condensation and evaporation processes follows quite closely the Sundqvist scheme although using specific humidity instead of relative humidity in the computations. The microphysics (conversion of cloud condensate into precipitation) is different from Sundqvist, following somehow the formulations used in Cloud Resolving Models (CRM). The formation of precipitation is clearly separated into five process that make the diagnosis and improvement of the parameterization easier. Although, only one predicted variable is used for cloud condensate, four different species are represented. A diagnostic approach is used to compute cloud fraction.

Cloud fraction parameterization

A diagnosis of fractional cloud cover is computed preceding the condensation process. The parameterization is the standard CCM3 formulation based on Slingo (1987). Some modifications to this scheme are introduced as an attempt to remove well known biases of the CCM3 simulations. The cloud fraction in the Slingo scheme depends on relative humidity, vertical motion, static stability and convective properties. In RK a modification is introduced to take into account the detrainment rate from convective updrafts. There is also a modification for the diagnosis of marine stratocumulus which is based on the mean stratification in the lower troposphere. The cloud fraction scheme requires that at least 1% cloud fraction is diagnosed in any gridbox where relative humidity exceeds 99%.

Condensation rate

The parameterization of the rate at which water vapour is converted to condensate (and vice-versa) is essentially identical in Sundqvist and RK, though the equations are developed using

relative humidity in Sundqvist and using specific humidity in RK. Here the equations will be derived following the method of RK. The tendency equations for specific humidity (q), temperature (T) and cloud condensate (q_c) can schematically be written as

$$\frac{\partial q}{\partial t} = A_q + S_q + C_q - (Q - E) \tag{3.24}$$

$$\frac{\partial T}{\partial t} = A_T + R_T + S_T + C_T + \beta(Q - E) \tag{3.25}$$

$$\frac{\partial q_c}{\partial t} = A_{q_c} + S_{q_c} + C_{q_c} + (Q - E) - P \tag{3.26}$$

where $\beta = L/c_p$, A represents the advection, C the convection and $Q - E$ the tendency from condensation and evaporation of cloud water that needs to be parameterized. R is the radiative forcing and S is the tendency term from other subgridscale processes as vertical diffusion. The calculation of the precipitation term, P, will be explained in the next subsection. In order to derive closure relations for the $Q - E$ term, let's write the saturation pressure, given by the Clausius-Clapeyron equation, as $q_s = q_s(T, p)$. One can differentiate in time and get

$$\frac{\partial q_s}{\partial t} = \frac{\partial q_s}{\partial T}\frac{\partial T}{\partial t} + \frac{\partial q_s}{\partial p}\frac{\partial p}{\partial t} = \alpha\frac{\partial T}{\partial t} + \gamma\frac{\partial p}{\partial t} \tag{3.27}$$

The grid box is partitioned into a cloud free part and a cloudy part where the air is assumed to be saturated with respect to water vapour.

$$q = (1 - f)q + f q_s \tag{3.28}$$

By differentiating (3.28), substituting (3.24),(3.25) and (3.26), and rearranging we get

$$\begin{aligned}(q - q_s)\frac{\partial f}{\partial t} + (1 - f)\frac{\partial q_c}{\partial t} + (1 + f\alpha\beta)(Q - E) = \\ A_q + S_q + C_q - f\alpha A_T + R_T + C_T - \gamma\frac{\partial p}{\partial t} \equiv M\end{aligned} \tag{3.29}$$

The right hand side of (3.29) contains all known tendencies from other processes. Whereas in Sundqvist the cloud fraction derivate is a function of other variables, in RK, f is more complex and its derivate is computed from f at two adjacent time steps

$$\frac{\partial f}{\partial t} = (f^{new} - f^{old})/\Delta t \tag{3.30}$$

The left hand side terms of (3.29) can be determined by two closure assumptions similar to those in Sundqvist.

Closure assumption 1: The fraction of M operating within the cloud part of the volume acts to condensate/evaporate cloud water

$$(1 + f\alpha\beta)(Q - E)_{cloudy} = fM \tag{3.31}$$

Closure assumption 2: When the cloud is growing or dissipating, the cloud water in the volume adjusts to keep the cloud water of the cloudy part

$$(1 + f\alpha\beta)(Q - E)_{clear} = \frac{\partial f}{\partial t}\tilde{q}_c \tag{3.32}$$

qwhere $\tilde{q_c} = q_c/f$ is the incloud condensate. The unknown term $(Q - E)$ can now be written as

$$Q - E = (Q - E)_{clear} + (Q - E)_{cloudy} + (Q - E)_{rain} = \frac{(fM + \frac{\partial f}{\partial t}\tilde{q_c})}{1 + f\alpha\beta} + E_{rain} \tag{3.33}$$

since Q_{rain} =0. E_{rain} is determined in the parameterization of microphysics.

Parameterization of cloud microphysics

In RK four types of condensate are represented: suspended liquid and ice (q_l and q_i), and liquid and ice falling (q_r and q_s). Currently, only the total suspended condensate, $q_c = q_l + q_i$, is a prognostic variable as in Sundqvist. At the beginning of the computations, q_c is decomposed into liquid and ice assuming that ice phase increases linearly from 0^{o}C to -20^{o}C. Precipitation is assumed to be snow at temperatures below freezing.

Five processes are considered to convert condensate into precipitation: autoconversion, coalescence, local production of ice and ice collecting rain and snow. The conversion rates are parameterized following bulk microphysics formulation used in smaller scale CRM. Autoconversion, collision between cloud drops locally to form precipitation particles, is parameterized following Chen and Cotton (1987). Coalescence, collection of cloud liquid water by falling rain particles is parameterized in accordance with Tripoli and Cotton (1980). Local conversion of ice to particles with mass sufficient to reach a fall speed relative to the air, ice autoconversion, is modelled following Lin et al. (1983), in a form proposed for liquid processes by Kessler (1969), but with a temperature dependence similar to the one proposed by Sundqvist (1988). Accretion, the collection of ice by snow, and the collection of liquid by snow, follow Lin et al. (1983). Evaporation of precipitation is function of relative humidity and is parameterized as in Sundqvist et al. (1989). All snow is assumed to melt in layers where the temperature exceeds 0^{o}C. For further details see Rasch and Kristjánsson (1998).

3.3.3 The Kain-Fritsch convection scheme

The Kain Fritsch scheme is a one-dimensional entraining/detraining plume model that is specifically designed for application in mesoscale convective parameterization schemes. Only one cloud is resolved in this scheme, as the mesoscale models for which it is designed have quite small grid boxes. These usually contain only a small number of different clouds. Resolving a whole ensemble of clouds also would be too time consuming and costly.

The available buoyant energy (ABE) in the grid box is assumed to be consumed in time τ_c, where τ_c is the advection time at the Lifting Condensation Level (LFC) if this lies between 1800 and 3600 seconds. If the advection time is less than 1800 seconds then τ_c is set at 1800 seconds, if it is more than 3600 seconds it is set at 3600 seconds. The parameterized mass flux of the single cloud usually is not enough to remove all ABE in time τ_c. To achieve the (almost) full consumption of the ABE the mass flux is adjusted to a level where the remaining ABE is less than 10% of the initial value. The increase in mass flux can be seen as an increase in the number of clouds that are present in the grid box.

Trigger functions initiating convection

In the scheme, two functions are used to trigger convection. The second trigger can be extended with an orography part, but due to the absence of one parameter in the description of

the orography it cannot be implemented yet. These triggers represent the subgrid scale variability induced by turbulence, orography or inhomogeneities in relative humidity and temperature.

The first part of trigger 1 represents the temperature perturbations associated with vertical wind speed perturbations. It depends on the grid box averaged vertical wind speed:

$$W_0 = W(LCL)\Delta x/25000 - W_{KLCL}, \tag{3.34}$$

where $W(LCL)$ is the vertical speed at the LCL, Δx is the grid box length and W_{KLCL} is defined as:

$$W_{KLCL} = \begin{cases} 0.02 & \text{if } Z_{LCL} \geq 2000 \text{ m} \\ 0.02 * Z_{LCL}/2000 & \text{if } Z_{LCL} < 2000 \text{ m} \end{cases} \tag{3.35}$$

Z_{LCL} is the lifting condensation level. W_0 is dependent on the grid box length and scale by 25 km. It is assumed that the vertical wind component is linearly dependent on the grid box length. The final temperature perturbation belonging to the vertical wind component trigger is defined as (Fritsch and Chappell, 1980):

$$\Delta T_w = 4.64\,(W_0)^{1/3} \qquad \text{if } W_0 \geq 0 \tag{3.36}$$

The second part of trigger 1 depends on the relative humidity and can be thought of as caused by the variance in the relative humidity distribution. In this trigger the relative humidity (Rh) at the LCL is defined as:

$$Rh(LCL) = Q_{env}/Q_s(LCL) \tag{3.37}$$

where $Q_{env}(LCL)$ and $Q_s(LCL)$ are the environmental and the saturation mixing ratio at the LCL. Another parameter that is used in the trigger calculation is the derivative of the saturation mixing ratio as a function of the temperature. This is defined as:

$$\partial Q_s/\partial T = Q_s \left(\frac{C_{liq} - B_{liq} D_{liq}}{(T(LCL) - D_{liq})^2} \right) \tag{3.38}$$

where $T(LCL)$ is the temperature at the LCL and B_{liq}, C_{liq} and D_{liq} are constants (17.502, 4780.8 and 32.19 respectively). The final trigger temperature perturbation is:

$$\Delta T_{rh} = \begin{cases} 0.25\,(Rh(LCL) - 0.75)\,Q_{mix}/(\partial Q_s/\partial T) & \text{if } 0.75 \leq Rh(LCL) \leq 0.95 \\ (1/Rh(LCL) - 1)\,Q_{mix}/(\partial Q_s/\partial T) & \text{if } Rh(LCL) > 0.95 \end{cases} \tag{3.39}$$

This trigger increases from zero to a maximum at a relative humidity of 95% and falls of rapidly towards zero again above this value. It represents subgrid clouds that can be present when the relative humidity is higher than a certain threshold (usually 75% or 80 %). The two parts of trigger 1 in equations 3.36 and 3.39 are combined to see if a cumulus cloud can develop. The complete trigger 1 is defined as:

$$\Delta T_w + \Delta T_{rh} + T_{pack}(LCL) > T_{env}(LCL). \tag{3.40}$$

When this condition is met, convection is possible and the mass flux calculations are carried out. In equation 3.40, $T_{pack}(LCL)$ is the temperature of the lifted parcel at the LCL and $T_{env}(LCL)$ is the environmental temperature at the LCL.

The second trigger that is used in the cumulus parameterization depends on the boundary layer turbulence and variance in orography. Under the right conditions boundary layer turbulence can

become quite strong, causing powerful eddies to reach the boundary layer top and penetrate into the free atmosphere. These eddies can contain enough energy to overcome the energy necessary to reach the LCL and the Level of Free Convection (LFC). The magnitude of the upwards velocity in boundary layer eddies is characterized by the free convective velocity scale (w_*) that is defined as:

$$w_* = \left(\frac{g z_i}{\overline{\theta_v}} \overline{w' \theta_v'} \right)^{1/3} \tag{3.41}$$

where g is the gravity acceleration, z_i is the height of the boundary layer and $g/\overline{\theta_v} * \overline{w'\theta_v'}$ is the surface buoyancy flux. If the mid-level height of the source layer of the updraft lies within the boundary layer, then:

$$\Delta_{turb} = w_*, \tag{3.42}$$

where Δ_{turb} is the perturbation caused by the boundary layer eddies. If the mid-level height of the source layer lies above the boundary layer, then Δ_{turb} is defined as:

$$\Delta_{turb} = w_* exp - \left[30000 \left(\frac{\partial \theta_v}{\partial z} \right) / \theta_{e,neg} + 5 * 10^{-7} (z_{source} - z_i)^2 \right]. \tag{3.43}$$

Here z_{source} is the depth of the source layer for the convection, $(\partial \theta_e / \partial z)$ is the virtual potential temperature gradient between the top of the boundary layer and the mid-level height of the source layer and $\theta_{e,neg}$ is given by:

$$\theta_{e,neg} = \sum_{k(z_i)}^{k(z_s)} T(k) Exn(k) \Delta P(k) / [P(k(z_s)) - P(k(z_i))], \tag{3.44}$$

where Exn is the Exner function, $\Delta P(k)$ is the depth of layer k (in Pa), $k(z_i)$ is the height of the boundary layer in full model levels and $k(z_s)$ the model level just below the mid-level height of the source layer. To enable the triggering of convection due to the boundary layer turbulence, Δ_{pbl} has to be larger than the root of the Convective Inhibition Energy (CINE).

The second part of trigger 2, which is not utilized completely in HIRLAM due to the absence of subgrid scale orographic variance, but included in the code for when this field becomes available, represents the variance in the vertical wind speed induced by the subgrid scale orographic variance. It is defined as:

$$\Delta_{oro} = \begin{cases} \Delta_{base} & \text{if } \sigma_{oro} < 100 \\ \Delta_{base}[((\sigma_{oro} - 100)/300) + 1] & \text{if } \sigma_{oro} \geq 100 \end{cases} \tag{3.45}$$

Here, σ_{oro} is the orographic standard deviation and Δ_{base} is defined as:

$$\Delta_{base} = 0.2(\Delta x / 10000)(1 - [z_{source}/5000]) \tag{3.46}$$

where Δx is the grid box length. If σ_{oro} is not present then $\Delta_{oro} = \Delta_{base}$

The complete second trigger is given by:

$$\Delta_{pbl} = (\Delta_{turb} + \Delta_{oro}) \left(1 + C(\nabla)^{1/3} \right). \tag{3.47}$$

where ∇ is the large scale divergence/convergence of the layer being lifted and C is a constant ($= 10$). Δ_{pbl} then is compared to $\theta_{e,neg}$ and convection is triggered when $\Delta_{pbl} > \theta_{e,neg}$.

Mass flux calculations

With W_0 from equation 3.34 the radius (R) of the initial updraft mass flux is calculated. This radius is defined as:

$$R = \begin{cases} 1000 & \text{if } W_0 \leq 0 \\ 1000 + 1000W_0/0.1 & \text{if } 0 \leq W_0 \leq 0.1 \\ 2000 & \text{if } 0.1 \leq W_0 \end{cases} \tag{3.48}$$

After determinig this radius, the initial updraft mass flux at the LCL is calculated with the assumption that the initial vertical wind speed in the updraft is 1 ms^{-1}:

$$M_{u0} = \rho(LCL)\pi R^2, \tag{3.49}$$

where M_{u0} is the initial updraft mass flux and $\rho(LCL)$ is the density of the air at the LCL.

When the initial mass flux and radius have been determined, the up- and downdrafts are calculated. In these calculations it is assumed that any mixture of updraft and environmental air that becomes negatively buoyant is detrained from the cloud. The mixtures that remain positively buoyant entrain into the cloud. These mixtures develop near the edges of the updraft, where environmental air is entrained into the cloud. As this environmental air usually is subsaturated, cloud droplets evaporate in the mixtures. Mixtures with a high percentage of environmental air therefore become can become colder than the environment, making them negatively buoyant and causing them be detrained from the cloud.

One of the important parameters for the updraft calculations is the rate at which environmental air is made available for entrainment into the cloud. This rate is assumed to be (Simpson, 1983; Kain and Fritsch, 1990):

$$\delta M_e = M_{u0}(0.03\delta p/R), \tag{3.50}$$

where δM_e is the rate at which environmental air is made available for entrainment and δp the depth of the layer in Pa. Equation 3.50 prescribes that when there is no detrainment, the updraft mass flux doubles if it travels 500 hPa upwards. The updraft mass, with which this environmental air mixes, must become available at the same rate, which leads to(Kain and Fritsch, 1990):

$$\delta M_t = \delta M_e + \delta M_u. \tag{3.51}$$

Here δM_t is the total rate of mass entrainment into the mixing region and δM_u is the rate of mass entrainment from the updraft into the mixing region. As mentioned previously, the amount of mass that detrains out of the cloud will be dependent on the mixtures of environmental and updraft air. For these subparcel mixtures a Gaussian distribution is assumed (Kain and Fritsch, 1990):

$$f(x) = A_g\left(e^{-(x-m)^2/2\sigma^2 - k}\right), \tag{3.52}$$

where x is the fraction of environmental air in the mixtures, m is the distribution mean (0.5), σ is the standard deviation (1/6) and k an constant so that $f(0) = f(1) = 0$, so $k = e^{-4.5}$. A_g is defined such that:

$$\int_0^1 f(x)dx = 1, \tag{3.53}$$

which means that $A_g = (0.97\sigma\sqrt{2\pi})^{-1}$. The distribution in equation 3.52 gives a specification of the rates at which various mixtures are generated. Assuming that the subparcel size is independent of the mixing proportion allows the total mass distribution to be obtained simply by multiplying the frequency distribution δM_t (Kain and Fritsch, 1990):

$$\delta M_e + \delta M_u = \delta M_t \int_0^1 f(x)dx. \tag{3.54}$$

The individual components of this distribution are given by:

$$\begin{aligned} \delta M_e &= \delta M_t \int_0^1 x f(x)dx \\ \delta M_u &= \delta M_t \int_0^1 (1-x) f(x)dx. \end{aligned} \tag{3.55}$$

From these equations the total entrainment into and detrainment out of the updraft can be calculated. As the negatively buoyant parcels detrain from the updraft, the updraft detrainment rate (M_{ud}) is determined from:

$$M_{ud} = \delta M_t \int_{x_c}^1 (1-x) f(x)dx. \tag{3.56}$$

Similarly, the environmental entrainment rate (M_{ee}) is calculated from:

$$M_{ee} = \delta M_t \int_0^{x_c} x f(x)dx, \tag{3.57}$$

where x_c is the fractional amount of environmental mass that just yields a neutrally buoyant mixture. A very small x_c (very dry air) will therefore cause a high detrainment rate while very moist environmental air will enhance the updraft through a large M_{ee}.

Shallow convection

If the trigger function for deep convection is satisfied, but the resulting cloud depth is less than that specified for deep convection (Cloud depth of ~3000m), then the source layer supporting the deepest cloud is saved and shallow convection is tested for. Presently the shallow convective cloud is treated as a detraining plume from cloud base. The intensity of shallow convective activity is set by the closure assumption that shallow convection consumes the maximum TKE available in the model Boundary Layer in a preset time period of 40 minutes.

Conversion of condensate to precipitation

The amount of precipitation that is lost in a certain layer is given by (Ogura and Cho, 1973; Kain and Fritsch, 1990):

$$\delta r_c = r_{c0}(1 - e^{-c_1 \delta z / w}), \tag{3.58}$$

where δr_c is the condensate that is lost (precipitated out), δz is the depth of the layer, r_{c0} is the condensate concentration at the bottom of the layer plus one half the degree of supersaturation at the top, w is the mean vertical velocity in the layer and c_1 is a proportionality constant (0.01).

Cloud glaciation

Glaciation in the clouds is parameterized by assuming a linear transition from θ_e with respect to water to θ_e with respect to ice. This transition takes place between 268K and 248K. A hybrid value is used during the transition. θ_e is defined as (Kain and Fritsch, 1990):

$$\theta_e = \theta exp[1.0723 * 10^{-3}(1 + 0.81r)Lr/T], \tag{3.59}$$

in which:

$$\begin{aligned} L &= (1-\nu)L_v + \nu L_s, \\ r &= (1-\nu)r_{sl} + \nu r_{si}. \end{aligned} \tag{3.60}$$

L_v and L_s are the latent heating for vaporization and sublimation processes respectively, r_{sl} and r_{si} are the saturation vapor pressures over liquid water and ice respectively and ν is the level of glaciation which is zero if T = 268K, one if T = 248K and varies linearly in between.

The freezing of condensate takes place similarly. The magnitude of the latent heat release/consumption at any level in the transition region is a function of the combined influence of freezing/melting and evaporation/condensation/sublimation necessary to maintain saturation at the new temperature and hybrid saturation mixing ratio.

3.4 Fluxes of momentum, heat and moisture at the surface

Turbulent surface fluxes in numerical weather prediction models are traditionally computed from drag formulae relating the surface fluxes to the mean states of the surface and of the atmosphere at the observation height (in a numerical model, typically the lowest model level). The drag coefficient C_γ for a scalar variable γ is defined by the equation

$$\overline{w'\gamma'} = C_\gamma \Delta\gamma |\vec{V}_N| \tag{3.61}$$

in which $\overline{w'\gamma'}$ is the vertical turbulent kinematic flux of γ and $\Delta\gamma = \gamma_s - \gamma_N$.

The scalar variables and the wind vector on the right hand side of (3.61) are time averages like the flux term on the left hand side, but for convenience the averaging symbol has been omitted on the right hand side. The sign convention in (3.61) is such that upward fluxes are positive.

Over the ocean $\vec{V}_N$ should, strictly speaking, be measured in a frame moving with the ocean surface current and T_s should be the surface skin temperature.

The surface specific humidity q_s is not so obvious to specify as compared to surface wind speed $\vec{V}_s$ which may be assumed to be zero.

Over sea the saturation value $q_{sat}(T_s)$ with respect to the surface temperature is assumed. Over land a soil wetness parameter $F_{wet} = a_{w1} + (1 - a_{w1})(\frac{W_s}{W_{sat}})^{a_{w2}}$. W_s and W_{sat} are the soil water (m) of the top model soil layer and the saturation soil water value, respectively ($a_{w1} = 0.05$, $a_{w2} = 8$). The constraint is always imposed that the saturation value $q_{sat}(T_s)$ is never to be exceeded.

$$q_s = F_{wet} q_{sat}(T_s) + (1 - F_{wet}) q_N \tag{3.62}$$

In the current formulation

$$C_\gamma = C_{MN} \left(1 + \ln \frac{z_{0M}}{z_{0H}} \Big/ \ln \frac{z}{z_{0M}}\right)^{-1} \Psi_\gamma \left(Ri, \frac{z}{z_{0H}}, \frac{z}{z_{0M}}\right) \tag{3.63}$$

This formulation is used for different surface types and forecast scalar parameters. ($\gamma = M$ for momentum, correspondingly H and Q stand for heat and moisture transfer, respectively).

The two factors in front of the function Ψ_γ form a neutral drag coefficient C_{γ_N}. For identical roughness lengths z_{0M} and $z_{0\gamma}$ the second factor in the brackets becomes unity and $C_{\gamma_N} = C_{MN}$.

$$C_{MN} = \left(\frac{k}{\ln(\frac{z}{z_{0M}})}\right)^2$$

$k \sim 0.4$ is the von Kármán constant.

The functions used for Ψ_γ follow the work of Louis (1979) and Louis et al. (1982).

The following form applies to the unstable atmospheric boundary layer:

$$\Psi_\gamma = 1 + \frac{a_{\gamma_U} Ri}{1 + b_{\gamma_U} C_{\gamma_N} (Ri \frac{z}{z_{0M}})^{\frac{1}{2}}} \tag{3.64}$$

In (3.64) above Ri is the surface bulk Richardson number. $a_{mU} = 10$, $b_{mU} = 75$, $a_{HU} = a_{QU} = 15$ and $b_{HU} = b_{QU} = 75$.

A special situation occurs as the mean horizontal wind speed goes to zero in an unstable boundary layer. It turns out that realistic surface fluxes can be achieved in the framework of the existing formula provided that z_{0M} in (3.64) is replaced by the length scale d_γ, (Nielsen, 1999), where $d_M = 1.03\frac{\nu}{u_{fc}}$, $d_H = d_Q = 1.14\frac{\nu}{u_{fc}}$, and ν is a kinematic viscosity of air. u_{fc} is a velocity scale associated with free convection

$$u_{fc} = \left(\frac{g}{\theta_v} \Delta\theta_v \nu\right)^{\frac{1}{3}} \tag{3.65}$$

In (3.65) g is the acceleration of gravity and $\Delta\theta_v$ expresses the virtual potential temperature difference between the atmosphere and the sea surface.

For the stable boundary layer the corresponding formula is

$$\Psi_M = \frac{1}{1 + \frac{a_{MS} Ri}{\sqrt{1 + b_{MS} Ri}}} \tag{3.66}$$

In (3.66) above $a_{MS} = 10$, $b_{MS} = 1$.

For heat and moisture the following function is used

$$\Psi_\gamma = \frac{1}{1 + a_{\gamma S} Ri \sqrt{1 + b_{\gamma S} Ri}} \tag{3.67}$$

$a_{HS} = a_{QS} = 10$ and $b_{HS} = b_{QS} = 1$.

For flux computations over land and sea ice the potential of the formulae above to incorporate effects of separate roughness lengths for momentum, heat and moisture, is not used. Currently the roughness lengths for heat, moisture and momentum are taken to be equal. The actual value of roughness depends strongly on land surface type according to basic theories of the planetary boundary layer (Garrat, 1977). The actual values are computed in the climate field generation. Values of heat and moisture roughness up to 0.5 m are allowed for. However, it turns out that larger values for momentum roughness are needed to describe surface drag in a formulation based on roughness length. This is because the effect of subgrid scale orography is significant. The current algorithm for such computations makes the momentum roughness proportional to the variance of subgrid scale orography. In this way it turns out that momentum roughness over land can achieve values of several meters. The roughness length for ice is taken to be 0.03 m which is substantially higher than current estimates for a plane ice surface, in order to account for horizontal irregular features occurring in natural ice fields.

To complete the parameterization for surface flux computations over sea the ratio of the roughness lengths z_{0M}/z_{0H} and z_{0M}/z_{0Q} must be related to known quantities. Measurements over sea indicate that the surface roughness lengths for momentum, sensible heat and moisture are different, (DeCosmo et al., 1996). The main reason is probably that form drag (i.e., momentum transport by pressure forces) in the presence of sea waves enhances the momentum transfer, while the heat and moisture transfers at the air-sea interface are controlled by molecular diffusion alone.

According to observations and laboratory experiments z_{0M} depends on the sea state. At low wind speeds laboratory experiments indicate that the sea surface becomes aerodynamically smooth with z_{0M}=0.11 ν/u_*. At sufficiently high wind speeds the sea surface becomes aerodynamically rough. In the latter conditions the Charnock formula $z_{0M} = \beta u_*^2/g$, (Charnock, 1955) with a value of β in the range of 0.011 to 0.032, are widely used. In numerical models values of β in the interval of 0.014 to 0.0185 are recommended (Garrat, 1992). We introduce a transition in z_{0M} from a smooth to a rough sea surface depending on wind speed $|\vec{V}_N|$, where index N refers to the height of the lowest model level. The interpolation formula for z_{0M} is specified as

$$z_{0M} = (1 - \xi(\mathrm{u}))\, 0.11\, \nu/u_* + \xi(\mathrm{u})\beta\frac{u_*^2}{g}\,, \tag{3.68}$$

with

$$\xi(\mathrm{u}) = \left[\max\left(\min\left(\frac{u - u_s}{u_r - u_s}, 1\right), 0\right)\right]^{1/2}\,. \tag{3.69}$$

In (3.68) and (3.69) $u = |\vec{V}_N|$, $u_s = 3.0$ m/s and $u_r = 5.0$ m/s. According to (3.68) and (3.69) the sea surface is considered to be smooth for $u \leq 3$ m/s and rough for $u \geq 5$ m/s. The transition interval is somewhat smaller than the interval from 2.5 to 5.5 m/s usually found in the literature e.g., Garrat (1992). Over open sea (here defined as fraction of sea equal to 1) β has been set to 0.014, otherwise $\beta = 0.032$. The former value fits measurements over the open ocean quite well, e.g. Yelland and Taylor (1996), Grachev et al. (1998). A higher value of β (larger momentum flux) in shallow coastal waters is supported by measurements (Oost, 1998; Hansen and Larsen, 1997; Maat et al., 1991), although the governing parameter for z_{0M} seems to be the wave age.

We assume that z_{0H} and z_{0Q} can be related to z_{0M}. Observational evidence of a relation between z_{0M} and z_{0Q} has been presented for example by DeCosmo et al. (1996). Dependence of z_{0H} and z_{0Q} on z_{0M} are also key features in 'surface flux over sea' models by e.g. Clayson et al. (1996), Liu et al. (1979), Brutsaert (1975). In our parameterization z_{0H} and z_{0Q} are related to z_{0M} by formulae suggested in Garrat (1992). In terms of the roughness Reynolds number Re_* the relations take the form

$$\ln\frac{z_{0M}}{z_{0H}} = \alpha_H Re_*^{1/4} - 2\,, \tag{3.70}$$

$$\ln\frac{z_{0M}}{z_{0Q}} = \ln\frac{z_{0M}}{z_{0H}} - \alpha_Q Re_*^{1/4}\,, \tag{3.71}$$

in which Re_* is defined as

$$Re_* = \frac{z_{0M} u_*}{\nu}\,, \tag{3.72}$$

i.e., by the sea surface momentum roughness length (z_{0M}), the surface friction velocity (u_*) and the molecular kinematic viscosity coefficient for air ($\nu \simeq 1.5 \cdot 10^{-5}\,\mathrm{m}^2/\mathrm{s}$). Over a rough sea α_H=2.48 and α_Q=0.2. With the transition from a smooth to a rough sea surface given by (3.69),

and over a smooth surface with the requirement $z_{0\gamma} \cdot u_*/\nu = S_\gamma$ for $\gamma = H, Q$ and $S_H \approx 0.2$ and $S_Q \approx 0.3$ (Garrat, 1992), the coefficients α_H and α_Q in (3.70) and (3.71) become

$$\alpha_H = 0.05\xi(\mathrm{u}) + 2.43\,, \tag{3.73}$$

$$\alpha_Q = -0.50\xi(\mathrm{u}) + 0.70\,. \tag{3.74}$$

3.5 The turbulence scheme

3.5.1 Introduction

A realistic representation of the Planetary Boundary Layer physics is essential for a good performance of Numerical Weather Prediction models. Formation of air masses, profiles of the wind, temperature and humidity in the lower atmosphere, boundary layer clouds, forecast of 2 meter parameters, fog and dew/frost formation depend on a proper formulation of the PBL scheme.

The present turbulence scheme is based on prognostic turbulent kinetic energy (TKE or E) combined with a diagnostic length scale: a TKE-l scheme. One of the advantages of such a approach [compared to e.g. a local Ri based scheme such as the Louis (1979) scheme] is that e.g. entrainment at the top of the boundary layer is represented. In addition, the TKE-l scheme is thought to be more suitable for the present high resolution, short time range predictions by HIRLAM, as it includes a prognostic equation of the turbulence including local and nonlocal (advective) terms.

The present TKE-l scheme is a modification of the CBR scheme developed for HIRLAM by Cuxart, Bougeault and Redelsperger (Cuxart et al., 2000). The CBR scheme uses the Bougeault and Lacarrère (1989) length scale formulation, in which the length scale is computed from distance which an upward (and downward) adiabatic parcel can travel before being stopped at a level where it has lost all its kinetic energy by buoyancy effects. The new scheme includes important changes in this length scale formulation. The reason for those are discussed in Lenderink and de Rooy (2000); the main being that CBR did not fulfill surface layer scaling for neutral conditions (see below). In addition some more technical changes are made compared to CBR: i) changing levels of TKE from full to half levels (see below) and, ii) changing diffusion solver in order to incorporate density changes with height (the original CBR is based on the Boussinesq approximation). We will denote the new scheme in the following CGL.

3.5.2 TKE framework

The turbulence scheme scheme computes turbulent fluxes:

$$F_\phi = -K_\phi \frac{\partial \phi}{\partial z} \tag{3.75}$$

where ϕ is u, v, θ etc. This gives for the tendencies a diffusion term. In a TKE-l scheme the eddy diffusivity $K_{m,h}$ for momentum/heat is computed from:

$$K_{m,h} = l_{m,h}\sqrt{E}, \tag{3.76}$$

where $l_{m,h}$ is a diagnostic length scale (defined below) and the E is the turbulent kinetic energy, which is computed from a prognostic equation (Stull, 1988):

$$\frac{\partial E}{\partial t} = \underbrace{-\overline{u'w'}\frac{\partial u}{\partial z} - \overline{v'w'}\frac{\partial v}{\partial z}}_{S} + \underbrace{\frac{g}{\theta_v}\overline{w'\theta_v'}}_{B} \underbrace{-\frac{\partial}{\partial z}(\overline{w'E} + \overline{w'p'}/\rho)}_{T} - \epsilon. \tag{3.77}$$

The shear production S, the buoyancy production/consumption B, the transport T by turbulence and pressure forces, and the dissipation D (or ϵ) are parameterized by

$$-\left[\overline{u'w'}\frac{\partial u}{\partial z}+\overline{v'w'}\frac{\partial v}{\partial z}\right]=K_m\left[\frac{\partial^2 U}{\partial x^2}+\frac{\partial^2 V}{\partial y^2}\right]$$

$$\frac{g}{\theta_v}\overline{w'\theta_v'}=-K_h\frac{g}{\theta_v}\frac{\partial\theta_v}{\partial z}\equiv -K_h N^2$$

$$-\frac{\partial}{\partial z}(\overline{w'E}+\overline{w'p'}/\rho)=2K_m\frac{\partial E}{\partial z}$$

$$\epsilon=c_d\frac{E^{3/2}}{l_m} \tag{3.78}$$

with c_d a constant. The factor 2 in the transport term is added to account for the effects of the pressure induced transport of E. Note also the definition of the Brunt-Vaisala frequency N^2.
The surface boundary condition for TKE is given by

$$E_{surf}=c_o u_*^2+0.35w_*^2 \tag{3.79}$$

with $c_o=3.3$ [see (Duynkerke and Driedonks, 1987)]. (In HIRLAM a slightly more complicated boundary condition is used, including some more stability effects.) In neutral conditions, and taking the limit to the surface, one gets as a balance between shear production and dissipation:

$$\frac{u_*^4}{K_m}=c_d\frac{E^{3/2}}{l_m} \tag{3.80}$$

With $K_m=l_m\sqrt{E}$ and $E=E_{surf}=c_o u_*^2$ one obtains $c_d=c_o^{-2}$.
We define a neutral constant $c_n\equiv c_d^{1/4}=c_o^{-1/2}$. It can be shown that with $l_m=c_n\kappa z\simeq 0.5\kappa z$ the neutral limit near the surface (i.e. $K_m=\kappa u_* z$, where u_* is surface friction velocity) is obtained.

3.5.3 The length scale

As mentioned above, the Bougeault and Lacarrère (1989) (hereafter B&L) length scale employed in CBR is changed to the new mixing length scale formulation that is presented below. This length scale combines a length scale for stable conditions l_s with an integral length scale for unstable conditions l_{int}. The length scale is computed by:

$$\frac{1}{l_{m,h}}=\frac{1}{\max(l_{int},l_{min})}+\frac{1}{l_s}, \tag{3.81}$$

where l_{min} is introduced to enable a smooth transition between stable and unstable condition (e.g, the averaging operator would yield a zero length scale for stable conditions when l_{int} is zero). It is defined by

$$\frac{1}{l_{min}}=\frac{1}{l_{limit}}+\frac{1}{c_n\kappa z}. \tag{3.82}$$

with $l_{limit}\simeq 50-100$m. Near the surface $l_{min}=c_n\kappa z$ (respecting the neutral limit). The above form has shown to give numerically stable and spatially continuous results.

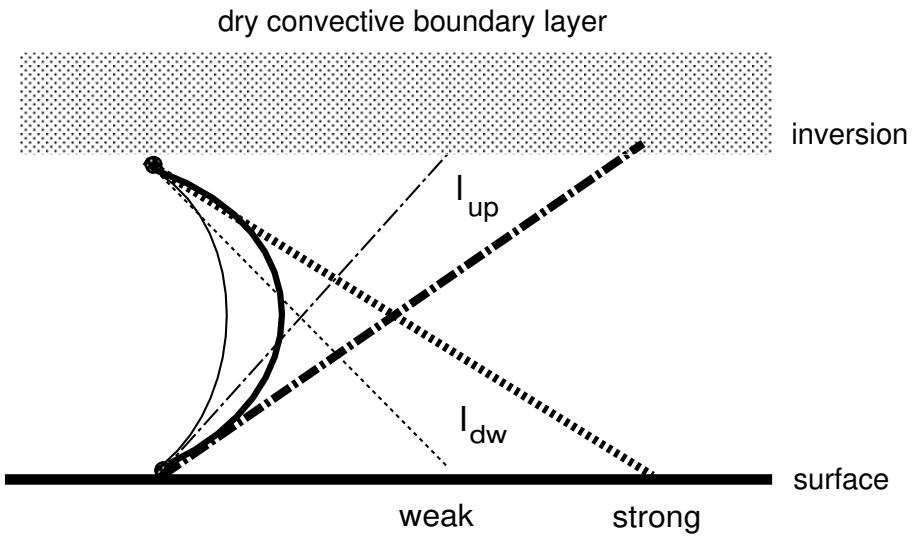


Figure 3.1: Illustration of the computation of the new length scale. Two "parcels" are released: one starts at the inversion, the other at the surface. The length scale increases as a function of the stability. Thick lines correspond to a convective, thin lines to a more neutral situation.

Integral length scale for unstable conditions

The length scale l_{int} (computed from integrals over stability) is designed to represent mixing in the range from close to neutral to unstable conditions. It is defined by an "averaging" over two length scales l_{up} and l_{dw} by

$$\frac{1}{l_{int}} = \frac{1}{l_{up}} + \frac{1}{l_{dw}} \tag{3.83}$$

These two length scale are defined as integrals over stability by:

$$l_{up} = \int_{z_{bottom}}^{z} F(\text{Stab}) dz'$$

$$l_{dw} = \int_{z}^{z_{top}} F(\text{Stab}) dz' \tag{3.84}$$

where F(Stab) is a function of a stability measure, and z_{bottom} and z_{top} are the lower and upper boundary of the mixing domain. The stability measure in present versions of the HIRLAM model is the Brunt-Vaisala frequency N^2, so (like the B&L length scale in the original CBR) based on buoyancy effects only.

The behavior of the length scale and its connection to the B&L length scale is illustrated below. In Fig. 3.1 a convective boundary layer capped by a strong inversion is shown. In this case z_{top} is the inversion height, z_{bottom} the surface. If we assume (just for illustrating purposes) that F is a constant, l_{up} and l_{dw} are proportional to the distance to the inversion and surface, respectively. If it is a strong inversion, the length scale is now very similar to the B&L length scale because in that case all upward parcels will stop just above the base of the inversion and the downward parcels will hit the surface. In this respect our method can be considered as a "poor man's" parcel method, obtaining rather similar results for convective situations to B&L, though at a much lower computational cost.

In our method F is a function of Brunt-Vaisala frequency N^2 ($= \frac{g}{\theta_v}\frac{\partial \theta_v}{\partial z}$). In Fig. 3.1 two situations are shown, one strongly convective representative for a midday situation, one only weakly convective representative for the late afternoon. In these two situations, the behavior of l_{up}, l_{dw} and l_{int} is shown. So roughly, F becomes larger with increasing instability; for the near neutral conditions at late afternoon F should represent neutral scaling, whereas for more convective conditions F is chosen larger.

The following form F is chosen:

$$\begin{aligned} F_m(Ri) &= \alpha_n - \alpha_r \alpha_d N^2 (1 + c\alpha_d N^2), \quad \text{stable} \\ &= \alpha_n - \alpha_r \alpha_d \arctan(\alpha_d N^2), \quad \text{unstable} \end{aligned}$$

with α_n chosen such that the resulting length scale respects the neutral limit, and α_r and α_d tuned so that the stability dependency near neutral and the convective limit is reasonable. The *arctan* is chosen because it levels off for high instability, allowing a reasonable convective limit. The constant c is between $0-1$, and when set to 1 causes a more rapid decrease of the length scale in stable conditions. It turned out that, because F is based on buoyancy only, it was necessary to include a correction due to wind shear in the length scale formulation. This correction decreases mixing for high wind shear convective situations.

Length scale for stable conditions

The length scale for stable conditions:

$$l_s = c_{m,h} \frac{\sqrt{E}}{N} \tag{3.85}$$

with $c_{m,h}$ a constant not necessarily the same for heat and momentum. In HIRLAM $c_h = 0.2$ is the standard value, and $c_m = 4c_h$, which implies a Prandtl number $Pr = 4$ for the (very) stable limit. This is done to represent the enhanced mixing of momentum by wave activity for (very) stable conditions. For example, data presented by Kim and Mahrt (1992) and Schumann and Gertz (1995) suggest a dependency of the Pr about 2-4 Ri. This dependency increases mixing under very stable conditions, which appears to be important for the forecast skill of AGCMs (Beljaars and Viterbo, 1998; Delage, 1997). On the other hand, other measurements (Nieuwstadt, 1984) and flux profile relations (Dyer, 1974) imply an almost constant Prandtl number.

3.5.4 Numerics

The stability parameters and the eddy diffusivities are all evaluated explicitly based on fields entering the CGL/CBR scheme. Ideally, these fields therefore should be based on balanced "full time level" fields, see (Lenderink and Holtslag, 2000). The diffusion solver is implicit, but uses explicit eddy diffusivities. Additional tendencies and surface fluxes (explicitly computed in the surface flux routine) are included in the diffusion solver. The diffusion solver may also be used to compute the surface momentum flux implicitly by assuming a no-slip (zero velocity) boundary condition at the surface, and prescribing an surface exchange coefficient. (The exchange coefficients can be easily diagnosed from the surface stress entering the scheme and the lowest level velocity.) This is the preferred boundary condition for momentum since the explicit flux boundary condition turned out to give rise to instability in operational forecasts.

In order to get a realistic balance between turbulent diffusion and other (explicit) processes, like dynamics and radiation, the diffusion solver should either use these tendencies in the diffusion solver, or the diffusion should work on the updated fields (partial time stepping procedure).

Details on the time stepping procedure can be found in the appendix of Lenderink and Holtslag (2000).

Because TKE is used to compute fluxes at the half levels, it was decided to change TKE levels from full levels in CBR to half levels in CGL. This increases the numerical accuracy inside the TKE scheme significantly, in particular near inversions. Outside the turbulence scheme no adaptations are made, and no attempts are made to interpolate back to the full levels when leaving the turbulence schemes. It was considered that the numerical errors associated with these interpolations are much more harmful, than the errors associated with advecting TKE with the "wrong" (full instead of half level) winds outside the scheme. (Thereby putting the emphasis on the local balance of TKE, and suffering slightly on the non-local advective terms, which are considered to be less important at the present horizontal resolution.)

3.5.5 Future developments

The present versions of CBR/CGL all compute diffusion of liquid water, together with diffusion of the dry variables. Because diffusion is a linear operator, this adds up to diffusion on moist variables, like total water and liquid water potential temperature. In this sense the scheme is a moist scheme. However, the length scales (in all versions up to now) are based on dry processes only, and do not include cloud condensation effects. The consequence is that mixing in stratiform is seriously underestimated. The scheme will therefore also underestimate entrainment at the cloud top, and will give rise to shallow moist boundary layers. A complete moist version of the scheme, including moist effect on the stability parameters, is in preparation.

The present form of the stability dependency in l_{int} is based on rather pragmatic arguments, only satisfying several limits (neutral, convective, and stable) but does not guarantee an optimal transition between these regimes. To enable a realistic dependency over a large stability range, the stability parameter may be changed from buoyancy to the Richardson number. This enables a matching to surface flux profile relations, and removes the arbitrariness of the present formulation. A manuscript describing this is in preparation (Lenderink, 2002).

3.6 Surface and Soil Processes

The HIRLAM surface scheme makes use of the so-called *mosaic of tiles* or *aggregation of fluxes* approach, first introduced by Avissar and Pielke (1989), and adopted by Claussen (1991); Koster and Suarez (1992). The primary motivation for the tiling approach is to promote suitable balance enhancements in horizontal complexity and to increase the physical realism of modelled surface energy and water fluxes. Such approach represents the surface heterogeneity within one grid element assuming the existence of different land-use patches which evolve independently and couple directly to the atmosphere of the model. The land-use patches within one grid square affect each other only through the atmosphere. Because the dependence of surface fluxes on land characteristics is non-linear, estimates of the area averaged fluxes calculated with mean land characteristics (*aggregation of parameters*) do not yield the same results as those obtained by aggregation of fluxes (Rodriguez-Camino and Avissar, 1999). Even more, the aggregation of parameters approach to surface heterogeneity can fail in certain cases (Stössel and Claussen, 1993; Blyth et al., 1993) when surface conditions vary strongly. This is the case of terrain borders, such as those between ice sheets and open sea or on coasts. The surface fluxes calculated in the new HIRLAM surface treatment has proven to compare well to NOPEX measurements (Bringfelt et al., 1999).

In principle, the scheme may treat an arbitrary number of surface types in each grid square (Bringfelt, 1996). Several global and local physiographic databases were merged to assign a

geographically- dominating vegetation type and soil texture for each subgrid surface (see the corresponding Section). At present, five surface types are considered within each grid square: sea/lake water, ice, bare land, forests and agricultural terrain/low vegetation. However, the code maintains a high flexibility to allow easy implementation of more complex tiling approaches in the future if required.

For sea/lake, there is no energy balance equation for calculating the water temperature: this quantity is kept constant. The surface roughness of the water is calculated using the Charnock formula (Charnock, 1955), but with modifications at low wind speeds and near coasts, see the preceding section on surface fluxes. For ice, a three layer model for temperature is used based on the one-dimensional equation for vertical heat diffusion with constant heat capacity and diffusivity in time and space (Källén, 1996).

For the three land-surface types the ISBA scheme was chosen (Noilhan and Planton (1989), identified as NP89 from here onwards; Noilhan and Mahfouf (1996)), because it has been tested extensively against laboratory and field-data for a number of land-use and soil types. Soil temperature and soil water content are treated with force-restore models. The soil is divided in two layers: one surface layer(d_1), with a depth typically of 1 cm, that responds to the diurnal cycle, and a total layer (d_2) extending down to a depth of about 1 m following a time scale of some days. There are prognostic equations for the surface temperature (including the vegetation canopy), T_s, its periodical mean value, T_2, surface soil water content, w_s, total soil water content, w_2, and rain (dew) water retained on the vegetation canopy, w_r. The corresponding equations are:

$$\frac{\partial T_s}{\partial t} = C_t G - \frac{2\pi}{\tau}(T_s - T_2) \tag{3.86}$$

$$\frac{\partial T_2}{\partial t} = \frac{T_s - T_2}{\tau} \tag{3.87}$$

$$\frac{\partial w_s}{\partial t} = \frac{C_1}{\rho_w d_1}(P_g - E_g) - \frac{C_2}{\tau}(w_s - w_{geq}) \tag{3.88}$$

$$\frac{\partial w_2}{\partial t} = \frac{1}{\rho_w d_2}(P_g - E_g - E_{tr}) - \frac{C_3}{\tau} max[0, (w_2 - w_{fc})] \tag{3.89}$$

$$\frac{\partial w_r}{\partial t} = veg \cdot P - E_r,\ w_r \leq w_{rmax} \tag{3.90}$$

The thermal coefficient C_t is expressed by $C_t = 1/(\frac{1-veg}{C_g} + \frac{veg}{C_v})$, where C_v and C_g are the contributions of the vegetated (veg) and bare ground fractions ($1 - veg$), respectively. C_g depends on soil texture and soil water content (NP89). Strong limitations to its maximum value have been imposed compared to the original formulation to prevent excessive diurnal cycle in screen variables. C_v takes a constant value for every type of vegetation which has been tuned. G is the sum of the fluxes at the surface in the soil-vegetation medium: net radiation flux (R_n), sensible heat flux (H), latent heat flux ($L_v E$) and flux due to the soil water change of phase (F_i). F_i can be expressed by $L_i \rho_w (1 - veg) d_2 \partial w_i / \partial t$, where ρ_w is water density, L_i is the latent heat of fusion and $\partial w_i / \partial t$ is the variation of the total ice water content. Following the ideas of Viterbo et al. (1999), the total ice water content can be assumed to be $w_i = f(T_s) w_2$, where $f(T_s)$ is a function taking the value 1 for temperatures well below $0^o C$ (all soil water content is in solid phase), 0 for temperatures well above $0^o C$ (all soil water content is in liquid phase) and with some smooth transition around $0^o C$. To avoid undesirable coupling between the temperature and water equations w_2 is additionally assumed to be equal to the field capacity

value, w_{fc}, in the expression for F_i. The final expression for the flux due to the soil water change of phase is: $F_i = [L\rho_w(1-veg)d_2 w_{fc} df(T_s)/dT_s]\partial T_s/\partial t$. This term can be included in the left hand side term of the surface temperature equation and incorporates the barrier effect through the pulse-like function of surface temperature, $df(T_s)/dT_s$, to simulate the soil water content freezing/thawing around 0^oC. The formulation of the hydric coefficient C_1 is similar to NP89, with the improvements of Braud et al. (1993); Giard and Bazile (2000). C_2 depends on soil water content and soil texture and follows the original formulation. C_3 is the relaxation coefficient of total water content to the field capacity (w_{fc}), to simulate the gravitational drainage. It depends only on soil texture (Mahfouf and Noilhan, 1996). For the soil water restore term, the ISBA model uses the equilibrium soil water, w_{geq}, (when gravity balances capillarity forces) instead of total soil water content, w_2, although both soil water contents are quite close (except for coarse soil texture (sand)). τ is a time-restore constant of one day. P_g is the precipitation reaching the soil. P is the precipitation at the top of the vegetation. E_g is the evaporation rate from the ground, which dries the shallow layer only. E_{tr} is the canopy transpiration flux from vegetation, which acts through the plant roots. E_r is the evaporation flux from the fraction δ of the foliage covered by intercepted water and w_{rmax} is the threshold value of the interception reservoir.

Snow sub-fractions are allowed within the ice fraction and within each of the land-surface fractions. Snow is represented by a single layer with a prognostic equation for snow depth. Two options are additionally available for the snow treatment: one modifying the thermal coefficient and albedo over snow covered fractions, and another including additional equations for snow density and snow albedo to simulate the ageing process of the snow (Douville et al., 1995) The snow parameterization is complemented in both options with a 6 hourly snow depth analysis based on a successive correction method making use of in situobservations.

In each time step, the sensible (H_i) and latent (E_i) heat fluxes from each subgrid surface are weighted according to their fractional share of the grid square to form the total surface fluxes. The aggregated fluxes were used at the lowest model level as lower boundary condition for the boundary layer vertical diffusion and the radiation schemes. At this level the wind, temperature and humidity were kept constant within the whole grid square. The surface fluxes H_i and E_i are based on the differences in temperature and humidity, between the lowest model level and the surface values for each sub-surface type, respectively. The atmospheric stability functions used for the computation of the aerodynamic resistance are given in Louis (1979) and Louis et al. (1982) (see a previous section) The vegetation surface resistance for transpiration includes the effect of the solar radiation, water stress in the root zone, water vapour deficit and air temperature, following NP89.

For the calculation of H_i and E_i the local roughness length due to vegetation is used. The weighted momentum flux is calculated using the orographic roughness length, $z_{0_{orog}}$, and an aggregated local vegetation roughness length, $z_{0_{veg}}$. $z_{0_{veg}}$ is obtained by averaging from the local roughness lengths for the subsurfaces, z_{0_i}, according to Mason (1988): $1/(ln\frac{l_b}{z_{0_{veg}}})^2 = \sum f_i/(ln\frac{l_b}{z_{0_i}})^2$, where l_b is the blending height or the height at which the flow becomes approximately independent of horizontal position (Claussen, 1991). In reality, the blending height varies according to the patchiness scale and local stratification. Assuming that the blending height is of the order of 1/100 of the horizontal scale of the roughness variations (Claussen, 1995), for typical landscape variations of 1000 m the blending height is of the order of 10 m, well below the lowest model level. Diagnostic fields, such as 2m temperature and relative humidity, 10m wind, etc., are also computed and are available for each grid box fraction separately.

The ISBA scheme for land-surface parameterization depends very strongly on the initial soil water content, see, e.g. Rodriguez-Camino and Avissar (1998). Consequently, some kind of soil water content assimilation is needed for its use in an operational environment. HIRLAM has

adopted the approach of Mahfouf (1991) for initialising the soil water content. He proposed an optimal interpolation scheme relating the increments of soil water to short-range forecast errors of 2m temperature and relative humidity. In bare ground areas, the error in 2m temperature and relative humidity is associated to error in the surface soil layer, whereas in vegetated areas it is related to the root layer soil water content error. Two aspects are critical in this method. The first is the definition of the optimum coefficients in the matrix relating the errors in 2m temperature and relative humidity to the analysis increments in the surface and root layer soil water content. The second is the elimination of systematic errors in 2m temperature and relative humidity not due to incorrect soil water. The optimum coefficients relating errors originally proposed by Bouttier et al. (1993) were further smoothed and adapted by Giard and Bazile (2000). Big jumps and daily oscillations in soil water analysis are prevented by filtering of soil water increments computed at 6 hours cycles. Systematic errors can recursively be computed and eliminated before the soil water analysis (Giard and Bazile, 2000). This method for initialising soil water content is only applied when 2m temperature and relative humidity is mainly forced by soil water in the ground. Therefore, in cases of ice or snow cover, frozen soil water content, dew, significant precipitation, strong low level winds, no evapotranspiration or polar night, the analysis of soil water content is switched off. A crude correction to T_s and T_2 is also carried out, proportional to the error in 2m temperature.

The accurate analysis of 2m temperature and relative humidity is the necessary first step to assimilate soil variables. The analysis of 2m temperature and relative humidity developed in HIRLAM (Navascues, 1997) is based on optimal interpolation. Observations of 2m temperature and relative humidity are usually of very local nature as they depend critically on surface characteristics. To palliate such local effects, a correction to the observation increments was introduced taking into account the vertical distance between the station height and the model interpolated orography. Such reduction is specially beneficial in the case of complex terrain. Furthermore, an anisotropy of the structure function may be used in the optimal interpolation, that includes the effect of coastline and inland topography.

3.7 Under development: Parametrization of the mesoscale orography effects

The mesoscale orography parametrization scheme of HIRLAM is adapted from the Arpege-Aladin forecast system. Documentation of the original scheme can be found in Cordeneanu and Geleyn (1998) and in the Arpege technical documentation (Geleyn, 2000). A more detailed description of its HIRLAM installation is given in a HIRLAM technical report by (Rontu et al., 2002). The parametrization includes the generation and dissipation of vertically propagating buoyancy waves, resonance effects and blocked-flow drag. The original scheme includes also a parametrization of lift effect, relevant for models with coarser horizontal resolution than HIRLAM, plus simple parametrizations of convective gravity wave drag and mesospheric drag.

The tendency of the horizontal wind $\vec{v}(x, y, z)$ in HIRLAM consists of contributions from explicitly resolved dynamics and parametrized momentum fluxes due to turbulence and mesoscale orography:

$$\frac{\partial \vec{v}}{\partial t} = (\frac{\partial \vec{v}}{\partial t})_d + (\frac{\partial \vec{v}}{\partial t})_t + (\frac{\partial \vec{v}}{\partial t})_m \tag{3.91}$$

where the indexes d, t and m refer to dynamics, turbulence and mesoscale orography effects, respectively. By definition, the turbulent and orographic tendencies are related to vertical divergence of corresponding subgrid-scale momentum fluxes or turbulent and orographic stresses

(defined here as forces exerted by the surface on the flow):

$$(\frac{\partial \vec{v}}{\partial t})_t = -\frac{1}{\rho}\frac{\partial \vec{\tau}_t}{\partial z}, \vec{\tau}_t = \rho(\overline{\vec{v}' w'})_t \tag{3.92}$$

$$(\frac{\partial \vec{v}}{\partial t})_m = -\frac{1}{\rho}\frac{\partial \vec{\tau}_m}{\partial z}, \vec{\tau}_m = \rho(\overline{\vec{v}' w'})_m, \tag{3.93}$$

where ρ is the air density, w is the vertical velocity, a prime $'$ denotes subgrid-scale deviation and an overline gridbox average. The stress, $\vec{\tau} = \vec{\tau}_m = (\tau_{mx}, \tau_{my})$, due to mesoscale mountain effects is represented by the scheme in the following way:

$$\vec{\tau}(x, y, z) = \gamma(x, y, z) \cdot \vec{\tau}_s(x, y) \tag{3.94}$$

where $\vec{\tau}_s$ represents the generation of orographic stress at the surface and the parameter γ determines the vertical distribution of the momentum flux (orographic stress).

In this parametrization scheme the surface value of the orographic stress is first calculated. Then the contribution of wave breaking and resonance is estimated and the parameter γ modified accordingly. The effect of flow blocking, although physically not directly related to the mountain waves, is taken into account by modifying the parameter γ at levels below the estimated blocking height.

3.7.1 Parametrization of mountain waves

Generation of the wave stress

The generation of wave stress $\vec{\tau}_s$ [Pa] at the surface is calculated according to:

$$\vec{\tau}_s(x, y) = K_g \cdot \rho_s \cdot N_s \cdot \vec{v}_{fs} \cdot h_m^2, \tag{3.95}$$

where the index s refers to the effective near-surface values (see below), K_g $[m^{-1}]$ is a tuning parameter depending on the model resolution (see Table 3.1), ρ_s $[kgm^{-3}]$ is the surface density, N_s $[s^{-1}]$ is an effective buoyancy frequency, $N^2 = \frac{g}{\theta}\frac{\partial \theta}{\partial z}$, $\vec{v}_{fs}$ $[ms^{-1}]$ is the fictive surface wind and h_m $[m]$ is the subgrid-scale mountain height given by

$$h_m = K_h \cdot \sigma_h, \tag{3.96}$$

where σ_h is the standard deviation of orography and the dimensionless parameter K_h (see Table 3.1) is used to estimate the distance between valleys and mountain tops in a gridbox (Lott and Miller, 1997).

So-called effective values are used in order to avoid a dependency on vertical discretization. Thus, the surface parameters used in Eq. (3.95) are calculated by vertical integration with linearly decreasing influence on the surface value over a certain depth, which is defined to be from the model surface to the height of the subgrid-scale mountains (h_m). In addition, to account for the non-isotropic properties of the subgrid-scale orography (see Ch. 3.7.4) a fictive surface wind $\vec{v}_s$ parallel to the surface orographic stress is calculated. This calculation is based on the effective surface wind, the anisotropy of the orography α (Eq. 3.120) and the direction of the principal axis of the subgrid-scale obstacles with respect to the model grid θ (Eq. 3.121):

$$u_{fs} = Au_s + D(u_s cos2\theta + v_s sin2\theta) \tag{3.97}$$

$$v_{fs} = Av_s + D(u_s sin2\theta - v_s cos2\theta), \tag{3.98}$$

The angle between the wind and orography is implicitly included into these formulae (J.F.Geleyn, 2002, personal communication). The coefficients A and D result from analytical fits of elliptic integrals for the anisotropic case (Arpege technical documentation):

$$A = \alpha^2 + \frac{4(1-\alpha)(1+a_1\alpha) - \alpha(1+a_2)ln\alpha}{2\pi} \tag{3.99}$$

$$D = \frac{4(1-\alpha)(1+b_1\alpha) + 3\alpha(1+b_2)ln\alpha}{2\pi}, \tag{3.100}$$

with $a_1 \approx 1.44$, $a_2 \approx 0.22$, $b_1 \approx 0.67$ and $b_2 \approx 0.44$. The fictive surface wind velocity has a maximum value when the flow is directed accross a two-dimensional ridge or over a circular mountain. The minum velocity is obtained for the flow along a two-dimensional ridge.

Wave breaking

As long as there is no interaction between the vertically propagating buoyancy waves and the surroundings, the wave momentum flux is constant with height, $\gamma = 1$ in Eq. (3.94). In order to maintain the constant momentum flux the wave amplitude must increase as the density decreases upwards. At a certain level waves start to break due to convective or shear instability. A drag is exerted on the flow since wave energy is converted into small-scale turbulent motions. The momentum flux vanishes at a critical level, where the mean velocity in the direction of wave motion becomes zero.

The parametrization of wave breaking processes follows Lindzen's saturation theory (Lindzen, 1981). Breaking is assumed when the local Richardson number approaches unity,

$$Ri = \frac{N^2}{|\frac{\partial \vec{v}}{\partial z}|^2} = 1 \tag{3.101}$$

The wind shear experienced by a wave of horizontal wavelength λ and amplitude A when meeting the large-scale flow can be approximated by

$$|\frac{\partial \vec{v}}{\partial z}| \approx U/A, \tag{3.102}$$

where U is the mean velocity in the direction of the wave motion,

$$U = \frac{\vec{v} \cdot \vec{v}_{fs}}{|\vec{v}_{fs}|}. \tag{3.103}$$

It can be shown that the momentum flux associated with a stationary hydrostatic buoyancy wave can be approximated by:

$$\rho(\overline{u'w'})_m = -\frac{\rho N U A^2}{2\lambda}, \tag{3.104}$$

where ρ is the density and N is the buoyancy frequency. An equation for the saturation stress for stationary mountain waves is now obtained replacing A by using Eqs. (3.101 - 3.102):

$$\tau_{sat} = \rho(\overline{u'w'})_{m,sat} = -\frac{\rho U^3}{2\lambda N}. \tag{3.105}$$

The level of wave breaking (saturation) is defined as a level where the magnitude of the saturation wave stress becomes less than the corresponding (maximum) value at the surface. Assuming that the saturation stress has the same expression at the surface as at the upper levels, this condition can be expressed as:

$$\gamma = \frac{\rho}{N}\frac{N_s}{\rho_s}(\frac{U}{U_s})^3 < 1 \tag{3.106}$$

where the index s denotes surface values. The horizontal wavelength λ can be ignored in this ratio since it is fairly constant with height.

An idealised wave stress profile related to wave breaking in the present scheme is shown in Fig. 3.2.

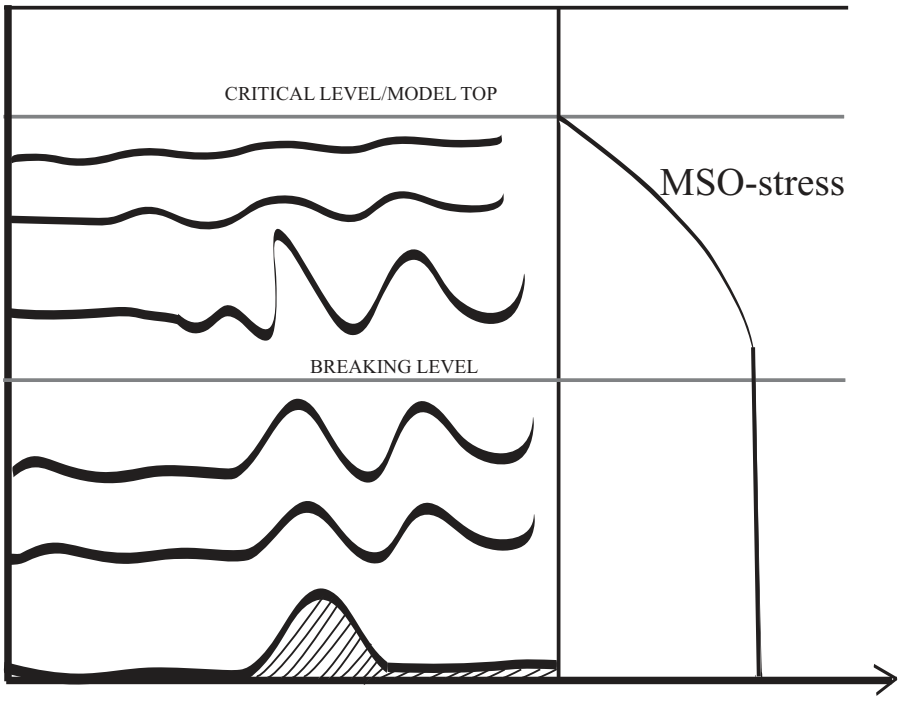


Figure 3.2: *Influence on the orographic drag of the generation and breaking of mountain waves. The parametrized MSO-stress remains constant in height until the waves start to break.*

Wave reflection

Wave reflection is a process that may modify the amplitude of the waves as well. As discussed in Peltier and Clark (1979) this process is strongly coupled to the region of wave breaking. Apparently, the upward propagating hydrostatic mountain waves are reflected from a breaking level. Under the right conditions the downward propagating wave is in-phase with the upward component and a constructive interference occurs. As a result of this, an increased wave drag is produced at low levels. It has been shown that the downward wave will be in-phase when the height of the breaking layer (z_{br}) is 75 % of the vertical wave length of the waves (λ_z). Not only this height is a reflecting layer but also a second level of self-induced resonance exists (Clark and Peltier, 1984). The condition for breaking level resonance can be expressed as

$$\frac{z_{br}}{\lambda_z} = \frac{3}{4} + n \tag{3.107}$$

or, in terms of phase angle θ,

$$\theta = mz_{br} = \frac{3\pi}{2} + 2\pi n \tag{3.108}$$

where n is a full and positive integer and m $= \frac{2\pi}{\lambda_z}$ is the vertical wave number. Correspondingly, maximum destruction is expected when the upward and downward propagating waves are completely out of phase, $\theta = \frac{\pi}{2} + 2\pi n$.

In the parametrization, the height of the breaking level is estimated according to Eq. (3.106). The phase angle is obtained by integration from the surface to breaking level:

$$\theta = \int_0^{z_{br}} m dz = \int_0^{z_{br}} \frac{N}{U} dz \tag{3.109}$$

where the vertical wave number has been approximated with $m = \frac{N}{U}$, an assumption valid for vertically propagating hydrostatic waves with horizontal wavelengths of about 10 km and above.

A periodic function $f(\theta)$ is defined for the resonance calculations with a maximum for $\theta = \frac{3\pi}{2}$ (amplification) and minimum for $\theta = \frac{\pi}{2}$ (destruction):

$$f(\theta) = \frac{1}{\sqrt{(1 + 2K_a sin(\theta) + K_a^2)}} \tag{3.110}$$

where an empirical parameter K_a (see Table 3.1) determines the amplitude of this function. For maximum resonance the amplification factor is 2.5 while for maximum destruction it is 0.625. When $\theta = 0$ or $\theta = 2\pi$ no reflection occurs since nodes are present at the boundaries (at surface and the breaking level). Thus the values given by the amplification function are ignored for these endpoints. Furthermore, it is assumed that the increased drag due to amplification is distributed in a linear fashion between the surface and breaking level. Wave resonance effects modify the parameter γ (Eq. (3.94)). The influence of wave reflection on the stress profile is sketched in Fig. 3.3a.

Note that the dynamics of horizontally propagating (trapped) lee waves (Scorer, 1949; Bretherton, 1969) is not handled separately by the present parametrization scheme. These waves are reflected from discontinuities of wind or stability in the lower troposphere and dissipated by turbulence downstream from the obstacle.

Critical level modifications

Close to a critical level the wave momentum flux approaches zero. In the parametrization a critical level is defined if $U = 0$, or also at the model top, where by definition $\rho = 0$. The parameter γ is set to zero at that level and above as the wave has been completely absorbed. Special treatment is needed when the stability becomes neutral or unstable: γ is set to zero also at a level with $N^2 \leq 0$ and above. Below this level γ is modified linearly as shown in Fig. 3.3b.

3.7.2 Blocked-flow drag

Usually, when a flow is approaching an obstacle not all of it can pass over it. Below some level the flow is instead forced to go around the object or stay blocked in valleys between the mountain ridges. This effect, which results in low level flow retardation, is referred to as flow blocking and the related drag is characterised as form drag. It is parametrized here following Lott and Miller (1997).

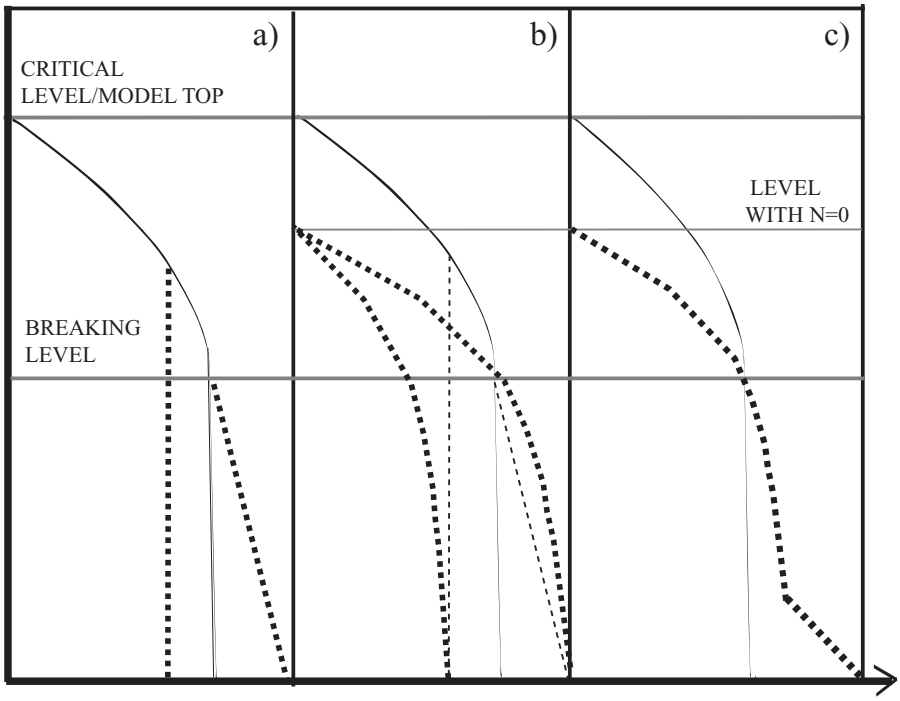


Figure 3.3: *Distribution of the parametrized MSO-stress in the case of wave reflection from the breaking level (a) and when an unstable layer with $N = 0$ is met (b). Thick dashed lines in (a) show the modification of wave stress in the case of maximum destruction and maximum amplification. In (b) the thick dashed lines show the linear modification of these stress profiles below the level $N = 0$. Thin dashed lines are a copy of the unmodified lines in (a). In (c) the combined effect of wave resonance, occurrence of an unstable layer above the breaking level and low-level blocking is sketched with the dashed line.*

A non-dimensional mountain height G is defined as:

$$G = N_s \frac{h_m}{U_p} \tag{3.111}$$

where N_s is the effective buoyancy frequency and h_m the mountain height defined by Eq. (3.96). U_p is the velocity of the upstream wind component perpendicular to the ridge, defined here as

$$U_p = \frac{U_s^2}{|\vec{v}_{fs}|} \tag{3.112}$$

This definition takes into account the anisotropy of the obstacle and the wind direction with respect to it, because the fictive surface wind includes these effects (Eqs. 3.97 - 3.100). The parameter G defined by Eq. (3.111) has a maximum value (with a given stability) when the flow is directed across a two-dimensional mountain ridge or over a circular obstacle.

The flow is assumed blocked if G is larger than a critical value $G_c = 1/K_b$ (see Table 3.1). The height of the blocked layer h_b (the upstream height of a streamline dividing blocked flow from the flow passing over the mountain) is in this case (Hunt and Snyder, 1980)

$$h_b = h_m(1 - \frac{G_c}{G}), \tag{3.113}$$

In general, for each layer below the level $z = h_b$ a force per unit volume, the form drag $\vec{D}_b$ exerted by the obstacle on the flow, is defined as

$$\vec{D}_b(z) = -K_b l(z) \rho(z) \frac{\vec{v}|\vec{v}|}{2}. \tag{3.114}$$

where K_b is a constant form drag coefficient, $l(z)$ is the horizontal width of the obstacle seen by the flow at upstream height z above the surface (mean orography height $\overline{h}$), ρ is the air density and $\vec{v}$ is the upstream velocity. The width $l(z)$ depends on the mountain height, anisotropy, slope and wind direction in respect to the mountain ridge. Within the present scheme the slope is not defined and $l(z)$ is therefore estimated assuming an elliptical mountain profile. After some calculations (J.F. Geleyn, personal communication) we arrive to the expressions (3.115-3.117):

$$l(z) = \int_z^{h_b} \sqrt{\frac{h_b - z'}{z' + h_m}} dz' \approx \sqrt{\frac{(1 - z/h_b)^3}{1 + z/h_m}}. \tag{3.115}$$

The total stress τ_b due to the blocked flow could now be obtained by integrating Eq. (3.114) from the surface to the blocking height. In the model, the contribution from blocked flow drag to the tendencies of wind components is required at each model level separately. Therefore the blocked flow stress is combined with the wave stress at each level by multiplying the parameter γ below the blocking height by a coefficient $\alpha(z)$:

$$\alpha(z) = 1 + a_b \sqrt{\frac{(1 - z/h_b)^3}{1 + z/h_m}}, \tag{3.116}$$

where a nondimensional coefficient a_b is given by the combination of parameters

$$a_b = \frac{K_b |\vec{v}_{fs}| (1 - G_c/G)}{N_s h_m}. \tag{3.117}$$

The low-level flow blocking modifies the stress profile as shown in Fig. 3.3c. In this schematic figure an example of combined wave breaking, constructive reference and low-level blocking is depicted.

3.7.3 Coefficients of the scheme

In the parametrization scheme there are five different tuning parameters mentioned in previous chapters: K_g, K_h, K_a, K_c and K_b. The coefficients K_g and K_h influence all effects handled by the parametrization. K_g is related to the generation of orographic stress (Eq. 3.95). K_h defines the depth of the subgrid-scale mountains. It influences the generation of MSO stress (Eq. 3.95), the depth over which the effective surface parameters are calculated and the calculation of the non-dimensional mountain height (Eq. 3.111). The other three parameters influence only the process they are related to. The proposed values of the five coefficients are shown in Table 3.1. The value of the parameter K_g is good for a horizontal resolution of about 20 - 40 km.

3.7.4 Derivation of the orography variables for the MSO scheme

From a data base of orography height three types of variables are derived for the use in HIRLAM model: surface geopotential for the resolved model dynamics; orographic roughness length for the turbulence parametrizations; standard deviation and parameters describing anisotropy and

Table 3.1: Coefficients of the scheme

parameter	related to	proposed value
K_g	generation of orographic drag	$3.5 \cdot 10^{-6}$ m^{-1}
K_h	obstacle height	3
K_a	wave resonance	0.6
K_c	nondimensional mountain height	3
K_b	blocked-flow drag	6.0

average orientation of the meso-scale orography for the MSO parametrizations. Different filtering is applied on the source data when calculating these parameters. This is done in order to ensure that the parametrizations are not overlapping each other or the resolved dynamics.

The aggregation of the parameters related to mesoscale orography is done on a target grid, which has a lower resolution than the source grid, and which is rotated with respect to the source grid.

Definitions

There are three parameters representing mesoscale orography in the scheme. The first of them is the standard deviation of subgrid orography σ_h. The other two parameters describe the non-isotropic properties of sub-grid orography and are derived from the tensor of orographic gradient correlation (Lott and Miller, 1997):

$$H_{ij} = \overline{\frac{\partial h}{\partial x_i}\frac{\partial h}{\partial x_j}}. \tag{3.118}$$

h is the orography, as it is represented in the source grid, and x_i and x_j describe the coordinate axes of this grid. If we define

$$\begin{aligned} T &= \frac{\partial h}{\partial x_i}\frac{\partial h}{\partial x_i} + \frac{\partial h}{\partial x_j}\frac{\partial h}{\partial x_j} \\ D &= \sqrt{\left(\frac{\partial h}{\partial x_i}\frac{\partial h}{\partial x_i} - \frac{\partial h}{\partial x_j}\frac{\partial h}{\partial x_j}\right)^2 + 4\left(\frac{\partial h}{\partial x_i}\frac{\partial h}{\partial x_j}\right)^2}, \end{aligned} \tag{3.119}$$

then the coefficient of anisotropy is defined:

$$\alpha = \sqrt{\frac{\overline{|T-D|}}{\overline{|T+D|}}}. \tag{3.120}$$

The anisotropy value becomes unity for a circular (fully isotropic) mountain and zero for a two-dimensional ridge. The direction between the gradient of orography, as referred to the mean is defined:

$$\theta = \arctan\left(\frac{\frac{\partial h}{\partial x_j}\frac{\partial h}{\partial x_j} - \frac{\partial h}{\partial x_i}\frac{\partial h}{\partial x_i} + D}{2\frac{\partial h}{\partial x_i}\frac{\partial h}{\partial x_j}}\right) \tag{3.121}$$

Source grid aggregation and filtering

The aggregation of the parameters is done on a target grid (model grid), which has a lower resolution than the source grid, and which is rotated with respect to the source grid. It is performed on basis of a source orography, that is filtered in order to determine the subgrid-scale orography variables in a proper scale for the parametrization. Otherwise, the parametrization may simulate unrelated effects. The demands for the filter are, however, different from those from the model orography. Removal of both the higher and the lower spectrum seems appropriate here (Scinocca and McFarlane, 2000).

The smallest-scale orographic structures (with a horizontal scale below a few kilometres) create turbulence but no mountain waves (Queney, 1948). It is thus appropriate to filter out smaller scales in the orography field using a low-pass filter. This is done by aggregating the orography from GTOPO30" (USGS, 1998) onto a grid with approximately 2.7 km resolution. The grid is rotated at the same time in order to assure better homogeneity with respect to areal representation, which is a necessity for the determination of σ_h.

In addition, the HIRLAM dynamics can be expected to resolve those scales in the orography, which are resolved in the (properly filtered) target grid. The properties of the high-pass filter should preferably be consistent with the filter applied on the model orography, i. e. the filter should start to work where scales begin to be resolved by the model.

The following high-pass filter, which is based on a Gaussian characteristics, was found to fulfil the demands from above:

$$h = h_L - [h_L]. \tag{3.122}$$

h_L denotes the low-pass filtered orography and $[h_L]$ is defined

$$[h_L(x)] = \frac{\int_{-\infty}^{\infty} g(x-x')h_L(x')dx'}{\int_{-\infty}^{\infty} g(x-x')dx'}, \tag{3.123}$$

where

$$g(x-x') = exp\left(-\frac{2|x-x'|}{\delta}\right). \tag{3.124}$$

The scale parameter δ determines the area taken into account when determining $[h_L(x)]$. It is set to $\delta = 2\Delta x$, where Δx is the resolution of the target grid.

The response function of this band-pass filter, including the grid aggregation as well as Eq. (3.122), is shown in Fig. 3.4 together with the response of the Shapiro filter used for the surface geopotential of the resolved model orography. The lower x-axis denotes relative wave number with respect to the resolution of the low-pass filtered elevation Δh_L, and the upper x-axis denotes relative wave number with respect to the resolution of the target grid Δx. The spectrum shown in the chart extends from the long wave range at the left hand side to the short wave range up to $k\Delta h_L = 0.5$, corresponding to $2\Delta x$-waves in the h_L-grid, which represent the shortest possible waves in the band-pass filtered orography, and it extends further into the small scale spectrum with $k\Delta h_L > 0.5$, which is defined as leading to orographic turbulence. The figure depicts two filter transition zones. The first transition occurs between the resolved–unresolved wave spectrum at approximately $k\Delta x = 0.3$. This value corresponds to $3\Delta x$-waves with respect to the target grid. It is the transition between the Shapiro filter and the high-pass part of the band-pass filter. The second transition occurs between the mesoscale and the turbulent orography spectrum at $k\Delta h_L = 0.5$, where the low-pass part of the band-pass filter is active.

Its response is an idealized depiction of a transition, where Gibbs phenomena are expected to blur the transition. The spectral region occurring between these two transitions is denoted "unresolved mesoscale orography", and is supposed to be represented in the model via the MSO parametrization.

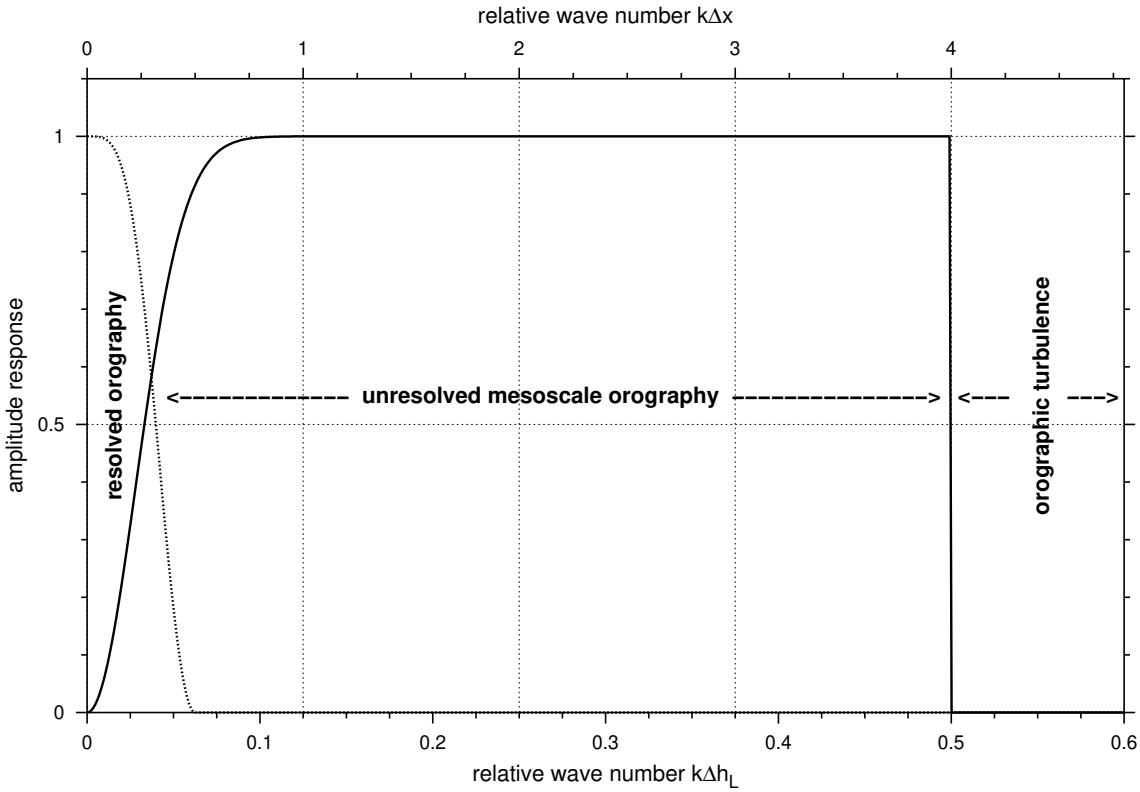


Figure 3.4: *Response function of the Shapiro filter (dotted curve) used to filter model orography (resolved orography), and the band-pass filter for mesoscale orography (solid curve) used within the MSO parameter aggregation. The latter consists of a low-pass filter using grid averaging on the Δh_L-grid, and a high-pass filter described by Eq. (3.122). Δx is chosen to $8\Delta h_L$ in the chart without loss of generality. See text for further details and a detailed description of the x-axis.*

Chapter 4

Data Assimilation and Initialization

4.1 HIRLAM Optimum Interpolation

4.1.1 Introduction

The HIRLAM Optimum Interpolation (OI) upper-air analysis is adapted from the previous operational global analysis scheme at ECMWF. It is a minimum variance scheme in terms of expected analysed interpolation errors, see e.g. Daley (1991). Several unique features were developed by Lorenc (1981) which greatly enhanced the performance of OI compared with earlier less optimal implementations. The main novel features were superobservation formation, quality control, not only with respect to the background, but also against other observations using OI in a modified mode (constrained not to use background information). The major reason for the success of this implementation is due to the box method, where the the OI equations are consistently applied inside boxes and the resulting analysis increments at gridpoints in each box are computed optimally.

The OI analysis relies a lot on the specified covariance statistics for background and observation errors. Those were considerably enhanced by Hollingsworth and Lönnberg (1986); Lönnberg and Hollingsworth (1986) and further tuned by Lönnberg (1988). Spatial correlations are modelled ussing series of Bessel functions which are more scale selective than the negative exponentials previously used. Magnitudes of background and observation errors have been tuned using spatial covariances of innovation vectors (observation minus background departures), chiefly for radiosondes over uniform areas. Some later tuning of magnitudes of of magnitudes of observation errors has also been done in the HIRLAM OI (Undén, 1999).

4.1.2 Optimal analysis weights

The OI method is a minimum variance estimate of increments of the state vector. In the mass and wind analysis it consists of surface pressure, height and wind components at the levels of the forecast model. The analysis increments (at a gridpoint and model level, commonly indexed k in the following) are defined as a linear combination of the of observed departures from the background (or innovation vector in the Kalman filter terminology, e.g. Daley (1991).

$$a_k = b_k + \sum_{i=1}^{N} w_{ki}(y_i - b_i) \tag{4.1}$$

where a_k is the analysis, b_k the background and w_{ki} the weights given to observations i=1,N. y_i are the observations and b_i the background values at those observation points, i. In the standard

OI methodology the analysis and background values in eq. (4.1) are transformed to departures from the "truth" by subtracting the true value at gridpoint, k, on both sides and subtracting and adding the true values at observation points i inside the summation to arrive at an equation for errors of analysis, background and observations (Lorenc (1981), eq. 3).

Note: In the equation above it is assumed that the background is perfectly available both at the grid-point (first term) and at the observation points (the second term on the R.H. side). The interpolation from background model grid to observation location is neglected in the statistical derivation and this is in common with all other known OI implementations. It is in contrast with standard Kalman filter formalism, see e.g. Daley (1991) and with 3D/4D-Var implementations, where the effect is implicitly catered for (Courtier et al., 1998).

The equation is then squared and an ensemble average is formed in order to use statistical quantities. The squared (normalised by background errors) analysis error is then:

$$(\epsilon_k^a)^2 = 1 - 2\mathbf{w}_\mathrm{k}^\mathrm{T}\mathbf{p}_\mathrm{k} + \mathbf{w}_\mathrm{k}^\mathrm{T}\mathbf{M}\mathbf{w}_\mathrm{k} \tag{4.2}$$

where $\mathbf{p}_\mathrm{k}$ are the background error correlations between all the observations, i, and the gridpoint k. It is assumed that observation and background errors are uncorrelated in (4.2). The matrix $\mathbf{M}$ represents the sum of the background error correlation matrix between observation points and the observation error covariance matrix normalised by the background error standard deviations at the two observation points. By minimizing this equation by differentiating $\partial(\epsilon_k^a)^2/\partial w_{ki}$ and setting it to zero for i=1,N one gets a set of linear equations for the weights $\mathbf{w}_\mathrm{k}$ at gridpoint k for all the observations:

$$\mathbf{w}_\mathrm{k} = \mathbf{M}^{-1}\mathbf{p}_\mathrm{k} \tag{4.3}$$

These weights describe the optimal combination of observation innovation vectors in (4.1) and the resulting analysis increments are statistically optimal in the least square sense.

Using the weights from (4.3) in (4.1) gives

$$a_k - b_k = \mathbf{M}^{-1}\mathbf{p}_\mathrm{k}(\mathbf{y} - \mathbf{b}) \tag{4.4}$$

or

$$a_k - b_k = (\mathbf{y} - \mathbf{b})^\mathbf{T}\mathbf{M}^{-1}\mathbf{p}_\mathrm{k} \tag{4.5}$$

Only the last factor, $\mathbf{p}_\mathrm{k}$, changes with the gridpoint position, k. The observation departures and the inverse of the correlation matrix between observation points are independent of the gridpoints and their product can thus be pre-calculated as the so called analysis coefficients:

$$\mathbf{C} = \mathbf{M}^{-1}(\mathbf{y} - \mathbf{b}) \tag{4.6}$$

and the analysis can then be computed at all gridpoints (and levels) in the analysis volume using the same vector of analysis coefficients:

$$a_k - b_k = \mathbf{C}^\mathbf{T}\mathbf{p}_\mathrm{k} \tag{4.7}$$

4.1.3 Analysis errors

The expression for the optimal weights $\mathbf{w}_\mathrm{k}$ in (4.3) can be substituted back in eq. (4.2) and a somewhat simpler equation for the optimal analysis error can be found:

$$(\epsilon_k^a)^2 = 1 - \mathbf{w}_\mathrm{k}^\mathrm{T}\mathbf{p}_\mathrm{k} \tag{4.8}$$

This equation is used to compute the estimated analysis errors on a reduced resolution each analysis cycle.

4.1.4 Superobservations

Lorenc (1981) designed a way of reducing data numbers for areas and observation types, where there are very dense and highly correlated sources of information. The OI analysis equations (4.1) and (4.3) can be used to analyse an optimal value of an observed quantity at an average position of a number of SYNOP or SHIP observations. The thus analysed observed quantity is the optimal combination of all the constituting observations and this value, together with its estimated analysis error, replace the constituting observations in all further processing. The important point is to make sure that the background values do not affect the superobservation; otherwise the assumptions for arriving at Eq. (4.2) are violated. Lorenc (1981) devised a modified statistical interpolation equation which used a constraint that the analysis and background errors should have zero correlation. By multiplying eq. (4.1) by the of background error and taking the ensemble average, one arrives at a condition on the weights (for the desired position, k):

$$\mathbf{w}_{\mathrm{k}}^{\mathrm{T}}\mathbf{p}_{\mathrm{k}} = 1 \tag{4.9}$$

Equation (4.2) for the interpolation errors is still valid and the equation is minimized as before but with the addition of the the constraint, eq. (4.9), multiplied by a Lagrange multiplier $2\lambda(1 - \mathbf{w}_{\mathrm{k}}^{\mathrm{T}}\mathbf{p}_{\mathrm{k}})$. This gives an equation for the constrained weigths:

$$\mathbf{w}_{\mathrm{k}} = (1 + \lambda)\mathbf{M}^{-1}\mathbf{p}_{\mathrm{k}} \tag{4.10}$$

The value of λ is found from substituting back in equation (4.9). One λ exists for each point or variable and the interpolation error is in fact identical to the value of λ, as can be seen from using (4.10) and (4.9) in (4.2).

Superobservations are formed from three or more closely located surface observations. That is, if they are inside one model grid box whose corners are the grid points of the HIRLAM forecast model. If at least three observations have pressures, a pressure superobservation is formed. Likewise, if at least three 10 m winds are present, a superobservations of wind is formed. The constituting data are subject to OI data checking (see below) and have to pass that before being considered for forming superobservations.

The superobservation is formed at the average position of the constituting observations and at average pressure. The data are linearly interpolated using eq. (4.1) above and assigned an observation error of λ. The superobservation is considered correct and excluded from further quality control. The constituting data are of course excluded from further processing, as their information has been transferred to the superobservation. The main purpose of this process is to reduce the data amounts in densely observed areas, since the covariance matrix between observation positions, $\mathbf{M}$, has a finite dimension and a should ideally contain all data in the area e.g. influencing one gridpoint.

4.1.5 OI data checking

Lorenc (1981) devised an optimal way of quality controlling data by estimating its value with the aid of surrounding data. The value is interpolated using a modified form of the OI interpolation equations. First, rather than minimizing the expected variance of analys error (with respect to the "true" value), the interpolation error is minimized:

$$< (a_k - y_k)^2 >= (\epsilon_k^o)^2 + 1 - 2\mathbf{w}_{\mathrm{k}}^{\mathrm{T}}\mathbf{m}_{\mathrm{k}} + \mathbf{w}_{\mathrm{k}}^{\mathrm{T}}\mathbf{M}\mathbf{w}_{\mathrm{k}} \tag{4.11}$$

It is obviously necessary to exclude the information from the datum being checked in this procedure; otherwise one would trivially check the datum against itself. Since the formation and inversion of the **M** matrix is rather expensive, the same matrix is re-used for a volume of data. This is done by adding a soft constraint that the weights of the datum being checked and possibly other already rejected data are set to zero. This condition can be written:

$$\mathbf{d}_l^{\mathrm{T}}\mathbf{w}_k = 0 \tag{4.12}$$

for each l, where l is the datum to be excluded and $\mathbf{d}_l$ is a vector whose l:th element is one and all other elements are zero. For a list of such excluded data (the datum itself and possibly other rejected data), m, it can be written, following Lorenc (1981):

$$\mathbf{D}^{\mathrm{T}}\mathbf{w}_k = 0 \tag{4.13}$$

where **D** is a matrix whose m:th colomn is $\mathbf{d}_l$. Minimizing equation (4.11) with this constraint gives an equation for the weights:

$$\mathbf{w}_k = \mathbf{d}_k + \mathbf{M}^{-1}\mathbf{D}\lambda \tag{4.14}$$

λ can be found by multiplying (4.14) by $\mathbf{D}^{\mathrm{T}}$ and using (4.12) to produce the modified weigths. Using these weights in (4.1) gives the final expression for the interpolated values at the observation positions. Also the expected interpolation error can be found from (4.11) above and using the two relations above to arrive at:

$$< (a_k - y_k)^2 >= (\epsilon_k^o)^2 + 1 - \mathbf{w}_k^{\mathrm{T}}\mathbf{m}_k \tag{4.15}$$

where $\mathbf{m}_k$ is a coloumn of **M**. If interpolated departures from the observation are greater than a certain multiple of the expected error variance in (4.15) (plus an empirical constant of 0.1 to allow for imperfect statistics employed), then the datum is considered to be in error and rejected.

The method is devised so that all data are scanned and the worst normalized departure considered first. If that one is bad enough for rejection, it is added to the list of data to receive zero weight (**D** in (4.12) above) and new interpolated departures (and expected errors) are computed, as the prior ones might be adversely affected by the erroneous data. The whole procedure is repeated until no furhter data exceed the rejection limit. In addition to rejected data, the list of zero-weight data may also contain so called blacklisted data. In the current implementation automatically blacklisted ships go through the checks in such passive mode; departures are calculated from other observations, but the blacklisted ships do not affect checking of any other data.

Although the OI data checking is rather sophisticated, the number of data rejected is small compared to the preceding screening against the background field.

4.1.6 Statistical model

The OI analysis equations depend on pre-determined forecast and observation error statistics in the covariance terms. The forecast errors are assumed to be (almost) separable, so that the covariance between two points can be expressed as the product between the horizontal and vertical correlations times the two standard deviations of error. The horizontal part is first described below, then the vertical. Observation errors are currently only assigned vertical correlations, no horizontal ones.

Forecast and observation error structures are based on Hollingsworth and Lönnberg (1986) and Lönnberg and Hollingsworth (1986) with further tuning by Lönnberg (1988). The horizontal structures are modelled with the aid of a series of Bessel functions of distance between observation points (r).

$$F(r) = \sum_{n=0}^{8} A_n J_0(k_n r/D) \tag{4.16}$$

where D is extent of the area and where a zero derivative condition is applied on F. This determines the values of k_n for each of the modes and A_n are the amplitudes of each mode, determined by the fit of the (historical) data.

A_0 determines the amplitude ot the horizontally constant (within the area D or in and around one analysis box). This is the Large Scale term and the total correlation is modelled as

$$F(r, \Delta p) = A_0 V_{LS} + \sum_{n=1}^{8} A_n J_0(k_n r/D) V_{SYN} \tag{4.17}$$

where V_{LS} and V_{SYN} are the vertical correlations for the Large Scale and Synoptic part of the forecast errors, respectively.

Thickness errors are derived from applying the statistical expectation operator on differences of heights and thereby using variances and correlations between height levels.

The wind errors are used for wind components, but the underlying model is more conveniently formulated in terms of stream function and velocity potential error, following Daley (1991). Both are modelled using a (common) series of Bessel functions, just as above for the height errors. The vertical parts of the correlations are however different for stream function and velocity potential. The error in wind is derived from Helmholtz' theorem:

$$\mathbf{v} = \nabla\chi + \mathbf{k} \times \nabla\psi \tag{4.18}$$

and the vind error variance is divided into a divergent part, νE_{v^2} and non-divergent part, $(1-\nu)E_{v^2}$ following Daley (1985); Undén (1989).

Stream function and velocity potential covariances are modelled to use the same horizontal correlations and using the same series of Bessel functions as for heights. Vertical correlations of velocity potential are however allowed to be different from height and stream function, see below.

The wind component correlations (in longitudinal i.e. projected on the line joining two points, and the orthogonal transverse directions) are derived by differentiating the stream function and velocity potential covariances using (4.18). This provides multi-variate coupling between wind components. For geostrophic coupling between height and stream function, this is achieved by modelling the coupling between two points, i and j, as

$$< \Phi_i \Psi_j >= \mu E_\Phi E_\Psi V F(r_{ij}) \tag{4.19}$$

and the factor μ is allowed to vary with horizontal mode in the Bessel function representation.

The vertical forecast error correlations are modelled as (see also Shaw et al. (1987):

$$V_{ij} = \frac{1}{e-1} e^{\frac{a}{a+(x_i - x_j)^2}} - \frac{1}{e-1} \tag{4.20}$$

For the divergent vertical correlations, a different formulation allowing negative lobes, is used:

$$V_{ij} = \left(\frac{1}{e-1} e^{\frac{a}{a+(x_i - x_j)^2}} - \frac{1}{e-1} \right) \cdot \frac{1 - 0.75(x_i - x_j)^2}{1 + 0.04(x_i - x_j)^2} \tag{4.21}$$

where x_i and x_j are values of a transformed pressure coordinate (from ln(p)) to give as good fit as possible to empirical correlations.

As both the standard deviations of forecast errors and the transformation x above, for the correlations, are expressed as polynomials of ln(p), the representation is continuous and observations and model levels are entered with their exact pressures.

4.1.7 Background and observation error statistics

Mean 6 hour first guess background error statistics have been computed from data according to Hollingsworth and Lönnberg (1986); Lönnberg and Hollingsworth (1986). The error magnitudes are modified by the analysis error computation according to equation (4.8). Then the forecast errors are assumed to grow during the cycling, and a simple and smooth model is used for this. First, a two-dimensional analysis error magnitude is computed by summing up the normalised analysis errors for 7 levels and three variables (height and u and v wind components and dividing by the number of levels and variables (21) involved. This is a mean error factor. Then a climatological error factor is computed from the specified climatological errors normalised by the mean 6 hour forecast errors, averaged in the same way for variables and levels.

Then the model of forecast error growth is:

$$e = e^a + \frac{\Delta t_{fc}}{\Delta t_{rand}} (\sqrt{2} e^c - e^a) \tag{4.22}$$

where a^a and e^c are the two-dimensional analysis and climatological errors creates as described above. The ratio of the two Δt:s control the error growth. Δt_{rand} is 6 days in midlatitudes and 2 days in the tropics with 6 hour cycling.

Finally, the 3-dimensional first guess background errors are created by multiplying the two-dimensional field by the the mean first guess errors at the appropriate levels. The error field is computed at a low resolution, 6 by 6° and 7 pressure levels due to the cost of calculating analysis errors in OI and it is a low precision estimate.

4.1.8 Data usage and quality control

The data variables from observations that enter the analysis are: 1) pressures at station heights converted to height at station pressure 2) heights at pressure levels from soundings 3) wind direction and speed converted to u- and v-wind components 4) temperature and dew points converted to relative humidity, 5) thickness between pressure levels and 6) precipitable water content from derived from satellite data (SATEMs are not used in all implementations). Temperatures could also in principle be used, but this option is not implemented in the HIRLAM OI analysis. There is also an option to convert cloud observations to pseudo humidities but this option is not activated either.

The data are extracted from the WMO observation types SYNOP, AIREP/AMDAR, SATOB, DRIFTER, TEMP, PILOT and SATEM as distributed on the Global Telecommunication System.

The actual analysis computations start by extracting the desired data from the observations and interpolating the background fields end error standard deviations to the observation positions

and pressures. Vertical interpolations are done at the surrounding gridpoints for *all* observations that need horizontal interpolation from that position. This was done for efficiency reasons, as a number of data could be at the same pressure but different horizontal positions. Having calculated background departures and standard deviations as well as ascribed standard deviations of observation errors, the scene is set to perform the background (or often called first guess) check. It is based on that the background departure should not exceed more than a certain number of standard deviations of the expected error variance at the observation position:

The normalised departure is:

$$\delta^o = (y - b)/\sigma_b \tag{4.23}$$

and the normalised observation error

$$\epsilon_o = \sigma_o/\sigma_b \tag{4.24}$$

where σ_b and σ_o are the standard deviations of first guess background and observation errors, respectively. Wind components are checked together by averaging u and v departures and u and v (normalised) observation errors.

If $\delta_o > (1 + \epsilon_o^2) \times ERRLIM(j)$ for j=1,3, then flags 1,2 or 3 are assigned. Otherwise 0. The values of ERRLIM are set varying by observation type and variable. 12.25, 25 and 36 for heights, 8, 18, 20 for wind and 2.25, 5.06 and 7.56 for thickness. For SATEM the values are reduced further by 0.2 to 1.0, depending on layer. For SATOBs the limits are multiplied by 0.1 except low level winds, 0.2. There is an asymmetric wind check for SATOBs, so that when the observed wind speed is 4 m/s lower than the background, the limits are reduced further (halved).

The flag 3 means always rejection, 2 rejection for non-standard level data but, but otherwise it is considered suspect, as is flag 1. The flag 1 and 2 data retained will be checked in combined checks and in the OI quality control check. It may also influence the selection of data for the analysis (in case of thinning, the lowest flagged data are preferred).

One such check is the multi-level check where subsequent layers around WMO standard pressure levels (TEMP code) are inspected for each TEMP, PILOT or SATEM. If four or more subsequent layers contain each at least one datum which is flagged, suspect or rejected, then the whole layer is rejected in the case of winds, the layer and all levels above for heights and the whole profile for SATEMs.

The OI rejection algorithm works in analogous way as the background check above, but using the estimated interpolation error instead:

$$\delta^i = (y - a)/\sigma_b \tag{4.25}$$

$$\delta_i > (\alpha + \epsilon_i^2) \times ERRLIM(j) \tag{4.26}$$

for j=1,3, where α is an empirical factor corresponding to 5 m in height error. The limits are 6,9 and 12, except for SATEM (4,6,8) and SATOB (1,1.5 and 2). The rejection limit is 2; all flagged 2 or 3 are rejected. If a SATEM report has one layer rejected, then the entire profile is rejected.

In addition to the above checks, a ship blacklist file is maintained. A ship is blacklisted if it has reported at least 3 times during the preceding 48 hours and has beend flagged a majority of these occasions. (Likewise a blacklisted ship can be cleared off the list once it is flagged 0 on a majority of occasions).

4.1.9 Data selection

First, certain observing practices are checked. SYNOP or TEMP height/pressure data extrapolated more than 800 m or 100 hPa below model terrain are discarded. For wind (and humidity) the limit is 25 hPa.

Observations are thinned in time and "duplicates" or what is considered as redundant information removed. In the time window where observations are used (usually 6 or 3 hours, the period of the analysis cycling) only one datum from a single level observation is kept. The one closest to the analysis time is used. Also the quality control flag is inspected; the datum with the lowest flag and closest to analysis time is kept. Note that this is done at datum level; the wind from a SYNOP may be from one hour and the pressure from a different time in some (very unusual) cases, due to availability and quality. For multi-level reports, TEMP and PILOTs, the selection is also for best quality and closest to analysis time. For both single and multi-level reports a check for duplicate is employed, so that only one datum for each variable and layer is kept from each station (based on its station identifier). Selection of multi-level reports is done for a number of specified layers, normally 15, so that datum closest to a WMO standard level (provided it has the lowest quality control flag) is chosen.

For moving platforms it is a little bit more tricky. Then an area of 0.5° is searched around each station so that only one report from the same identifier is kept over a distance of roughly 1° (providing that all data are present and considered correct). Aircraft data are thinned layer by layer, so that the condition applies for each WMO standard level layer. Then the important profiles from ascents/descents are still kept.

4.1.10 Analysis computations in boxes and overlapping

The analysis variables are height and u- and v- wind components, but the hydrostatic relationship allows for surface pressure and (virtual) temperature increments to be computed on model levels in the analysis evaluation. The three mass- and wind variables are analysed together, multi-variately. This is followed by a uni-variate analysis of relative humidity (see below). There is however a weak coupling since the mass and wind analysis is only working on virtual temperature and the final real temperature depends on the humidity analysis.

The analysis of single level pressure observations is augmented by adding artificial thin thickness (or temperature) data with zero increments from the background. This is to reduce the risk of interpreting the mass increments from observations at slightly different pressures as temperature information. For the same reason, such data are entered at the pressure of the first guess at the height of the observation, rather than the observed one.

The analysis equations are solved for one analysis box at a time. This is a compromise between the ideal global analysis and local analyses at each grid points. Inside the analysis boxes, the properties of the analysis structure function, variances, correlations and multi-variate coupling, are maintained. For the full analysis field a number of boxes need to be combined and then the specified qualities are lost to some degree in the overlap zones. A simple linear averaging is done in these zones, but this does not maintain the smooth structures for derivatives of the analysis increments such as vorticity and divergence.

Analysis boxes are constructed from a mesh of basic boxes, approximately of the size at least of the length correlation scales specified in the correlation functions, between 5 and 6° of latitude is usually chosen and varies slightly with the HIRLAM area. In the vertical there are two slabs, surface - 100 hPa and 300-10 hPa with an overlapping layer. The number of data in each box are scanned, and, if above a threshold value, the base box is divided into four with halved length of the sides. Again, if the number exceeds the threshold in the small boxes, one more subdivision

may take place.

Data are then selected in an area covering the whole box and a distance away from the box as well. A minimum selection radius from the box centre is defined which should always be adhered to, unless the maximum matrix size (of data) is exceeded. Data can be selected up to a maximum radius, until a minimum number of data has been found. The further out the data selection reaches, the better consistency between analyses in neighbouring boxes as they use more data in common. In very data dense areas, where the minimum distance may not be reached, or even not all data inside the box are used in extreme cases, the data density is likely to be that high that the analysis is well defined by the data alone and not relying as much on the structure functions.

Also an analysis area, extending beyond the box boundaries, is defined. All model grid points that lie inside that area are analysed from the box. Averaging is then done at each gridpoint using analyses from all boxes that include the grid point and with weights proportional to the distance from each box centre.

In the vertical, the overlapping is done by simple averaging for winds but for heights the overlapping is done in such a way so that analysed mean layer (virtual) temperatures are preserved.

4.1.11 Humidity analysis

The humidity analysis follows all the steps mentioned above, for the mass- and wind anlysis. The observations of SYNOP 2 m and TEMP level temperatures and dew points are converted to relative humidities. For SATEM, the Precitable water content is converted to relative humidity with aid of the vertical profile of the background (not the absolute values). A constant first guess error of 10 or 15% , depending on latitude, is assumed. No super observation formation is currently done. The analysis boxes are constructed in analogous way to the mass- and wind analysis, but obviously less critical since only humidity data enter the analysis and is done uni-variately. The analysis boxes are not overlapped for humidity. Quality control of super saturation and against the first guess background are done. The limit is 5 standard deviations. For the OI analysis check the rejection limit is 4 standard deviations.

Finally, the mass- and wind virtual temperature analysis increments and relative humidity increments are mutually converted to temperature and specific humidity through an iterative procedure.

4.2 Analysis of near-surface values

4.2.1 Introduction

The HIRLAM surface analysis is consistent with the mosaic approach adopted in the parameterization of soil/surface processes. As it is documented in a previous section describing the physical parameterizations, five different tiles are allowed within each grid square: water, ice, bare ground, low vegetation, and forest. Land fractions follow the ISBA scheme (Interaction Soil Biosphere Atmosphere) developed by Noilhan and Planton (1989). At each land tile, vegetation type and soil texture variables are defined. The vegetation type determines the albedo, local roughness length, vegetation coverage, leaf area index and minimum stomatal resistance, whereas the soil texture defines the hydrological properties like field capacity, wilting point and saturation values of the soil. The surface analysis module has been developed from the original surface parameters analysis operational at SMHI (Gustafsson, 1985). It has been substantially modified to adapt to the current parametrization requirements. It provides initial values to the following variables: water surface temperature (SST), fractions of ice and water, snow depth

over ice and land fractions, screen level temperature and relative humidity over land fractions, superficial and mean soil temperatures, and superficial and mean volumetric water contents.

4.2.2 Analysis Characteristics: SST

For the analysis of the water surface temperature a background field is created at each assimilation step as the previous analysis relaxed to its climatological value to allow for a seasonal evolution over persistent data void areas. Data used are observations contained in sea surface reports (SHIP and BUOY), and pseudo observations created from the ECMWF SST or Surface Temperature field, which is the daily NCEP global sea surface temperature analysis. They provide for a uniform data density over open ocean. Special observations, as e.g. locally derived SST values from AVHRR radiances, can be added optionally.

The observation operator consists of a bi-linear interpolation. It takes additionally into account the type of surface, either open ocean or inland lakes, of both grid points and observation. The fractions of land and lake fields allow the following classification of both observations and grid points in different classes: lake , inland, near coast, coastal and open sea:

Observations and grid points classification				
inland	near coast	coast	lake	open sea
$frland > 0.8$	$0.8 \geq frland > 0.5$	$0.5 \geq frland > 0.2$	$frlake > 0.02$	$frland \leq 0.2$

By using the above classification, some restrictions are imposed in the model interpolation of water surface temperature to the observations positions, in the sense that only grid points of similar type to the observation can be used to create the model equivalent of observed temperature.

The only quality control to data is a simple first guess check(each observed temperature should not be too different from the temperature in the previous analysis).

The default analysis method is successive corrections. The formulation of a priori determined observations weights follows Cressman (1959):

$$w(r) = \begin{cases} \frac{R^2 - r^2}{R^2 + r^2} & r \leq R \\ 0 & r > R \end{cases} \tag{4.27}$$

Three iterations with radius R at 600km, 400km, and 150km are applied.

In the current tiling scheme, water temperature and ice fraction represent two different fields. Their values over land and sea points refer to lakes and ocean, respectively. Thus, topography induces a discontinuity basically along the line coast, that should be taken into account. In particular, it is undesirable that a sea observation could influence surface temperature over lakes in the proximity (Eerola, 1995). In order to account for this land-sea anisotropy, observations weights depend not only on horizontal distance but they are further scaled using the grid points and observation classification. The following table is used for the scaling factors, $f(class_i, class_k)$, measuring the topographic features:

Topographic scaling factors					
obs.class	grid point class				
	inland	near coast	coast	lake	open sea
inland	1.	0.5	0.0	1.	0.0
near coast	0.5	1.	0.5	0.5	0.3
coast	0.0	0.5	1.0	0.0	0.9
lake	1.	0.5	0.0	1.	0.0
open sea	0.0	0.3	0.9	0.0	1.0

The analysis at grid point k, at iteration $j+1$, a_k^{j+1}, is then

$$a_k^{j+1} = a_k^j + \frac{\sum_{i=1}^{N} W_{ik}\sigma_o^{-2}(y_i - a_i^j)}{\sum_{i=1}^{N} W_{ik}\sigma_o^{-2} + \sigma_b^{-2}} \tag{4.28}$$

where $y_i - a_i^j$ is the background departure at the i observation position, σ_o and σ_b are respectively the observations and background error standard deviations, and W_{ik} are the weights given to observations i=1,N:

$$W_{ik} = w(r_{ik}) f(class_i, class_k) \tag{4.29}$$

In the first iteration, a_i^j is the created background from the previous assimilation cycle relaxed to the climatology.

Once the surface temperature over water has been analysed, the fractions of water and ice are diagnosed. The algorithm is linear between two limit temperature values. Different pairs of threshold values are used over the Baltic Sea, lakes and open ocean to take into account the water salinity.

4.2.3 Analysis Characteristics: Snow depth

In the mosaic scheme snow mass within the grid box comes from the ice fraction and the three land fractions. So, the analysed field is the fractional average over these tiles. The background snow mass is additionally relaxed to climatology and converted into snow depth by assuming a monthly varying snow density. The corresponding conversion factors (from water equivalent(m) to snow depth(cm)) are the following:

Conversion factor											
Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
450	430	416	360	320	320	320	700	700	620	550	470

The only information used are observations contained in land surface reports. The observation operator is simply a bi-linear interpolation, and data checking consists of a first guess check.

Again the analysis method is successive corrections with Cressman type dependence on horizontal distance of observations weights.

Observations weights are also scaled to account for the difference between model orography and observation height to prevent from getting analysed snow cover in the mountain surroundings if no stations reporting no snow are available (Eerola, 1995), by means of $f(\Delta z)$:

$$f(\Delta z) = \begin{cases} 0.0 & \Delta z \leq 0 \quad \mid \Delta z \mid \geq z_{crit} \\ \frac{{z_{crit}}^2 - \Delta z^2}{{z_{crit}}^2 + \Delta z^2} & \Delta z \leq 0 \quad \mid \Delta z \mid < z_{crit} \\ 1. & \Delta z > 0 \end{cases} \tag{4.30}$$

$$\Delta z = z^{mod} - z^{obs} \quad z_{crit} = 300m \tag{4.31}$$

The observations weights, $W_{ik}(r, \Delta z)$, are finally:

$$W_{ik}(r, \Delta z) = \begin{cases} f(\Delta z)\frac{R^2 - r^2}{R^2 + r^2} & r \leq R \\ 0 & r > R \end{cases} \tag{4.32}$$

Four scans are conducted at horizontal radius R 600km, 400km, 150km and 100km.

The analysed field converted back to equivalent water snow mass is copied into the ice and three land subtypes of surface.

4.2.4 Analysis Characteristics: T2m and RH2m

The main purpose of these screen level variables analysis is not only model validation or diagnosis but assimilation of soil temperatures and water contents at each of land tiles. The analysed field is the averaged model postprocessed field calculated from the three land subtypes surface temperature and specific humidity and snow mass. An effective local roughness length, z_0, is created from the local roughness length at each land tile, $z_{0,i}$, by means of a blending height, z_b, following the method proposed by Mason (1988) and used in the diagnosis of near surface parameters:

$$\frac{1}{ln^2\left(\frac{z_b}{z_0}\right)} = \sum_i \frac{A_i}{ln^2\left(\frac{z_b}{z_{0,i}}\right)} \tag{4.33}$$

being A_i the fractional area of i land subtype over the total land fraction in the gridsquare.

Relative humidity is the chosen humidity analysed variable. It allows for an easier modelling of first guess errors statistics. Data used are observed T2m and RH2m reported in SYNOPs corresponding to the analysis time. The model state projection onto observations space is described in Navascues (1997). It is a complex operator that tries to avoid artificial deviations in observation increments due to the misrepresentation of orography by the model. It mainly consists of application of the following operators:

- Horizontal bi-linear interpolation to observation position of atmospheric vertical profiles. Only grid points with not negligible fraction of land are used.
- Vertical interpolation of the whole profile from the model orography to the station height. This is carried out in the same way the boundary fields are vertically interpolated to the Hirlam grid, and thus preserving the stability structure in the boundary layer (Källén, 1996).
- Averaging over land of surface variables: Surface temperature and specific humidity, snow mass, and roughness length.
- Re-calculation of the surface temperature to keep the potential temperature lapse rate in the surface layer.
- Diagnosis of T2m and RH2m following Geleyn (1988).

The analysis method for both near surface parameters is univariate statistical interpolation. The principle has already been described above. It allows for an additional quality control to data, in which each observation is checked for agreement with its neighbours. This optimum interpolation check is only applied to data active and not rejected after the first guess check.

The autocorrelation model ρ is anisotropic, based on the relevant orographic features found when studying T2m and RH2m errors covariances corresponding to the old surface parametrization. The analytical formulation for the structure function consists of a gaussian dependence on both the horizontal and vertical distances $r, \Delta z$. Pragmatic characteristic length scales have been set as follows:

$$\rho(r, \Delta z) = exp[-0.5(\frac{r^2}{R^2} + \frac{\Delta z^2}{H^2})] \tag{4.34}$$

The values are $R = 100km$ for T2m and $R = 75km$ for RH2m and $H = 400m$ for both T2m and RH2m.

Values for observation and background errors variances have been assigned after some experimentation. Both the ratio of observations to first guess errors and characteristic length scales have been tuned to avoid too much noisy analysis increments, to minimize the rejected observations during the optimum interpolation check and to produce an analysis close to the observations.

The optimum interpolation check is performed by dividing the model area in boxes. Observations in a radius from the center of box are used together to decide the final status of data contained in the box. All active and no rejected observations are afterwards used in wider analysis boxes to solve the analysis problem.

4.2.5 Analysis Characteristics: Soil Temperatures

Superficial, T_s, and mean layer, T_d, soil temperatures are analysed separately for each subtype of surface following Giard and Bazile (2000). The method simply corrects the soil temperatures with the calculated 2m temperature analysis increment at every assimilation step:

$$\Delta T_d = \frac{\Delta T_{2m}}{2\pi} \tag{4.35}$$

$$\Delta T_s = \Delta T_{2m} \tag{4.36}$$

4.2.6 Analysis Characteristics: Soil Water contents

Superficial and total layer water contents, w_s and w_d, are assimilated separately at each tile. Water intercepted by vegetation, w_r, is not analysed, being simply copied from the first guess to the analysis. The method to initialize w_s and w_d is based on the sequential assimilation developed by Mahfouf (1991), with optimum coefficients approximated analytically as proposed by Bouttier et al. (1993), and further modified for operational implementation in the global French model Arpege by Giard and Bazile (2000). It was implemented in Hirlam by Ayuso (1995) simultaneously to the tiling surface package.

The analysed two meter temperature and relative humidity at each grid point are used as observations for the analysis of soil water contents. Soil moisture corrections are linearly calculated from the screen level analysis increments by means of an optimum interpolation analysis:

$$\Delta w_s = \alpha_s^T \Delta T_{2m} + \alpha_s^H \Delta H_{2m} \tag{4.37}$$

$$\Delta w_d = \alpha_d^T \Delta T_{2m} + \alpha_d^H \Delta H_{2m} \tag{4.38}$$

where $\alpha_s^T, \alpha_s^H, \alpha_d^T, \alpha_d^H$ are the optimum coefficients that minimize the mean square error of the estimation.

The statistics of the forecast errors on soil moisture was estimated by Mahfouf (1991). He used a Monte Carlo method consisting of a set of 24H perturbed one dimensional simulations, each of them started from a randomly modified soil moisture. Bouttier et al. (1993) analysed the distribution of the forecast errors statistics at each local solar time due to the variation of the incident radiation on the surface and the dependence of optimum coefficients on the surface characteristics. Then, they proposed an analytical formulation for $\alpha_s^T, \alpha_s^H, \alpha_d^T, \alpha_d^H$. The formulation of optimum coefficients in the current version differs from the original by Bouttier et al. (1993), and follows Giard and Bazile (2000). Coefficients depend on vegetation coverage, *veg*, local solar time, t^*, leaf area index, LAI, minimum stomatal resistance, R_{smin} and soil texture, *txt*:

$$\alpha_s^T = f(txt)(1 - veg)\left[a_0^T(t^*) + a_1^T(t^*)veg + a_2^T(t^*)veg^2\right] \tag{4.39}$$

$$\alpha_s^H = f(txt)(1-veg)\left[a_0^H(t^*) + a_1^H(t^*)veg + a_2^H(t^*)veg^2\right] \tag{4.40}$$

$$\begin{aligned}\alpha_d^T &= f(txt)\left\{(1-veg)\left[b_0^T(t^*) + b_1^T(t^*)veg + b_2^T(t^*)veg^2\right] + \right. \\ &\left. veg\frac{LAI}{R_{smin}}\left[c_0^T(t^*) + c_1^T(t^*)veg\right]\right\}\end{aligned} \tag{4.41}$$

$$\begin{aligned}\alpha_d^H &= f(txt)\left\{(1-veg)\left[b_0^H(t^*) + b_1^H(t^*)veg + b_2^H(t^*)veg^2\right] + \right. \\ &\left. veg\frac{LAI}{R_{smin}}\left[c_0^H(t^*) + c_1^H(t^*)veg\right]\right\}\end{aligned} \tag{4.42}$$

where $f(txt)$ accounts for the scaling of optimum coefficients for the different range of variation between wilting point, w_{wilt}, and field capacity, w_{fc}, for the different soil textures:

$$f(txt) = \frac{(w_{fc} - w_{wilt})_{txt}}{(w_{fc} - w_{wilt})_{loam}} \tag{4.43}$$

The coefficients in the polynomial terms, $a_i^T, a_i^H, b_i^T, b_i^H, c_i^T, c_i^H$, have been tuned by Giard and Bazile (2000), by fitting the adopted analytical function of the vegetation properties and local solar time to the original set of O.I. coefficients obtained by Mahfouf (1991).

A few constraints have been imposed to prevent undesired corrections in situations in which the atmosphere is not sensitive to errors in soil moisture. So, modifications to soil water content are not applied either in the cases of snow covered surface, strong wind speed, significant rain, attenuation of solar radiation by clouds, very short length of daylight or in case of dew deposition. Most of these conditions have been analytically introduced by mean of scaling factors w_1, w_2, w_3, decreasing the soil moisture corrections following the Arpege operational implementation (Bazile, personal communication):

$$w_1 = \begin{cases} \frac{daylen}{mindaylen} & daylen \le mindaylen \\ 1 & daylen > mindaylen \end{cases} \tag{4.44}$$

$$mindaylen = 6hours \tag{4.45}$$

$$w_2 = \begin{cases} 1. - \frac{\|\mathbf{v}^{10}\|}{\|\mathbf{v}_{max}^{10}\|} & \| \mathbf{v}^{10} \| \le \| \mathbf{v}_{max}^{10} \| \\ 0. & \| \mathbf{v}^{10} \| > \| \mathbf{v}_{max}^{10} \| \end{cases} \tag{4.46}$$

$$\| \mathbf{v}_{max}^{10} \| = 10 ms^{-1} \tag{4.47}$$

$$w_3 = \begin{cases} 1. - \frac{prec}{prec_{crit}} & prec \le prec_{crit} \\ 0. & prec > prec_{crit} \end{cases} \tag{4.48}$$

Where the critical value of accumulated precipitation over the past 6 hours, $prec_{crit}$, has been set to 6mm.

An analogous reduction of soil moisture increments in cloudy situations is modulated by the factor w_4:

$$w_4 = \begin{cases} 1. - \frac{cloud}{cloud_{crit}} & cloud \le cloud_{crit} \\ 0. & cloud > cloud_{crit} \end{cases}, \tag{4.49}$$

where $cloud_{crit} = 0.8$.

A test was further added to modify soil moisture only in the case of T2m and RH2m analysis increments have opposite sign.

The assimilation only takes place between wilting point and field capacity to avoid unrealistic corrections to the surface variables, provided that only within this range soil moisture is controlling the evaporation rate. In practice the lowest limit is set to veg*wilting point, to leave the minimum value allowed to have a dependence on the vegetation coverage.

The mean volumetric soil water content is changed by the mean of the last four analysis increments with the aid of a history file of soil wetness increments. It produces an smoother evolution of the soil moisture, as physically expected.

Systematic errors in screen level relative humidity are very much affected by errors in the specification of soil water content. However, two meter temperature systematic errors are very often due to other different reasons, as e.g. errors in the radiative forcing. Following Giard and Bazile (2000) a systematic long term T2m increment, ΔT_{2f}^n is calculated at the n^{th} assimilation cycle:

$$\Delta T_{2f}^n = (1 - r)\Delta T_{2f}^{n-1} + r\Delta T_{2m}, \tag{4.50}$$

where $r = 0.5$

Instead of the full two meter temperature increment, obtained in the n^{th} screen level parameters analysis, ΔT_{2m}, only the T2m effective increment, ΔT_{2m}^{eff}, defined as the deviation from this mean value is used to modify the soil moisture:

$$\Delta T_{2m}^{eff} = \Delta T_{2m} - \Delta T_{2f}^n. \tag{4.51}$$

4.3 Under development: HIRLAM 3-d variational data assimilation

The HIRLAM 3-dimensional variational (3D-Var) data assimilation is an original HIRLAM contribution (Gustafsson et al., 2001; Lindskog et al., 2001). The basic approach of HIRLAM 3D-Var is common with other operational 3D-Var implementations, i.e. an iterative reduction of model state misfit to available information, e.g. (Parrish and Derber, 1992; Courtier et al., 1998). The HIRLAM implementation of 3D-Var contains however several unique scientific developments generally applicable to LAMs (Limited Area Model). These novelties are related to the formulation and the computational efficiency of the background error constraint and to the treatment of lateral boundaries.

Atmospheric data assimilation problem can be formulated as a task of finding the model state vector $\mathbf{x}$, which would correspond the minimum of the cost function

$$J = J_b + J_o = \frac{1}{2}(\mathbf{x} - \mathbf{x^b})^T\mathbf{B^{-1}}(\mathbf{x} - \mathbf{x^b}) + \frac{1}{2}(H\mathbf{x} - \mathbf{y})^T\mathbf{R^{-1}}(H\mathbf{x} - \mathbf{y}). \tag{4.52}$$

Here J_b measures the distance to a background model state $\mathbf{x^b}$ and J_o measures the distance to the vector $\mathbf{y}$ of the observations. The observation operator H transforms a model state into the observed quantities. $\mathbf{B}$ is the matrix containing the background error covariances, while $\mathbf{R}$ is a matrix containing the observation error covariances.

4.3.1 Incremental formulation

HIRLAM variational data assimilation can be applied with the grid point or spectral HIRLAM forecast models, and it can be applied in its *incremental* form (Courtier et al., 1994) such that the

full resolution assimilation increment $\delta\mathbf{x} = \mathbf{x} - \mathbf{x^b}$ is represented at a lower horizontal resolution ($\delta\mathbf{x}'$) than the full model state. The cost function for the incremental formulation is defined in this case as follows:

$$J = J_b + J_o = \frac{1}{2}\delta\mathbf{x}'^T\mathbf{B}'^{-1}\delta\mathbf{x}' + \frac{1}{2}(H\mathbf{x^b} + \mathbf{H}'\delta\mathbf{x}' - \mathbf{y})^T\mathbf{R}^{-1}(H\mathbf{x^b} + \mathbf{H}'\delta\mathbf{x}' - \mathbf{y}) \tag{4.53}$$

The background error covariance matrix $\mathbf{B}'$ now has the squared dimension of the assimilation increment. The full non-linear H operates on the background field $\mathbf{x^b}$, while a linearized $\mathbf{H}'$ operates on the analysis increment $\delta\mathbf{x}' = \mathbf{x}' - (\mathbf{x^b})'$. The linearization is carried out with reference to the background field $\mathbf{x^b}$.

For efficient minimization of J with respect to the analysis increment $\delta\mathbf{x}'$, gradient of the cost function with respect to the model state increment vector $\delta\mathbf{x}'$ is required:

$$\nabla_{\delta\mathbf{x}'}J = \mathbf{B}'^{-1}\delta\mathbf{x}' + \mathbf{H}'^T\mathbf{R}^{-1}(H\mathbf{x^b} + \mathbf{H}'\delta\mathbf{x}' - \mathbf{y}) \tag{4.54}$$

4.3.2 Formulation of the background error constraint for a limited area model

In HIRLAM 3D-Var, the model state increment vector $\delta\mathbf{x}$ includes all spectral components of the following model increment variables to be determined by the variational data assimilation:

$$\delta\mathbf{x} = \begin{pmatrix} \delta\mathbf{u} \\ \delta\mathbf{v} \\ \delta\mathbf{T} \\ \delta\mathbf{q} \\ \delta\ln\mathbf{p_s} \end{pmatrix} \quad \begin{array}{l} \text{- the increment vector of the wind component in the x-direction} \\ \text{- the increment vector of the wind component in the y-direction} \\ \text{- the temperature increment} \\ \text{- the specific humidity increment} \\ \text{- the increment of the logarithm of the surface pressure} \end{array} \tag{4.55}$$

The HIRLAM model state includes some further variables, e.g., soil moisture and surface temperature. These grid point variables are not included among those to be controlled by the variational data assimilation, i.e. HIRLAM 3D-Var provides an upper air analysis.

4.3.3 Preconditioning and change of variable

For a fast convergence of the minimization algorithm, a pre-conditioning of the minimization problem is needed. An ideal pre-conditioning is obtained if the Hessian matrix (i.e., the second order derivative of the cost function) is an identity matrix. A good approximation of this is to ensure that the Hessian of the background error contribution J_b to the cost function is equal to the identity matrix (Lorenc, 1988). This can be achieved by defining a change of variable $\mathbf{U}$ to be applied to the assimilation increment $\delta\mathbf{x} = \mathbf{x} - \mathbf{x^b}$ such that it transforms the forecast error ϵ in model space into ϵ', a variable whose covariance matrix is an identity matrix. This change of variable can be written as $\epsilon' = \mathbf{U}\epsilon$, or equivalently $\epsilon = \mathbf{U}^{-1}\epsilon'$. This implies an equality between the corresponding covariance matrices:

$$\mathbf{B} = \overline{(\epsilon\epsilon^T)} = \mathbf{U}^{-1}\overline{(\epsilon'\epsilon'^T)}\mathbf{U}^{-T} = \mathbf{U}^{-1}\mathbf{U}^{-T} \tag{4.56}$$

if $\overline{(\epsilon'\epsilon'^T)} = \mathbf{I}$ and $\overline{(\ )}$ denotes ensemble averaging. $\mathbf{U}$ appears to be a left-hand inverse square-root of the forecast error covariance matrix.

$$\mathbf{B}^{-1} = \mathbf{U}^T\mathbf{U} \tag{4.57}$$

and the expression for J_b with respect to $\chi = \mathbf{U}\delta\mathbf{x}$ becomes:

$$J_b = \frac{1}{2}\delta\mathbf{x}^T\mathbf{B}^{-1}\delta\mathbf{x} = \frac{1}{2}\chi^T\mathbf{U}^{-T}(\mathbf{U}^T\mathbf{U})\mathbf{U}^{-1}\chi = \frac{1}{2}\chi^T\chi \tag{4.58}$$

The transform $\mathbf{U}$ converts the forecast error in physical model space to a variable, whose covariance matrix can be assumed to be an identity matrix. The design of $\mathbf{U}$ contains the following aspects:

- **Mass/wind balance**. Forecast error fields are known to be in near geostrophic balance in the extratropics and assimilation increment fields should ideally also be in near geostrophic balance. One way to obtain de-correlated forecast errors with regard to this balance is therefore to transform the full wind forecast error to an ageostrophic forecast error by subtraction of the geostrophic wind error, as determined from the mass field forecast error. The ageostrophic wind forecast errors are moreover assumed to be de-correlated with the temperature and surface pressure forecast errors.
- **Horizontal spectral transforms**. Forecast errors are known to be horizontally correlated in physical space. As the assimilation control variable is defined in spectral space, different spectral forecast error modes can be assumed to be statistically independent, given the assumption that spatial covariances are horizontally homogeneous in physical space (Boer, 1983).
- **Vertical transforms**. The residual correlations that need to be accounted for are those involving the same horizontal spectral modes at different vertical levels. Projecting each bi-Fourier forecast error on the eigenvectors of the corresponding vertical correlation matrix provides variables whose covariance matrix is diagonal. This diagonal matrix contains the associated eigenvalues. Dividing by the square-roots of these eigenvalues allows one to obtain a variable with a covariance matrix that is equal to the identity matrix.

4.3.4 Inverse change of variable and its adjoint

The complete transformation from the analysis increments in model space to the control variable is:

$$\chi = \mathbf{P}\,\mathbf{V}\,\mathbf{L}\,\mathbf{F}\,\mathbf{S}^{-1}\,\mathbf{A}\,\mathbf{F}^{-1}\,\delta\mathbf{x} \;=\; \mathbf{U}\delta\mathbf{x} \;=\; \mathbf{U}(\mathbf{x}-\mathbf{x}^{\mathbf{b}}) \tag{4.59}$$

where $\delta\mathbf{x}$ is the assimilation increment, $\mathbf{F}$ is the Fourier transform to spectral space, $\mathbf{F}^{-1}$ is the inverse Fourier transform, $\mathbf{A}$ is the subtraction of the geostrophic wind increment from the full wind increment, $\mathbf{S}^{-1}$ is the normalization with the forecast error standard deviation, $\mathbf{L}$ is the normalization by the square-root of the spectral density of the horizontal forecast error correlation, $\mathbf{V}$ is the projection on the eigenvectors of the vertical forecast error correlation matrix and $\mathbf{P}$ is the normalization by the square root of the vertical eigenvalue.

These transforms are never applied in their forward form during the variational data assimilation. At the start of the minimization process, all elements of the control vector χ are put equal to zero. Then, during each iteration, the inverse transform $\mathbf{U}^{-1}$ is applied to obtain the spectral assimilation increments in the model space:

$$\delta\mathbf{x} \;=\; \mathbf{U}^{-1}\,\chi \;=\; \mathbf{F}\,\mathbf{A}^{-1}\,\mathbf{S}\,\mathbf{F}^{-1}\,\mathbf{L}^{-1}\,\mathbf{V}^{-1}\,\mathbf{P}^{-1}\,\chi \tag{4.60}$$

Knowing that $\mathbf{B}^{-1} = \mathbf{U}^T\mathbf{U}$, the value of the cost function is evaluated as:

$$J = J_b + J_o = \frac{1}{2}(\chi)^T(\chi) + \frac{1}{2}(H\mathbf{x}^{\mathbf{b}} + \mathbf{H}\mathbf{U}^{-1}\chi - \mathbf{y})^T\mathbf{R}^{-1}(H\mathbf{x}^{\mathbf{b}} + \mathbf{H}\mathbf{U}^{-1}\chi - \mathbf{y}) \tag{4.61}$$

The adjoint of the inverse transform $\mathbf{U}^{-T}$ is then applied to obtain the gradient of the cost function with respect to the control variable χ:

$$\nabla_\chi J = \nabla_\chi J_b + \nabla_\chi J_o = \chi + \mathbf{U}^{-\mathbf{T}}\mathbf{H}^T\mathbf{R}^{-1}(H\mathbf{x}^{\mathbf{b}} + \mathbf{H}\mathbf{U}^{-1}\chi - \mathbf{y}) \tag{4.62}$$

$\mathbf{H^T}$ and $\mathbf{U^{-T}}$ include the application of the adjoint of the inverse Fourier transform (which is the same as a direct Fourier transform) to a grid point field of departures between the analysis and the observations.

4.3.5 Estimation of background error statistics

Estimation of background error statistics makes use of accumulated differences between +24 h and +48 h HIRLAM forecasts that are valid at the same time. These statistics of forecast differences are considered as an approximation to the true forecast errors in this "NMC method" (Parrish and Derber, 1992; Rabier et al., 1998). Standard deviations of the forecast differences are rescaled to match the amplitude of +6 h forecast errors. An empirical scaling factor of 0.6 is used in HIRLAM 3D-Var (Berre, 2000). This scaling factor varies slightly with season. Assumptions of horizontal homogeneity and isotropy of the forecast error correlations are made in the first version of the structure functions.

The bi-Fourier formulation of the background error constraint allows to apply *non-separable assimilation structure functions* in a limited area data assimilation. This improves the realism of the background error constraint since it enables to represent the height dependence of horizontal correlations and the scale dependence of vertical correlations in HIRLAM 3D-Var. This is in sharp contrast with the past systems where it was usual to assume that three-dimensional auto-correlation functions were *separable*. This meant that the auto-correlation functions could be written as the product of a horizontal correlation function that was independent of height, and a vertical correlation function which was supposed to be independent of horizontal coordinates.

The basic data set is accumulated of the operational SMHI HIRLAM runs with a resolution of $0.4^\circ \times 0.4^\circ$ in the horizontal and with 31 vertical levels. The data set covers the period 1 December 1997 through 28 February 1998 for the variables of the HIRLAM 3D-Var control vector.

4.3.6 The observation handling system

Observations in BUFR-format are input to OBSPROC-module which outputs a CMA-file (Central Memory Array) (Note: the OBSPROC-module is developed by and shared with ECMWF). While the CMA is being prepared, necessary information is checked for gross errors and, if necessary, observed variables are transformed into those used directly by the variational assimilation. The specified observation error standard deviations are assigned. CMA-file is input to HIRVDA (HIRlam Variational Data Assimilation). The HIRVDA observation *screening* performs quality control using various algorithms, including a comparison with the background field, data rejections and data thinning to select only those observed values that actually will be used by the variational data assimilation. The HIRLAM 3D-Var produces the analysis file and the CMA file is updated with the appropriate monitoring information, such as quality control flags and observations minus analysis departures. The *feedback* statistics are produced from the updated CMA files for diagnostic purposes.

The screening decisions in logical order are as follows:

- Logical checks; Verify that observations are situated within the analysis area and that SHIP reports originate from ocean areas.
- Representativity checks; Stations situated too far away from the model orography are rejected, for instance.
- Background quality control (BgQC).

- Blacklisting; Reject observations from stations situated in particular geographical areas or observations that from poor quality stations.
- Multilevel check; If more than four consecutive levels are flagged suspicious, all of these data are rejected.
- Station level; For land surface stations the observed station altitude pressure and the corresponding geopotential height are selected rather than the mean sea level pressure and the corresponding zero height.
- Data redundancy check; All colocated land surface observations which have passed the former tests are considered in this test. Only the observation closest to the time of analysis is accepted. Radiosondes and PILOT balloon reports are simultaneously checked. If both types of reports exist from the same station the radiosonde report is selected by preference. Furthermore, in the case of multiple reports of the same type, the one closest to analysis time is chosen.
- Moving platform check; A thinning is applied by retaining only the observation closest to the time of analysis from the same station identifier. The minimum horizontal distance allowed between two moving platform observation reports of the same type has initially been set to half a grid distance. In addition, for aircraft observations, a thinning is applied in the vertical by allowing a minimum vertical pressure difference of 50 hPa.

4.3.7 Background quality control

In the background quality control (BgQC) the background model state $\mathbf{x}^b$ is projected on the observed quantity y_i with the observation operator H. The squared background departure from the observation i is calculated and normalised with the sum of the observation error variance, $\sigma^2_{o,i}$, and the background error variance, $\sigma^2_{b,i}$. The observation, y_i, is rejected if it does not satisfy the following inequality:

$$L_1 \leq ([H(\mathbf{x}^b)]_i - y_i) \mid ([H(\mathbf{x}^b)]_i - y_i) \mid /(\sigma^2_{b,i} + \sigma^2_{o,i}) \leq L_2, \tag{4.63}$$

where L_1 and L_2 are the upper and lower rejection limits and $[H(\mathbf{x}^b)]_i$ denotes the projection of the model state on observation i. The rejection limits currently used are obtained from historical rejection series of the OI system, and are dependent on type of variable. Since satellite cloud track wind (SATOB) observations are systematically too weak, $-L_1$ is smaller than L_2 for this particular observation type. For all other data $-L_1$ is equal to L_2. The coefficients L_1 and L_2 are tuning constants. In the case of wind, both components are checked jointly. For high level winds with wind speed greater than 15 m/s, a wind direction check is applied.

The σ_b-values vary spatially representing the differences in flow characteristics in different geographical areas, as well as from variations in the station density. This external file of σ_b-values based on ECMWF forecasts provides a global coverage allowing the assimilation system to run over different geographical areas.

4.3.8 Parallelisation

The HIRVDA requires that all observations are kept in the memory simultaneously. In order to achieve a load balance among the processors of a parallel computer, a parallelisation strategy that takes the actual observation distribution into account is applied.

The parallelisation of the J_b grid point space calculations, as well as the corresponding spectral model calculations, builds on a sub-division of the model domain into equally sized sub-areas and

a distribution of one such sub-area to each processor of the parallel computer. The main parallelisation strategy for the treatment of observations in the HIRLAM 3D-Var is a combination of the following two extreme strategies:

A: Distribute all the observations in accordance with the horizontal position of the observations to the processors on which the corresponding sub-area of the grid point fields resides. Carry out all the observation operator calculations on the same processor.

B: Distribute the observations, independent of their horizontal positions, to the available processors in such a way that each processor will have the same amount of computational work for the observation operators. One possibility is to distribute an equal number of observations of each type to each processor. Fetch the grid point values needed for observation operator calculations by message passing.

Strategy A is simple and identical to the parallelisation of the forecast models. The disadvantage is a poor load balance of the observation operator calculations. Strategy B has a good load balance in the observation operator calculations, the disadvantage being the need to communicate grid point information between the processors; the efficiency of this may depend on the speed of the communication network between the processors.

For the HIRLAM 3D-Var the horizontal interpolation part of the observation operator is carried out with strategy A, and the remaining parts of the observation operator calculations are carried out with strategy B. This "main" parallelisation strategy is memory consuming, since the horizontally interpolated grid point fields need to be calculated in advance and stored in the computer memory for all observation positions. Therefore the parallelisation strategy A is available as an option, for computers with too small memory for the main strategy.

The parallelisation of the HIRLAM variational data assimilation is based on explicit message passing, i.e. intermediate calculation results are communicated between the computer processors by means of messages.

4.3.9 The observation error constraint

The observation error constraint J_o measures the distance between the model state $\mathbf{x}$ and the observations $\mathbf{y}$. The observation operator H provides the model counterparts of the observed quantities. With the *tangent-linear approximation* the observation error constraint becomes

$$J_o = \frac{1}{2}(H\mathbf{x^b} + \mathbf{H}\delta\mathbf{x} - \mathbf{y})^T\mathbf{R^{-1}}(H\mathbf{x^b} + \mathbf{H}\delta\mathbf{x} - \mathbf{y}), \tag{4.64}$$

where $\mathbf{H}$ denotes the tangent-linear observation operator, linearised around the background field $\mathbf{x^b}$. Since $\mathbf{H}$ is linear, the gradient of the observation error constraint J_o with regard to the model state increment vector $\delta\mathbf{x}$ is

$$\nabla_{\delta\mathbf{x}} J_o = (\mathbf{H})^T\mathbf{R^{-1}}(H\mathbf{x^b} + \mathbf{H}\delta\mathbf{x} - \mathbf{y}). \tag{4.65}$$

The observation errors are assumed to be un-correlated. The covariance matrix $\mathbf{R}$ for the observation errors is thus diagonal with the specified σ_o-values on the diagonal, and J_o reduces to a sum over all observations.

4.3.10 The observation operators

The observation operator H and the tangent-linear observation operator $\mathbf{H}$ are sub-divided into a sequence of sub-operators. Formally, we may write:

$$H = H_{spec} I_v P_{calc} \mathbf{I_h} \mathbf{F^{-1}}, \tag{4.66}$$

where $\mathbf{F}^{-1}$ is an inverse Fourier transform, $\mathbf{I_h}$ denotes horizontal (bi-linear) interpolation of model level data from grid points to the horizontal positions of the observations, P_{calc} calculation of pressures and geopotentials at model full and half levels (by integrating the hydrostatic equation), I_v vertical interpolation to the levels of the observed data values (linear in logarithm of pressure) and H_{spec} any other specialised operators for each type of observation. H will only be applied in the case of a spectral forecast model.

With regard to the conventional observations, only a few special types of sub-operators are applied at present: the extrapolation of geopotential below the model lower boundary surface and the calculation of wind, temperature and humidity at the reporting levels of SYNOP, SHIP and DRIBU reports.

Geleyn (1988) describes a method to calculate the vertical interpolation between the lowest model level and the surface. We use this method to interpolate model values for the observed 10 m wind (u_{10m},v_{10m}) as well as the 2 m temperature (T_{2m}) and relative humidity (RH_{2m}) from SYNOP, SHIP and DRIBU reports. This highly non-linear operator based on the Monin-Obukhov similarity theory for the surface layer takes turbulence and terrain characteristics into account. The following two modifications of the method introduced by Geleyn (1988) are made. The roughness length over open sea is iteratively determined from the surface friction velocity (u^* in the Charnock formula). In addition, the formula is improved with an additional term dependent on "gustiness stability", to ensure that in low wind and unstable conditions the heat fluxes do not drop to zero because of a vanishing roughness.

4.3.11 Data usage and observation errors

The following observations are assimilated by the HIRLAM 3D-Var system:

- TEMP reports; all significant level data of temperature, wind and specific humidity.
- SYNOP, SHIP and DRIBU reports; the station level height, multiplied by the gravitational acceleration, g, is used as a geopotential observation at the observed station level pressure. 10 m winds from SHIP reports. [The observed 10 m wind (u_{10m},v_{10m}), 2 m temperature (T_{2m}) and 2 m relative humidity (RH_{2m}) might also be used.]
- AIREP reports; wind and temperature observations.
- PILOT reports; all significant level wind data, observed at pressure or height levels.
- SATOB reports; wind observations.

The specified σ_o-values are given in Appendix 3. These σ_o-values include both the instrumental and the representativeness errors. For continuity reasons, the σ_o-values are interpolated in the logarithm of the pressure to the observed pressure. Above 10 hPa and below 1000 hPa the errors are assumed constant. An empirical regression relation

$$\sigma_o = -0.0015\ T + 0.54 \ \text{ for } \ 240\ \text{K} < T < 320\ \text{K} \tag{4.67}$$

is used to specify the σ_o for relative humidity. This relation is obtained through a statistical analysis of the dependence of the errors on temperature. For temperatures below 240 K, σ_o takes the value 0.18 and for temperatures above 320 K it takes the value 0.06. The σ_o for specific humidity is derived from the tangent-linear relationship that relates specific and relative humidities.

The following observations are assimilated by the HIRLAM 3D-Var system *in research mode*:

- (A)TOVS; brightness temperatures of selected channels.
- GPS (Global Positioning System); Zenith Total Delay observations.
- Doppler-radar wind; VAD and radial wind superobservations.
- Scatterometer; QuikScat.

4.3.12 Variational quality control

The variational assimilation algorithm is formulated on the basis of Bayes theorem (Lorenc, 1986), identifying the cost function J with the logarithm of a probability density function (p.d.f.). The variational quality control (VarQC) accounts for the possibility of observational gross errors, represented by a flat p.d.f., in the data presented to the variational analysis (Lorenc and Hammon, 1988), in addition to random errors, represented by a Gaussian p.d.f.. For an individual observation value, y_i, given the state vector $\mathbf{x}$, the p.d.f. is

$$p_i = P\frac{1}{D} + (1-P)\frac{1}{\sqrt{2\pi}\sigma_i}e^{(-0.5{z_i}^2)}, \tag{4.68}$$

where P stands for the *a priori* probability of observation i having a gross error. $z_i = \frac{y_i - [H\mathbf{x}]_i}{\sigma_i}$ is the departure between the observed value, y_i, and the model state projected on observation i, normalised by the observation error standard deviation (σ_i). The factor D represents the range of possible values, all with the same probability in the case of gross error. It is assumed that the data presented to the variational analysis have been previously checked, in order to keep the absolute value of z_i less than $D/(2\sigma_i)$ (D needs to be consistent with the parameters L_1 and L_2 of BgQC).

The effect of the VarQC is to scale the magnitude of the gradient of the observation part of the cost function obtained assuming a Gaussian error distribution only. An observation can be considered rejected when the scaling factor

$$r = \frac{\nabla_{z_i} J_{oi}^{QC}}{\nabla_{z_i} J_{oi}}, \tag{4.69}$$

or the *a posteriori* weight, r, is (for instance) less than 0.75. Here J_{oi}^{QC} and J_{oi} are the contributions from observation i to the observation part of the cost function with and without VarQC, respectively. These are given by identifying the observation part of the cost function with the logarithm of p_i, with P larger than zero and equal to zero, respectively. The value of r depends on the *a priori* parameters P and D as well as on the normalised departure z_i. Following the procedure used at ECMWF (Andersson and Järvinen, 1999) for this purpose, P and D have been estimated from the historical rejections series of operational HIRLAM OI analyses, for the different variables of all kind of observation reports. The values obtained clearly show a dependence on observation system and variable, as well as on observation density. The latter dependence has not yet been taken into account in the HIRLAM 3D-Var. For wind data, it is assumed that gross errors affect both of the wind components and these are checked jointly.

Because of the global data usage in 3D-Var, all observations are used in VarQC to support or reject an observation. Due to the iterative nature of the variational procedure, the *a posteriori* weight, r, evolves with the number of cost function evaluations. Observations are not definitively rejected or accepted at a fixed calculation step. At present, since a spatially varying P has not yet been introduced, there is no direct dependence on the background error or the analysis error.

The procedure is to switch on the VarQC after 20 iterations, and to switch off VarQC again after 70 iterations, before reaching the convergent solution of the minimization. This procedure

is related to the undesired fact that VarQC introduces multiple minima to the cost function as it is no more strictly quadratic.

4.4 Digital filter initialization

The requirement to modify meteorological analyses to avoid spurious high frequency oscillations in numerical forecasts has been known from the beginning of numerical weather prediction. The spurious oscillations are caused by gravity wave components of large amplitude, resulting primarily from minor errors in the observations. The most popular method of initialzation up to recently was the normal mode method (Machenhauer, 1977). A method equivalent to the normal mode method, called implicit normal mode initialization (INMI), was introduced by Temperton (1988). For INMI it is not necessary to derive the normal modes explicitly; the problem is reduced to the numerical solution of Helmholtz equations. This method was used in the Hirlam model up to 1999, when an alternative method of initialization was introduced. Digital filter initialization (DFI) was devised by Lynch and Huang (1992) and Huang and Lynch (1993). An enhanced method using a simple optimal filter, the Dolph-Chebyshev filter, was applied to the initialization problem by Lynch (1997) and a more efficient way of performing the initialization was presented by Lynch et al. (1997).

The method of digital filter initialization, which is now in use operationally at several major weather prediction centres, has several significant advantages over alternative methods: the method is applied consistently to all prognostic model variables (cloud water, rain water, turbulent kinetic energy, etc., as well as mass and wind); it is applicable to non-hydrostatic models without difficulty; there is no need to compute or store normal modes; no vertical separation is required, and all model vertical modes are filtered; exotic discretizations and irregular model geometries can be facilitated; finally, the simplicity of the scheme makes implementation and maintenance easy. The implementation in the Hirlam model is described comprehensively in a Hirlam Technical Report by Lynch et al. (1999) (denoted by TR42).

4.4.1 The Dolph-Chebyshev Filter

The details of the Dolph-Chebyshev filter are presented in Lynch (1997) We present only essential details here. The low-pass filter response is defined in the frequency domain by

$$H(\theta) = \frac{T_{2M}\left(x_0 \cos\left(\theta/2\right)\right)}{T_{2M}(x_0)} \tag{4.70}$$

where T_{2M} is a Chebyshev polynomial, $x_0 > 1$ and θ_s is such that $x_0 \cos(\theta_s/2) = 1$. By means of the definition of $T_{2M}(x)$ and basic trigonometric identities, $H(\theta)$ can be written as a *finite* expansion

$$H(\theta) = \sum_{n=-M}^{+M} h_n \exp(-in\theta). \tag{4.71}$$

The coefficients $\{h_n\}$ may be evaluated from the inverse Fourier transform

$$h_n = \frac{1}{N}\left[1 + 2r \sum_{m=1}^{M} T_{2M}\left(x_0 \cos\frac{\theta_m}{2}\right) \cos m\theta_n\right], \tag{4.72}$$

where $|n| \leq M$, $N = 2M + 1$, $\theta_m = 2\pi m/N$ and $r = 1/T_{2M}(x_0)$. The weights $\{h_n : -M \leq n \leq +M\}$ define the Dolph-Chebyshev or, for short, Dolph filter. In the Hirlam model, the

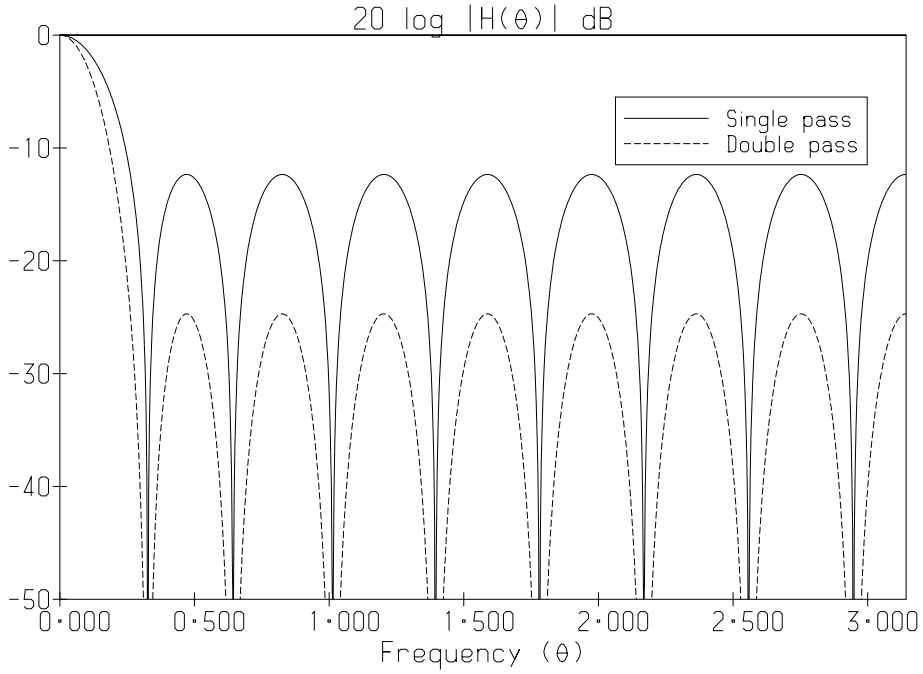


Figure 4.1: *Logarithmic response (dB) as a function of frequency for Dolph filter with span* $T_S = 2$h*, order* $N = 2M + 1 = 17$ *and cut-off* $\tau_s = 3$h*. Results for single and double application are shown.*

filter order $N = 2M + 1$ is determined by the time step Δt and forecast span T_S. The desired frequency cut-off is specified by choosing a value for the cut-off period, τ_s. Then $\theta_s = 2\pi\Delta t/\tau_s$ and the parameters x_0 and r are given by

$$\frac{1}{x_0} = \cos\frac{\theta_s}{2}, \tag{4.73}$$

$$\frac{1}{r} = \cosh\left(2M\cosh^{-1}x_0\right). \tag{4.74}$$

This may be used to evaluate the *ripple ratio*, which is a measure of the maximum amplitude in the stop-band $[\theta_s, \pi]$:

$$r = \left[\frac{\text{side-lobe amplitude}}{\text{main-lobe amplitude}}\right] \tag{4.75}$$

The Dolph filter has minimum ripple-ratio for given main-lobe width and filter order.

The DFI procedure employed in the Hirlam model involves a double application of the filter. Thus, we examine both the frequency response $H(\theta)$ and its square, $H(\theta)^2$, as the effect of a second pass through the filter is to square the frequency response. The parameters chosen for the DFI tests below are span $T_S = 2$ h, cut-off period $\tau_s = 3$ h and time step $\Delta t = 450$ s $= \frac{1}{8}$ h. So, $M = 8$, $N = 17$ and $\theta_s = 2\pi\Delta t/\tau_s \approx 0.26$. [In conventional units, the cut-off frequency is $f_s \approx 100\,\mu$Hz]. The response and square response are shown in Fig.4.4.1. The ripple ratio has the value $r = 0.241$. For a double pass, the minimum attenuation of high frequencies (components with $|\theta| > |\theta_s|$) is about 25dB, more than adequate for elimination of HF noise. In the Appendix of Lynch (1997) it is proved that the Dolph window is an *optimal* filter whose pass-band edge, θ_p, is the solution of the equation $H(\theta) = 1 - r$.

The digital filter initialization is performed in two stages by applying the filter to time series of model variables. In the first stage, a backward integration from $t = 0$ to $t = -T_S$ is performed, with all irreversible physics switched off. The output $\bar{x}$ is valid at time $t = -\frac{1}{2}T_S$. In the second stage, a forward integration is made from $t = -\frac{1}{2}T_S$ to $t = +\frac{1}{2}T_S$, starting from the output of the first stage. Now the output is valid at the centre of the interval $[-\frac{1}{2}T_S, +\frac{1}{2}T_S]$, *i.e.*, at $t = 0$. The output of the second pass of the filter is the initialized data. The values of the prognostic variables at the lateral boundaries are left unchanged during the digital filtering process. Complete details of the technical implementation are given in TR42.

4.4.2 Application of the method

The digital filter initialization was compared to the implicit normal mode initialization (NMI) scheme, and to forecasts with no initializaiton (NIL). Forecasts starting from the analysis valid at 1200 UTC on 10 February, 1999 were compared. The maximum change in surface pressure induced by DFI was 2.2hPa, with an *rms* change of about 0.5hPa. The changes to the other analysed variables are in general comparable in size to analysis errors, and considerably smaller in magnitude than typical changes brought about by the analysis itself. The impact of DFI on the 24 hour forecast was much smaller: the maximum difference in surface pressure between the uninitialized and initialized 24-hour forecasts was less than 1hPa and the *rms* difference was only 0.1hPa. For a full tabulation of results, see TR42.

The basic measure of noise is the mean absolute value of the surface pressure tendency. In Fig.4.4.2 we plot the course of this quantity for three forecasts. The solid line represents the forecast from uninitialized data (NIL), the dashed line is for a forecast starting from data initialized using the implicit normal mode method (NMI), and the dotted line is for the digitally filtered data (DFI). It is clear from this measure that DFI is most effective in removing high frequency noise.

A more comprehensive evaluation was undertaken using a FASTEX dataset (Joly et al., 1997). The FASTEX analyses covered the period 10–17 February, 1997. Using a 3-hour data assimilation cycle, two sets of forecasts were generated using the DFI and NMI initialization schemes. The results were very encouraging: the forecast skill for the DFI scheme was markedly superior to that of the control run. Scores were either better than or comparable to those using the older implicit NMI scheme. In Fig.4.4.2 we plot the *rms* and bias errors for mean sea-level pressure as a function of forecast time, averaged over all thirty forecasts. The lines marked with crosses are for the reference run, and those marked with spots are for the DFI run. The *rms* scores are significantly better for DFI, and the improvement is greater for longer forecast times. At 48 hours, there is a reduction in *rms* error of about 0.5hPa, which is substantial, representing an increase in predictive skill of six hours. The bias error (dashed curve) is reduced slightly.

Following further satisfactory testing and evaluation of the digital filtering initialization scheme, it was introduced as the reference method in Version 4.6.3 of the Hirlam model in October, 1999.

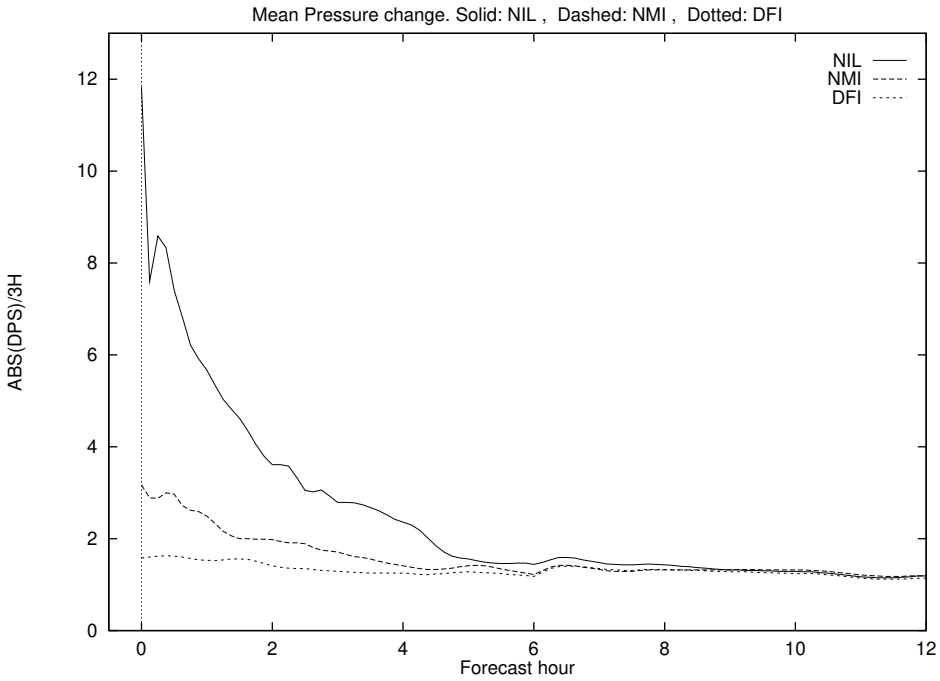


Figure 4.2: *Mean absolute surface pressure tendency for three forecasts. Solid: uninitialized analysis (NIL). Dashed: Normal mode initialization (NMI). Dotted: Digital filter initialization (DFI). Units are hPa/3 hours.*

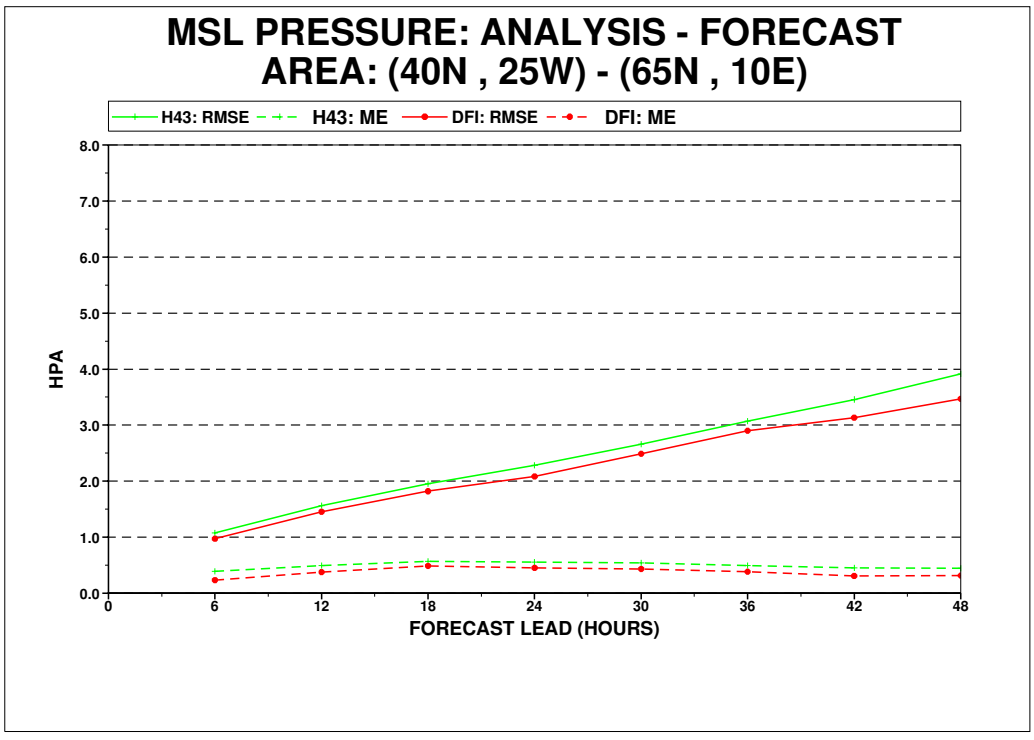


Figure 4.3: *Root-mean-square (solid) and bias (dashed) errors for mean sea-level pressure, as a function of forecast time, averaged over all thirty forecasts. Lines marked with crosses are for the reference run (NMI); those marked with spots are for the DFI run.*

Chapter 5

Input and output of the HIRLAM model

5.1 Observational Data

The HIRLAM analysis system requires all observational data to be in BUFR format (ref: Manual on Codes, International Codes, Volume L.2, WMO-N0. 306). Data extraction is based on a table system, the tables providing a key to enable the unpacking of selected elements from the observations. Plain text versions of tables are provided either by the data producers or the data processing Centres and are freely disseminated within the meteorological community. A set of tables, covering a wide range of observation types, are included with the HIRLAM suite. Tables are regularly revised and expanded to keep pace with changes in coding practices and to retain backward compatibility with older data types.

5.2 Physiography and Climate Data

The physiographic data base for HIRLAM originates in several different data sets from different sources. Some sources were already utilized in older versions of the HIRLAM system (Källén, 1996). Only a part of the old data sources is, however, still in use. The complete list of the current data sources is given below:

- Global 30 Arc Second Elevation data (GTOPO30), (USGS, 1998)
- The Global Land Cover Characteristics (GLCC), (USGS, 1997)
- Surface temperature, (Crutcher and Meserve, 1970; Taljaard et al., 1969)
- US NAVY Tape, (Joseph, 1980)
- Sea surface albedo, (Geleyn and Preuss, 1983)
- Precipitation climate data, (Jaeger, 1976)
- Surface soil moisture data, (Mintz and Serafini, 1981)
- SST climate data, (Alexander and Mobley, 1974)
- FAO-UNESCO soil data base, digitized for Europe, (FAO-UNESCO, 1987)

In addition, local data sources may be added (Sattler, 1999). The exploitation of the data sources is depicted in Fig. 5.1. The utilized horizontal resolutions of the data sources are given in the figure. The original data sources are not used in the climate generation of HIRLAM because of their heterogeneity in data representation as well as in data format. They are preprocessed instead, and a number of new fields are derived during this step (see Fig. 5.1). This procedure reduces the size of the data representation. The resulting Climate Data Base, CDB) is in a uniform Hierarchical Data Format (HDF) (NCSA, 1998) and is represented in a geographical mapping, that may be regular or rotated. The CDB supplies the input to the climate generation of HIRLAM. The most important steps when building the CDB are described in the following paragraphs.

It is important to note that the CDB is an open database and thus contains data, that originate from different meteorological centres, and which therefore may be preprocessed differently. The data format and the map representation are, however, standardized.

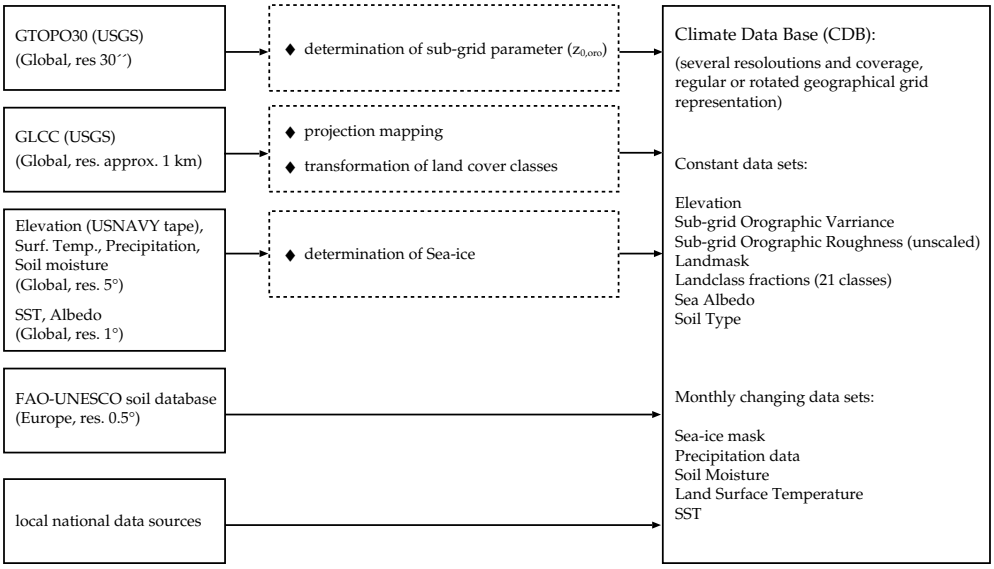


Figure 5.1: The physiographic data sources and their utilization. The original data are transformed if necessary and a unified climate data base is built in a preliminary step. The climate data base is an open data base and makes up the input to the HIRLAM climate generation.

Most of the original data sources are available in a regular geographic map projection. The GLCC data however is available in an equal area projection and transformed to a geographical grid. GLCC contains different land classification schemes, of which the *Global Ecosystems Legend* (Olson, 994a,b) is used to create a reduced classification scheme (Sattler, 1999), which originates in the classification scheme used by Wilson and Henderson-Sellers (Wilson and Henderson-Sellers, 1985). This reduced and modified Wilson and Henderson-Sellers classification scheme (RWHS) contains the following 21 land classes:

cropland	forest/fields	semi-desert
cropland irrigated	forest mixed	desert
grassland short	evergreen needle	ice
grassland tall	deciduous needle	bogs
evergreen shrubs	evergreen broadleaf	water
deciduous shrubs	deciduous broadleaf	ocean
field/woods	tundra/wetland	urban

In addition to the elevation, that is retrieved from GTOPO30, an unscaled isotropic sub-grid orographic roughness length $z_{0,oro,unsc}$ is determined with a modified formulation that goes back to Tibaldi and Geleyn (Tibaldi and Geleyn, 1981):

$$z_{0,oro,unsc} = \frac{1}{2}\sqrt{\frac{n_p + 0.001}{A}} var_h. \tag{5.1}$$

n_p is the number of relative height maxima within the actual grid square, which covers the area A. The variance var_h refers to the same grid square and is determined with

$$var_h = \frac{1}{n}\sum_{i=1}^{n} h_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} h_i\right)^2. \tag{5.2}$$

h_i denotes the height of pixel i and n the number of pixels within the actual grid square. Supplementarily to $z_{0,oro,unsc}$, the variance is included in the CDB together with the information about the average number of pixels that where used in the calculation of var_h.

An alternative data set for $z_{0,oro,unsc}$ is the variance $var_{23\times 23}$, which refers to a grid of 23×23 pixels of the GTOPO30 grid. At least one of these two data sets is available within the current CDB.

The Sea-ice fraction $f_{sea-ice}$ is deduced from SST with the following simple function:

$$f_{sea-ice} = \begin{cases} 1, & SST < -1.7\,^\circ C \\ -0.5822\,SST, & -1.7\,^\circ C \leq SST < 0\,^\circ C \\ 0, & 0\,^\circ C \leq SST \end{cases} \tag{5.3}$$

This Sea-ice fraction field can be used, if no sea-ice analysis is available. The other data in the climate data base can be directly retrieved from the original data sources.

The task of the climate generation is to create a set of climate data fields in the desired HIRLAM grid in GRIB format on the basis of the fields from the CDB by using interpolation and different averaging and selection schemes.

A typical set of parameters created in the climate generation is shown in Table 5.1. The actual set depends on the configuration of the climate generation. Some of the parameters like SST or soil water content are just interpolated from data of the CDB without any further processing. Others like geopotential, albedo and roughness length are deduced from data of the CDB.

Those parameters, which are derived from the RWHS land classes (see Table 5.1), are described in parameter tables, which can be adapted to different surface schemes. Currently, a set of parameter tables exists within the HIRLAM reference system for a simple surface scheme and for the ISBA scheme (section 3.6).

The derived parameters are described in the sequel. The geopotential is determined by the simple relation

$$\Phi = gh, \tag{5.4}$$

Table 5.1: List of typically created fields during the HIRLAM climate generation.

parameter	comment
surface geopotential	
surface albedo	derived from RWHS classes over land
surface roughness	derived from RWHS classes and $z_{0,oro}$, monthly
surface temperature	monthly
sea surface temperature	monthly
deep soil temperature	monthly
climatic deep soil temperature	monthly
soil water content	monthly
deep soil water content	monthly
climatic deep soil water content	monthly
snow depth	monthly
soil type	currently only for Europe
fraction of land	derived from RWHS classes
fraction of lakes	derived from RWHS classes
fraction of ice	derived from RWHS classes
fraction of no vegetation land	derived from RWHS classes
fraction of vegetation	derived from RWHS classes
fraction of forest	derived from RWHS classes

where g is the gravitational constant and h the model grid height.

The surface temperature and the soil temperatures are determined as a monthly mean. They are derived from the monthly surface temperature fields from the CDB, which refer, however, to the heights of the US NAVY Tape. So the actual mean surface temperature for month m on the desired HIRLAM grid is determined with

$$T_{s,m} = T_{usnav,m}\Gamma\left(h - h_{usnav}\right), \tag{5.5}$$

where Γ denotes the adiabatic lapse rate of the standard atmosphere and h_{usnav} is the height from the US NAVY Tape interpolated to the desired grid. The monthly mean deep soil temperature $T_{d,m}$ and the climatic deep soil temperature $T_{c,m}$ are derived from $T_{s,m}$ by considering a phase lag within the soil layer (Bringfelt et al., 1995):

$$T_{d,m} = 0.1\overline{T_s} + 0.9\left(0.2T_{s,m-1} + 0.8T_{s,m}\right) \tag{5.6}$$

$$T_{c,m} = 0.23\overline{T_s} + 0.77\left(0.5T_{s,m-1} + 0.5T_{s,m}\right). \tag{5.7}$$

$\overline{T_s}$ denotes the annual mean of the surface temperature.

Snow depth is derived from precipitation cliamte data P and surface temperature T_s, and the procedure goes back to (Brancovic and van Maanen, 1985). The derivation is performed after P and T_s have been interpolated to the desired grid. It makes use of several empirical functions. The first one is a snow creation function, which determines the fraction of precipitation that is

assumed to fall as snow:

$$C_{sn} = \begin{cases} 1, & T_s < T_{s,sn} \\ 0.5\left(1 - sin\left(\pi \frac{T_s}{(T_{s,r} - T_{s,sn})}\right)\right), & T_{s,sn} \leq T_s < T_{s,r} \\ 0, & T_s \leq T_{s,r}. \end{cases} \tag{5.8}$$

$T_{s,r}$ is the surface temperature, above which only rain occurs. It is assumed to $10^\circ C$. $T_{s,sn}$ is the corresponding temperature, below which only snow is assumed to fall. It is set to $-10^\circ C$.

The second empirical function used in the determination of snow depth represents the monthly amount of snow melt:

$$M_{sn} = \begin{cases} 0, & T_s < T_{s,sn} \\ 0.3125(T_s - T_{s,sn})^2, & T_s \geq T_{s,sn}. \end{cases} \tag{5.9}$$

The snow depth $d_{sn,m}$ for month m is then calculated in an iterative process using

$$d_{sn,m} = 10 P_m C_{sn,m} - 15 M_{sn,m} + 0.95 d_{sn,m-1}, \tag{5.10}$$

which is performed over a number of years. Regions with permanent snow cover are taken into account. More details on this procedure can be found in (Brancovic and van Maanen, 1985). T_s is set to $0^\circ C$, where $d_{sn} > 0$.

The determination of roughness length is different, depending on what surface scheme is chosen. There are, however, the two basic components influencing roughness in the model representation: roughness arising from sub-grid orographic variance and roughness due to vegetation or surface type. The former can be determined from the respective unscaled field of the CDB (Eq. (5.1)) using a logarithmic scaling:

$$z_{0,oro} = a \left(z_{0,oro,unsc}\right)^B, \tag{5.11}$$

where $a = 0.4038$ and $B = 0.715$, or alternatively:

$$z_{0,oro} = c \left(var_{23\times 23}\right)^D, \tag{5.12}$$

with $c = 0.4038$ and $D = 0.7767$. The current HIRLAM reference system climate generation makes use of the latter approach.

The vegetation/surface type roughness $z_{0,veg}$ is described via parameter tables, one for each month, depending on the surface scheme. For the simple surface scheme, the monthly vegetation roughness values for the RWHS land classes are listed in Table 5.2. Values for water are not included in this table, because roughness over water is determined in HIRLAM using the Charnock formulation (see section 3.6).

Sub-grid orographic roughness and vegetation/surface type roughness may finally be combined according to Tibaldi and Geleyn (Tibaldi and Geleyn, 1981):

$$z_0 = \sqrt{z_{0,oro}^2 + z_{0,veg}^2}. \tag{5.13}$$

Albedo over land is determined like vegetation/surface type roughness by parameter tables, describing values for each RWHS class. Parameter tables exist for annual means as well as for monthly means. The parameter table used for the simple surface scheme, which uses annual means, is shown in Table 5.3.

Table 5.2: Monthly values for vegetation roughness in meter for use with a simple surface scheme.

RWHS land class	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
cropland	0.018	0.016	0.013	0.012	0.024	0.042	0.06	0.06	0.054	0.042	0.024	0.018
cropland irrigated	0.018	0.016	0.013	0.012	0.024	0.042	0.06	0.06	0.054	0.042	0.024	0.018
grassland short	0.012	0.013	0.014	0.015	0.018	0.019	0.02	0.02	0.018	0.016	0.014	0.013
grassland tall	0.051	0.056	0.061	0.064	0.08	0.098	0.1	0.099	0.84	0.065	0.06	0.056
evergreen shrubs	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
deciduous shrubs	0.04	0.047	0.052	0.056	0.073	0.096	0.1	0.099	0.08	0.057	0.051	0.047
field/woods	0.051	0.051	0.050	0.050	0.062	0.080	0.097	0.096	0.89	0.077	0.059	0.052
forest/fields	0.283	0.306	0.306	0.312	0.327	0.345	0.356	0.345	0.338	0.323	0.301	0.028
forest mixed	0.69	0.717	0.745	0.763	0.782	0.8	0.8	0.772	0.763	0.745	0.717	0.69
evergreen needle	1	1	1	1	1	1	1	1	1	1	1	1
deciduous needle	0.812	0.85	0.912	0.95	0.975	1	1	0.962	0.95	0.912	0.85	0.812
evergr. broadleaf	2	2	2	2	2	2	2	2	2	2	2	2
decid. broadleaf	0.667	0.693	0.738	0.764	0.782	0.8	0.8	0.773	0.764	0.738	0.693	0.667
tundra/wetland	0.02	0.02	0.02	0.02	0.02	0.027	0.033	0.04	0.03	0.02	0.02	0.02
semi-desert	0.04	0.05	0.05	0.06	0.07	0.09	0.1	0.1	0.08	0.06	0.05	0.05
desert	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
ice	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
bogs	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
urban	2	2	2	2	2	2	2	2	2	2	2	2

Table 5.3: Anual values for vegetation albedo of the RWHS classes for use with a simple surface scheme.

RWHS land class	Albedo	RWHS land class	Albedo
cropland	0.18	deciduous needle	0.16
cropland irrigate	0.18	evergreen needle	0.16
grassland short	0.20	evergreen broadle	0.14
grassland tall	0.20	deciduous broadle	0.16
evergreen shrubs	0.18	tundra/wetland	0.18
deciduous shrubs	0.17	semi-desert	0.20
field/woods	0.18	desert	0.34
forest/fields	0.17	ice	0.55
forest mixed	0.16	bogs	0.14
urban	0.20		

The soil type is currently derived from the CDB by identifying the most frequent soil type class within the HIRLAM grid point. A list of the soil type classes is given in Table 5.4.

Concerning the fraction fileds of surface types, there is a difference between the surface schemes, too. Land fraction, lake fraction and the ice fraction fields are interpolated from the CDB grid to the HIRLAM grid in any case. The other fraction fields are created depending on the surface scheme, for which the climate fields are to be created. They are derived from the RWHS classes and described in a parameter table, which is listed in Table 5.5.

5.3 Postprocessing

The output produced by Hirlam falls into two general categories. The data needed for starting a new run is automatically stored in so-called history files. The postprocessing package serves to provide flexible control over the output of additional model variables and related data.

Postprocessing provides horizontally distributed data in the GRIB format and data from selected

Table 5.4: Soil type classes of the FAO-UNESCO soil data

class	soil type	class	soil type
1	sand	6	equal parts of sand and loam
2	loam	7	peat
3	clay	8	ice
4	equal parts of sand and loam	9	rocks
5	equal parts of clay and loam	0	missing value

Table 5.5: Derivation of the land use fraction fields as for example used within the ISBA scheme from the RWHS classes

RWHS land class	forest	low vegetation	no vegetation
cropland	0	1	0
cropland irrigated	0	1	0
grassland short	0	1	0
grassland tall	0	1	0
evergreen shrubs	0	1	0
deciduous shrubs	0	1	0
field/woods	0.05	0.95	0
forest/fields	0.25	0.75	0
forest mixed	1	0	0
evergreen needle	1	0	0
deciduous needle	1	0	0
evergreen broadleaf	1	0	0
deciduous broadleaf	1	0	0
tundra/wetland	0	1	0
semi-desert	0	1	0
desert	0	0	1
ice	0	0	1
bogs	0	1	0
urban	0	0	1

grid points in a BUFR-like format. The latter stream can be requested with a time resolution of the dynamical time step, and is thus suited for generating meteograms. Technical aspects of the postprocessing are described in Källén (1996).

Horizontally distributed data in the GRIB format may be requested in multiple streams, having their own selections of variables, levels, sub-domain, and write-up times. Variables that are available through postprocessing are shown in the tables of Appendix 4.

The user may request postprocessed data on model levels or as interpolated or extrapolated to constant pressure surfaces. Variables are interpolated linearly in $\ln(\eta)$, where η is the hybrid coordinate.

Above the highest model level or below the lowest level most variables are simply kept constant at the value preailing on the highest or lowest model level, respectively. Exceptions to this rule are made for geopotential, temperature, winds, and the vertical pressure-coordinate velocity ω.

The mean sea level pressure (p_{msl}) is calculated assuming below the surface a dry hydrostatic atmosphere with a constant lapse rate of 6.5 K/Km. The lapse rate is modified for very warm

or cold surface temperatures. If the surface geopotential equals zero, mean sea level pressure equals the surface pressure p_s.

5.4 Model Diagnostics

Accurate prediction of weather phenomena and the conditions in general near the ground, even in the short range, call for a correct simulation of the processes involved in the energy and water cycles, notably turbulent and radiative transfer, condensation, cloud and precipitation formation. Physically sound formulations for the exchange of momentum, heat and water between the atmosphere and the surface are also essential for improving the forecasts by increasing the horizontal resolution in order to resolve fine scale topographic features. From another point of view, parameterized fluxes and exchange rates are potentially valuable output of a NWP system. For instance, NWP data are likely to replace in the near future other estimates of surface fluxes and boundary layer parameters in agricultural applications and atmospheric pollution dispersion modelling.

Fluxes and exchange rates, unlike state variables, are not routinely observed and are not assimilated. Hence they are generally not subject to direct verification, but must rather be validated by indirect means. Parallel experiments with different model configurations are an often used method to evaluate new model components. Typically the different configurations are intercompared in terms of their forecast verification statistics. While this procedure is suited for ranking different model configurations, it does not provide direct estimates of the accuracy to which any particular process is simulated. Sometimes a scrutiny of parameterized processes directly may help spotting and rectifying model deficiencies, e.g. Trenberth (1997).

In the Hirlam 5 system the user may specify the output of a variety of fields related to parameterized processes. This data includes surface fluxes of momentum, heat, and water, radiative heat fluxes at the outer boundary of the atmosphere, and three-dimensional distributions of selected parameterized effects. In order to make accurate budget studies possible, all fluxes and tendencies are stored in the form of time integrals over the forecast length. Output of these fluxes and tendencies is controlled by namelist and is performed as part of the postprocessing.

5.5 Verification

The distribution of development work over the meteorological institutes participating in the Hirlam project makes a common verification framework a necessity. Results of experimentation with the Hirlam system have to be reproducible and comparable throughout the stages of the development process and from one research group to the next or even within one geographically distributed research group. Furthermore, as the Hirlam NWP system has been developed for implementation in all the participating countries it must cater for a variety of operational and climatic conditions. To separate these conditions from the intrinsic qualities of the model and to allow the institutes and the project as a whole to learn from each other's implementation a common framework for model intercomparison is a prerequisite. Finally the management of cooperative research and development requires a consistent and well-defined approach to the monitoring of model quality. For these reasons a framework for verification is an integral part of the Hirlam reference system. This common verification framework comprises a standardization of verification methods on standardized sets of reference data. This report details the background and the design considerations behind the implemented verification procedures.

5.5.1 The purpose of verification

The purpose of verification is to answer questions about the quality of forecasts. The way to answer these questions, i.e. the verification method, depends as much on the background of the questions and the application of the answer as it does depend on the intrinsic characteristics of the forecasts. Brier and Allen (Brier and Allen, 1951) divide the reasons for forecast verification in three categories: administrative, scientific and economic.

The justification of all NWP modelling lies in the economic value of the resulting forecasts. Nevertheless, economic value is not a direct driver for model development. Rather, the value of the model is implied by the objective of the project to deliver a state-of-the-art model, i.e. a model that is consistent with current meteorological knowledge and advanced modelling expertise.

The administrative verification in the context of the Hirlam project monitors the progress of forecast quality. This monitoring of forecast quality provides an evaluation of the effectiveness of (the management of) the research and development work in the project and a justification of the modelling effort.

The main focus of verification in the Hirlam project is on the scientific development of the NWP system. The three major verification activities are: diagnostic verification, model validation and comparitive verification.

Diagnostic verification uncovers the strength and weaknesses of the model. The long term monitoring of model performance for administrative purposes may show up deficiencies under particular conditions. Diagnostic verification aims at tying these conditions to systematic model errors. As such it guides research and helps to set research priorities.

Model validation is the activity of the project scientists whereby the consistency of the actual model performance is checked against the improvements targeted by the model changes that were introduced. As with diagnostic verification validation deals primarily with conditional systematic errors. It is applied to selected periods where such conditions exist, i.e. to case studies.

Comparitive verification is the objective of so-called parallel tests. Proposed new releases of the Hirlam NWP system have to be compared to the current reference version over a collection of periods containing all relevant weather conditions. These comparisons may oppose the findings of validation tests. An improvement of one aspect of the model under restricted conditions is often not an improvement under all conditions. In such cases the new model is said to be 'not sufficient' for the old model, *c.f.* Murphy (Murphy, 1997). The reverse is certainly true, if model validation has already shown an improvement on certain aspects of the model. This basic uncertainty implies that the choice for a new model version is in part a value statement where improvements on one aspect of the model are weighed against the deterioration on others.

The fact that the method of verification depends on the type of question being asked and the fact that our measures of good and better are not always uniquely determined show that verification is not primarily a mechanical exercise with numbers. It is good to keep this in mind when reading further about the mechanics of verification in the Hirlam context.

5.5.2 Integration of observation and field verification

The multiplicity of verification methods and metrics, deriving directly from the multifaceted nature of the perception of forecast quality, may be divided in many ways in relation to for instance the forecast type, probabilistic or deterministic, the type of reference data, station observations or analyses, the type of performance, absolute or relative to a competing forecast system, or the approach chosen, measure oriented or distribution oriented. The division that is of

interest here is the distinction that is often made between point verification and field verification.

In point verification the forecast, in the case of NWP a field of a forecast variable on a regular grid, is interpolated to the position of an observing station and compared to the observation at that station. In field verification the NWP forecast field is compared to an analysis field, or similar, on a common grid.

There are two principle differences between these two approaches:

1. Observations contain information about all scales of the atmospheric motion, in particular about local variations that are not resolved on the model grid, whereas the scales in the analysis and forecast fields are matched by comparing them on a common grid. In the observation verification part of the difference between forecast and observation is then due to the mismatch between the representation of the forecast and that of the observations. This will be most prominent in the case of extreme events.

2. In observation verification the observations are treated as independent and, in general, so are the interpolated forecast values, so that the calculated differences between model and observations can be treated as statistically independent. In this view then the notion of (a collection of) observations or forecast values representing a correlated atmospheric phenomenon is lost. In both the analysis and forecast the atmospheric flow is represented as a field, not a collection of uncorrelated grid point values. Field verification, ideally, determines the match between the patterns in the forecast and in the analysis.

There are two ways to make the observation verification and field verification approaches at least methodologically equivalent. The first is to treat observations not as independent but as an irregular field and to analyse the scales of atmospheric motion that are represented by the observation network. The second way to integrate observation and field verification is to disregard the spatial patterns in the forecast and analysis field and apply the methods of observation verification to pairs of forecast and analysis values at individual grid points. While the first approach is preferable, because it is the more correct one, it has to rely on pattern evaluation and pattern matching techniques that are still very much under development, *c.f.* Ebert and McBride (Ebert and McBride, 2000). The second approach is therefore the more practical one and is most often applied. It can still give access to spatial patterns in the errors, but these can not be reduced to errors in the spatial patterns. When interpreting the results of either observation verification or field verification performed in this manner, especially when evaluating the significance of a comparitive verification, the neglected spatial correlations have to brought back into account.

5.5.3 Design considerations and implementation

The integration of observation and field verification by applying identical methods to forecast/observation pairs and to forecast/analysis grid point value pairs is the basis for the design of the common verification framework.

BUFR to GRIB

To achieve this unification in practice the observations are treated as a field on an irregular grid. They are translated from the BUFR format that is used in Hirlam to store observations to an extended GRIB format. Along with the field of observation values two extra fields are defined, one with the latitude values and one with the longitude values of the observation station's location. The Hirlam interpolation routines are adapted to allow interpolation from and to irregular grids. Thus the comparison can be made on any chosen grid, which is then treated as a collection of disjoint points.

Data

To guarantee that results from the verification in different research groups and in all the participating institutes are comparable only those observations that are available to all the institutes are used, i.e. observations that are distributed on the GTS. In particular surface observations from SYNOP stations and SHIPs and upper air observations from TEMP profiles are used. These data are retrieved from the BUFR observation file that is ingested to the model analysis. In principle this allows the comprehensive observation quality control of the analysis to be used to screen the observations. In the current implementation however the model analysis is used for a more crude quality control. Observations that differ more than a predetermined amount from the analysis, about 4 times the estimated standard deviation of the observations, are not used for verification.

A further standardization of the verification data set is realized by adapting the EWGLAM recommended list of observing stations as the basic reference for verification. This not only facilitates the comparison of forecast quality within the Hirlam project but also the exchange of verification results within the EWGLAM community.

Conditional selection of verification data

The Hirlam verification system allows for three types of conditional subsampling of the verification data set: by geographical area, by station list and by masking.

- Geographical subareas are defined by specifying the vertices of an enclosing polygon. Currently these areas are defined: Ireland & England, the Southern North Sea, the Northern North Sea, the Netherlands, Scandinavia, France, Spain and North Africa.
- Station lists are specified as a collection of station ID's. Currently the EWGLAM station list has been implemented. Alternative station list may be derived from other selection criteria applied to the observations, such as station height, station type or position within a river catchment area.
- Masking is at this time implemented as a general Hirlam utility outside the core verification package. Forecast fields are masked according to any union or cross-section of properties of climatology, forecast or analysis fields. This provides a flexible approach to complex conditional sampling. Simple masks are for instance the land-sea mask and the model height field. A more complicated mask could for instance be a combination of precipitation in excess of 0.3 mm/6 hr with temperatures between -2 °C and 2 °C.

Methods

The basic method of both observation and field verification is a location-by-location, or point-by-point, comparison. This implies that all the concepts and methods that have been developed for observation verification, *c.f.* Murphy (Murphy, 1997), are equally applicable to field verification. In the current implementation a measure-oriented approach is used. In a central statistics module verification scores are calculated by aggregation of the grid point scores over the verification subareas (see above.) At this time only three performance measures are applied: bias, rms and maximum.

The parameters that are verified at the surface are mean sea level pressure, station level pressure, ten meter wind speed, wind direction and wind vector, 2 meter temperature, 2 meter dewpoint temperature, 2 meter relative humidity, total precipitation and cloud cover. Parameters that are verified at standard pressure levels are wind speed, wind direction, wind vector, temperature, dew point temperature, relative humidity and geopotential.

The verification results are archived per level (surface or pressure level), subarea, forecast length and verification time. Along with each score the number of valid observations it is based on is

archived. This allows further aggregation, e.g. over a period of time. The defining characteristic of the verification method, dictated by the archiving strategy, is that scores are always calculated as areal averages. They cannot be reduced to individual points or observing stations.

5.5.4 Limitations and extensions

The unification of observation and field verification by treating both the observational network and the analysis fields in a similar way as collections of (disjoint) grid points allows for great flexibility in applying the extensive body of statistical methods that has been developed for observation verification equally to field verification. The major drawback of this approach is that information about the correlations of observations or model gridpoints is not retained, which is a principle obstacle to verifiying weather patterns. This method of verification is therefore not directly suited to answering questions about the quality of the representation by the model of the weather patterns and the appropriateness of the interpretation of model results in those terms. To do so would require a fundamentally different approach based on pattern recognition and pattern matching techniques.

However, some understanding of the errors in the representation of patterns may be gained using the point verification method if at least the spatial patterns of the errors are conserved and conditional sampling of these errors is applied. Using this technique for instance Cameron and Milton (Cameron and Milton, 2001) are able to show systematic errors in extratropical cyclonic development in the UK Met Office Global Model. To achieve this within the current Hirlam verification framework an alternative strategy for archiving, namely storing errors as fields, has to be implemented.

The flexibility that the verification framework offers is at the time underused. More may be learned about critical aspects of the model if additional performance measures are implemented, in particular skill measures, distribution oriented measures for precipitation verification and field oriented measures such as the anomaly correlation and the tendency correlation.

For comparitive verification it is furthermore important that for each of the performance measures confidence intervals are derived and presented alongside these measures whenever a comparison is made. Thought has to be given as to how to reintroduce into these confidence intervals the spatial correlations that are neglected in the point verification method.

The limitation of the verification data set to data distributed on the GTS is a barrier for the areal verification of precipitation and for true mesoscale verification. Within the verification framework a provision has to be made to include high resolution data sets, such as precipitation from national rain gauge networks or from radar, that are not included in the observation file that is used by the analysis.

The current data quality control is minimal. The exclusion of data with large deviations from the analysis precludes the verification of extreme errors and thereby possibly the verification of extreme events. This may be a problem for mesoscale forecasts. A better quality control of observations is therefore needed. This may be based on the analysis data quality control for those observations that are used by the analysis or on similar quality checks being applied to non-analysis data.

The purpose of data quality control is to separate problems with the data from model errors. Taking this one step further one could try to separate the intrinsic model errors from the errors in the verification data and from the errors due to the differences in representation between the model weather and the observed weather. The best way to achieve this is probably to analyse the observations at the scale of the weather patterns to be studied and to match the model forecasts to this scale before comparing the two.

Eventually the verification method that seems most appropriate to the verification of model

forecasts, at least in terms of the concepts that define the weather as it is understood, is by field verification, based on analysed observations using pattern recognition and pattern matching. This method can not be implemented as an extension to the current verification strategy, but it requires a completely new development. While this development is undertaken, a lot may still be learned that supports the advancement of Hirlam by further exploitation of the current common verification framework.

Chapter 6

The HIRLAM system

6.1 Introduction

A system that is intended to be used for operational weather forecasting consists of more components than just the analysis and forecast modules. Equally essential are components to embed those two modules in existing environments, like operational forecasting procedures, or requirements posed by the available hardware. Probably even more important for proper embedding in an operational and research environment is the availability of good, easily accessible, documentation.

The HIRLAM activities in the field of embedding can be divided into two groups:

- Research and development of meteorological and technical components
- Maintenance and development of a portable system to run the HIRLAM components

Within the HIRLAM project, activities of the first kind include the design and development of procedures for writing and generating documentation and for verification and diagnostics, and of methods to efficiently use the available hardware. Activities of the second kind are the maintenance and development of the software to run the HIRLAM system. Included in this are adaptations of the software to the continuous changes in the environment, such as the introduction of new data base formats, or the millennium problem. These activities are usually referred to as 'system management'.

The activities mentioned above will be discussed in the following subsections.

6.2 Documentation

In a distributed organisation like the HIRLAM consortium, the existence of well-accessible, relevant, and up to date documentation is more important than ever. This is true for all levels of documentation, ranging from the science to the program level.

The top level of scientific documentation consists of articles in the open literature. A powerful means of communications is the HIRLAM news letter, that appears 4 times a year. Many of the articles in it serve to document modifications to the HIRLAM system, both scientifically and technically.

An extensive set of technical documentation notes are available over the HeXnet. The HeXnet is the HIRLAM extranet, a WWW application over the internet, accessible from the participating institutes. These notes include instructions for local installation of the reference system, release

notes for every new release. Also, over the HeXnet reports of meetings are available, and a few discussion groups have been established on it.

The lowest level of technical documentation consists of (Unix) manpages to every program module in the HIRLAM system. These manpages are in HTML format; they contain many cross-links to the source code of the module that they describe. For this purpose, a pre-processor converts the source codes to HTML pages, with the reverse cross-links.

In spite of the existence of conventions for the format of scientific and technical documentation, and for the actual (mainly Fortran) coding, it proves virtually impossible to impose them, with so many researchers, from so many different institutes, involved in the HIRLAM project. To reduce the negative impact this has on code exchangability, HIRLAM is investigating methods to assist researchers and developers in writing good documentation, and to increase accessibility of codes written by someone else. The latter has resulted in the development of a Fortran-77 code browser, written in Java. One of its extensions currently under development consists of utilities to write codes and proper, relevant, manpages simultaneously. Methods to extend this to include scientific documentation are also researched. However, we believe that the only way to obtain a source code and at the same time the scientific documentation that is really valid for the source code, is by generating both from a program specification.

Although much documentation on how to use the HIRLAM system is already available, many a user finds it difficult to properly use it. To help users to set up their experiments, and to inspect the output of their HIRLAM runs, a graphical user interface is under development.

6.3 System Management

6.3.1 Requirements on the system

The HIRLAM system has been designed to support the main HIRLAM objectives: to provide an operational system for, and to provide a tool for research into, short-range, limited-area, numerical weather forecasting.

The operational system must be applicable in the HIRLAM member institutes. Because those institutes apply different hardware configurations for their operational HIRLAM systems, a consequence is that the HIRLAM system must be easily portable to a variety of hardware architectures. Also, because the HIRLAM members have their own formats for the observations data base, there should be a clear interface to input the observations. Similarly, there must be a clear interface to the data bases that hold the meteorological fields.

In addition to the operational applications, the HIRLAM system should be available for experimentation. On the one hand, experiments are carried out at the member institute's own hardware. On the other hand, the HIRLAM member states often are keen to use their allocation of ECMWF computer resources for HIRLAM experiments (almost all HIRLAM states are also member states of ECMWF). So the HIRLAM system should also use ECMWF resources efficiently.

Obviously, it saves on maintenance costs, and it supports integrity of the system, if all applications (operations, experiments at the home computer, experiments at ECMWF) can be achieved with essentially the same system. There is, however, a conflict between, on the one hand, demands for a stable system, posed e.g. by operational applications and by experimenters during the course of their experiments, and, on the other hand, demands for an up-to-date system, e.g. posed by experimenters at the start of their experiments. A version control system, implemented recently, tries to resolve this conflict.

6.3.2 The Reference System

There is a standard version of HIRLAM, referred to as the 'reference system', kept on the ECMWF computer system. The reference system is maintained by the 'HIRLAM system manager'.

To install the HIRLAM system on a local computer, a copy of the reference system should be obtained. HIRLAM members can create their own copy if they have access to ECMWF, other HIRLAM users should obtain it through the HIRLAM system manager. This copy contains instructions to install the system. The first step is the manual editing of a file, which serves to communicate properties of the available hard- and software (computer type, directory structure, etc.) to the HIRLAM system, by setting a number of environment variables. After that, the installation procedures are fairly straightforward. They are concluded with a test run, which serves both to complete the installation, and to verify the correctness of the installation.

6.3.3 Portability of the codes

To achieve maximum portability, HIRLAM source codes are almost fully standard Fortran-77, with a few, generally available, extensions (namelists and automatic arrays). A few routines are coded in C, but to avoid problems in the interface to Fortran routines,, they do not have strings ('character variables') in their argument lists, if calling or called by a Fortran routine. The Unix scripts are written in Bourne shell syntax. Documentation is in HTML, with some LATEX, WordPerfect, or postscript still around.

Reasonable efficiency on all currently available hardware is achieved by writing the compute-expensive parts of the HIRLAM system in Fortran vector code. Parallel execution is organised by automatic parallellisation on shared-memory machines, and by MPI on massively parallel hardware. In Engelen et al. (1997) it is demonstrated that vector codes may not achieve optimal performance on super-scalar hardware, with a performance loss of up to 30 % for specific code fragments, but the overall loss for HIRLAM is at most a few percent.

The current system has been designed to execute by default on a single platform, which, for acceptable turn-around, should be a state of the art compute server. Some parts, however, may not execute efficiently on a compute server. These parts include, on the one hand, actions to be taken only once, such as compilation of the codes, and, on the other hand, actions within the main execution stream, such as IO. Even though the one time actions would not delay operational production, their slow execution may be annoying for experimentation. These less efficient parts may take a relatively long time on a compute server, the more so because the compute intensive parts are so well optimised. Usually, computer systems that include a compute server also have other components that might be more suitable for those tasks than the compute server. Hence, a system to distribute HIRLAM execution over a computer network has been developed. This system, 'mini-SMS', is a subset of ECMWF's Supervisor Monitor Scheduler SMS. (Obviously, if compilation is to be executed on a) different machine, a cross-compiler would be required.

The suite of jobs in a full HIRLAM run comprises about 40 tasks. These have been organised into 'families', which leads to a genealogical tree of families and tasks (Fig. 6.1). Most tasks are only allowed to start after the tasks that provide their input have completed. This defines dependencies between the tasks. For example, the forecast model is dependent on the compilation of a large number of Fortran source codes. In Figure 6.2 the tree is shown again, but now re-arranged to display the dependencies between the tasks and families.

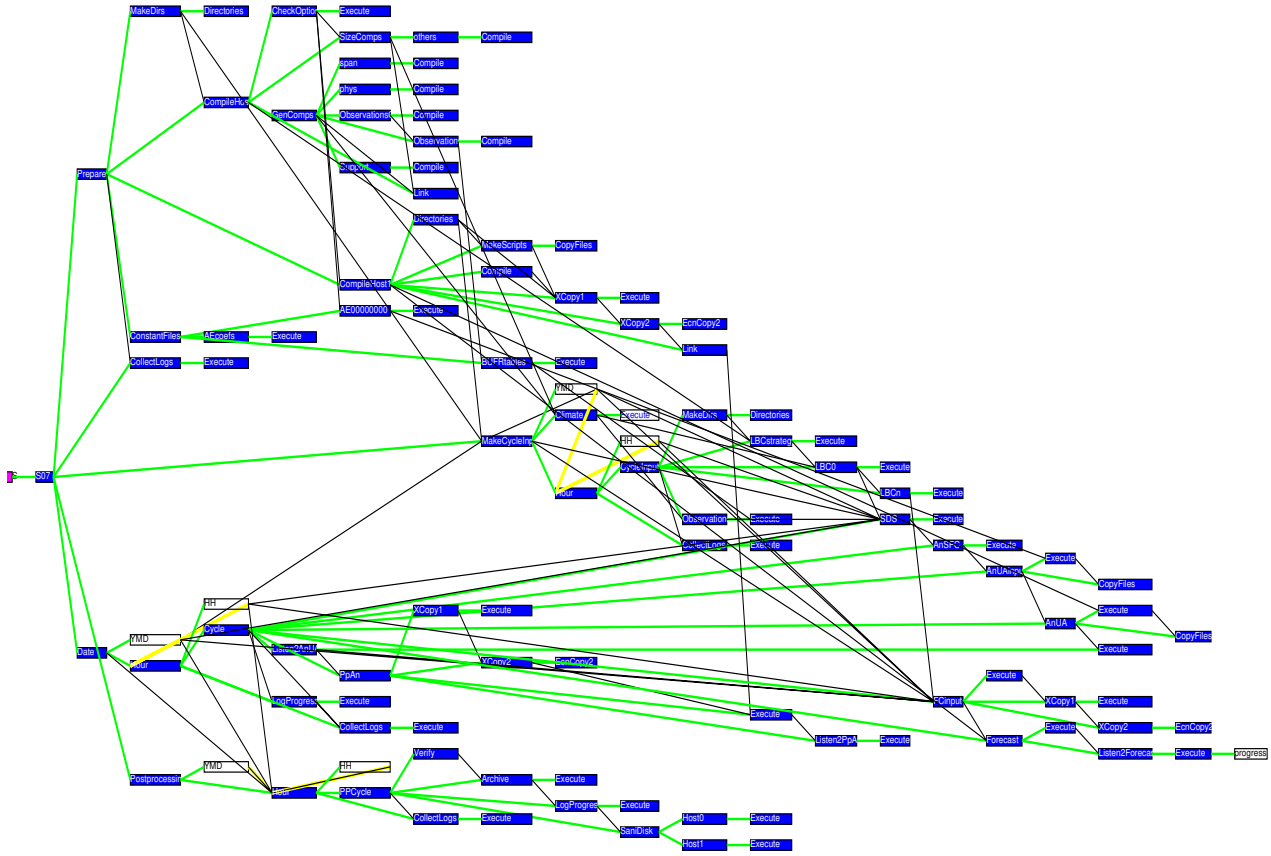


Figure 6.1: *The tree of HIRLAM tasks and families. Nodes are shown in dark (blue); end nodes are tasks. The open rectangles denote loop counters, e.g. to measure the progress of the forecast. Parent-child relations are shown by the moderately light (green) lines.*

6.3.4 Portability of the IO interfaces

By using standard WMO binary codes (GRIB, BUFR) for in- and output, the HIRLAM system has a clearly defined interface to local data bases for observations and fields. However, there are some restrictions:

Currently, the observations must be provided in a BUFR format that matches a fixed template for each observation type; the template is the one used by ECMWF for their observational data base. The reason to use this template is historical. It has the advantage that connection to the ECMWF data base, for experimentation at ECMWF, is easy; but the disadvantage may weigh heavier: HIRLAM users must have software to convert incoming messages (usually in WMO character codes, like SYNOP or AMDAR) to the BUFR template form. The requirement of having the observations in this template may be relaxed with the variational data assimilation scheme.

Fields, and notably those to specify the lateral boundary conditions, are input in the form of files, that contain GRIB messages. Internally, HIRLAM uses a data base structure, that

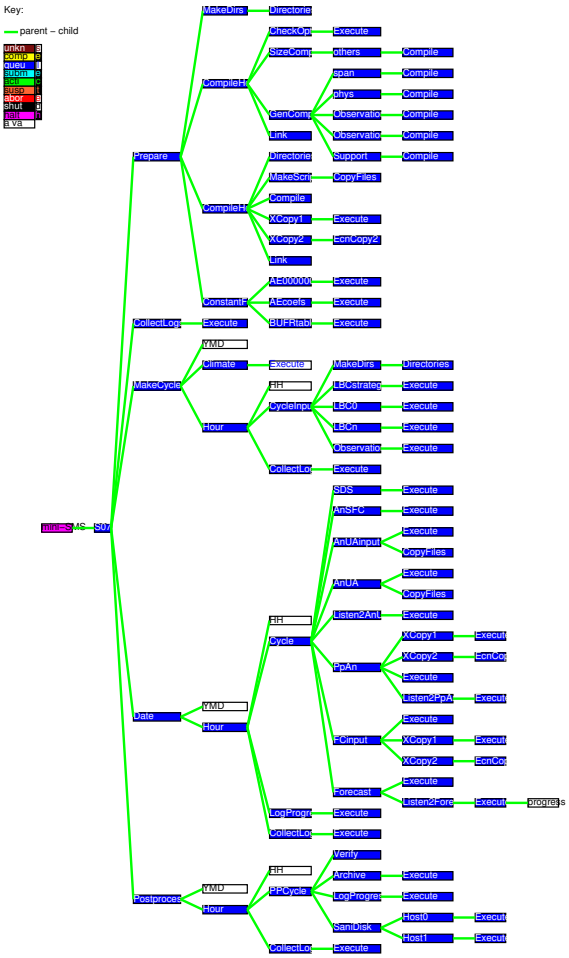


Figure 6.2: *As Fig. 6.1, but re-arranged to show dependencies between tasks and families. The dependencies are given by the dark (blue) lines. A family or task is only allowed to execute if all its dependencies are complete. The light (yellow) lines indicate that the node must not be executed if the dependency is met.*

adds administrative information to the files. This data base format does normally not destroy the GRIB structure, so no effort is made to remove that information when HIRLAM writes its output. On the contrary, the information may make subsequent applications process the HIRLAM output more efficiently. This is in particular true for applications that are part of the HIRLAM system, like the verification package, or indeed, the HIRLAM main stream for its subsequent analysis-forecast cycle.

A particular form of output is the time series file ('meteograms'), containing the value of some meteorological parameter at some grid points, for each model time step. WMO BUFR would be the appropriate format for this. Currently, however, HIRLAM does not comprise a utility to encode BUFR, Instead, the format of the time series is one resembling GRIB; it can be decoded with most standard GRIB decoding packages.

6.3.5 Version Control

The HIRLAM consortium maintains the reference system at ECMWF. Most participating institutes run a local copy of that system, with their own set of modifications, both for their operational numerical forecasting, and for their experiments to further develop the HIRLAM model. Those copies may have been taken a considerable time ago -two or more years is not unusual-, which is a consequence of the fact that one normally does not want to modify operational procedures frequently.

Now, when an institute decides to upgrade its local implementation, usually they would want a state of the art version, but still implement the local deviations to the reference system. Hence, a version control mechanism is called for. Moreover, the experimenters that use the reference system at ECMWF normally want to start from a recent version,. This requires that the reference system be updated frequently. On the other hand, once an experiment is running, experimenters prefer not to see changes in the system during their experiments. Here again, a version control mechanism is needed to resolve this conflict. And last but not least, such a mechanism is required to manage new developments, because, in the HIRLAM structure, those developments aredistributed over the participating institutes.

HIRLAM developed its own version control system, on top of the GNU revision control system (rcs). When a local copy is taken, that copy will be the latest version of the reference system. The owner of an experiment fixes the version number to be used at the start of the experiment, be it with the reference system at ECMWF, or with a local copy (default is to start from the latest version available; this is not necessarily the current reference level, because a local copy may have been taken some time ago).

Normally, an institute would upgrade the operational suite only after substantial parallel testing. The HIRLAM version control mechanism helps to guarantee that the suite, finally implemented, is indeed the same as was used for the parallel tests.

6.4 Parallelization

6.4.1 Data distribution

Originally the HIRLAM forecasting model was designed to run efficiently on shared memory vector machines. Rapid development in parallel processing made in necessary to enlarge the code to parallel computers. In fact, there exists two earlier parallel versions, which are used in operational environment, (Eerola et al., 1997; Bjørge and Skålin, 1995).

Data decomposition strategy is a crucial point on the distributed memory machines, because the data exchange between processors is much slower than computing. Also, in a code like HIRLAM, the decomposition strategy should be transparent on shared memory computers.

The basic data distribution model, used during the explicit dynamics and physics, is a two-dimensional latitude-longitude decomposition. Schematically, this is shown as phase 1 in Figure 6.3. Every processor holds one and only one sub-domain, and every sub-domain consists of contiguous vertical columns extending from the surface to the top of the atmosphere. Every sub-domain contains, if possible, the same number of grid-points, which assures a good load balance between processors. However, this does not guarantee an optimal load balance in the physics, where many switches make the amount of computations different in the different sub-areas even if the number of grid-points is the same. In the Eulerian dynamics HIRLAM uses second order finite differences in the horizontal and therefore a halo zone with a width of one row or column is enough, where data calculated by the neighbouring processor are used. In the Semi-Lagrangian version, the width of the halo zone is by default four, but can be changed if

necessary.

Figure 6.3: Data distribution during explicit dynamics (Phase 1), during the FFT (Phase 2) and during the Gaussian elimination (Phase 3)

In the semi-implicit part two other data decomposition strategies are needed, because of the way of solving the three-dimensional Helmholtz equation. This is done by decoupling the three-dimensional Helmholtz equation into a set of two-dimensional equations, which are solved using a direct method by applying Fast Fourier sine-transformation in x-direction and Gaussian elimination in y-direction.

During the Fourier transform or inverse Fourier transform phase, the data decomposition strategy shown as phase 2 in Figure 6.3 is used. Always, the whole row of data in x-direction is located in the same processor to allow a local Fourier transformation. Because there are no data dependencies in y- or z-direction, the number of rows can be made as equal as possible on different processors.

In the Gaussian elimination the data dependencies are only in the y-direction and the data distribution between processors looks as shown in phase 2 in Figure 6.3.

6.4.2 Data communications

From the previous discussion it is obvious that efficient data communication is needed, because all the data must be re-distributed between processors several times during every time-step. There exist several communication or message passing libraries. At the moment, MPI (Message Passing Interface) is regarded as a standard, and it has been implemented almost on every platform. The starting point in the HIRLAM parallel code is, that when running in shared memory mode, no communication libraries are needed. In the distributed memory environment, a real data communication is needed, and there are two basic options. The most general option is to use the GC-library, which is not a message passing library itself but offers a consistent interface to the most popular message-passing libraries (MPI, PVM, SHMEM, etc.). The parallel code has been most extensively used with the SHMEM-library, which is available on SGI and SGI/Cray computers. It is not a normal message passing system, which uses a pair of send/receive calls but it is one-sided in the sense that any individual processor can directly point to the memory of any other processor and put or fetch data. A version using the general GC-library is also available.

6.4.3 Flow of computations

In the distributed memory environment some extra steps are needed in the flow of computations compared to the original serial code.

The essential steps in an Eulerian run are:

1. Read the input parameters containing the number of processors used in the x- and y-directions and the dimensions of the grid.
2. Calculate the data decomposition strategies for phases 1, 2 and 3.
3. Read the input fields and distribute to each processor its own data. At the moment the reading and encoding of GRIB takes place on one processor.
4. If new boundaries are needed, read the boundary data as in step 3.
5. Calculate the explicit part of dynamics and exchange the data in halo zones with the neighboring processors when needed.
6. In case of explicit horizontal diffusion perform the fourth order diffusion by calculating Laplacian twice and exchanging data in the halo zones between the calculations.
7. Perform the semi-implicit calculations using the following steps:
   - data transposition from phase 1 to phase 2
   - FFT calculations in the x-direction
   - data transposition from phase 2 to phase 3
   - Gaussian elimination in the y-direction
   - data transposition from phase 3 to phase 2
   - inverse FFT calculations in the x-direction
   - data transposition from phase 2 to phase 1
8. Perform the physics computations using the phase 1 data decomposition.
9. If required do the implicit horizontal diffusion using the same data transposition as in the semi-implicit phase. Note that all the data transposition must be done for temperature, wind components, humidity and possible scalar fields.
10. Write out the model state and possibly the post-processed fields at given intervals
11. Repeat the phases 4 - 10 for every time-step.

A similar solution method has been developed also for the semi-Lagrangian integration.

6.4.4 Efficiency

An important measure of the efficiency of any code in parallel environment is scalability, ie. how the relative speedup is increased when the number of processors is increased. Figure 6.4 shows the scalability of the code on T3E system using SHMEM-library, when the number of processors increases from 32 to 192. The total number of grid-points in this case is 194*140*31. The straight line shows the ideal scalability, i.e. efficiency is doubled, when number of processors is doubled. In practice, this goal is very seldom achieved. The lower line shows, what is achieved with the parallel HIRLAM code. The efficiency value 4.6 when going from 32 processors to 192

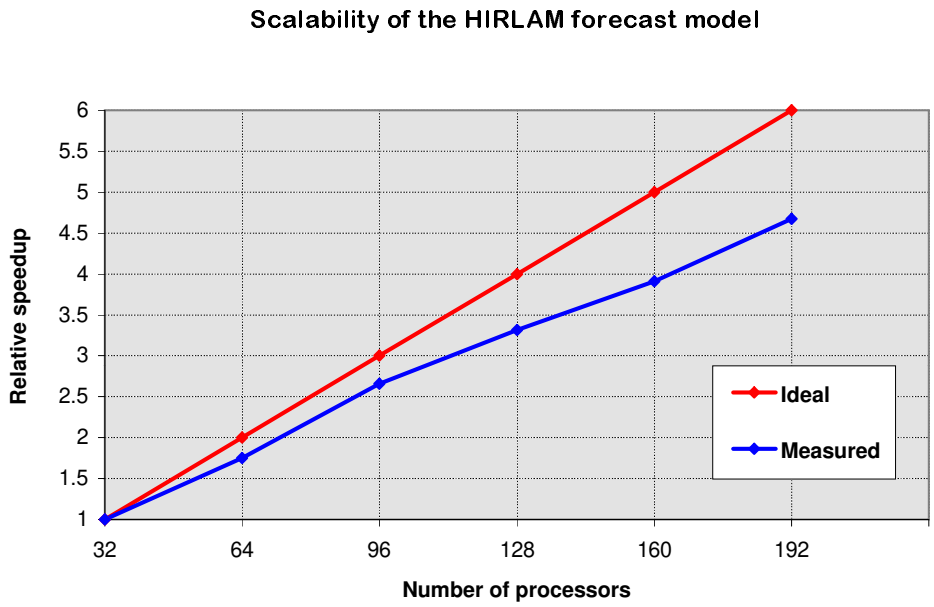


Figure 6.4: The efficiency of the MPP-version of the HIRLAM forecast model compared to the run with 32 processors. The upper curve shows the perfect scalability, while the lower one shows the result from the HIRLAM code.

processors, i.e. when the number of processors is increased with a factor of six, is a very good result. Note that this efficiency refers only to computations. Input/output is at the moment the bottle neck in all massively parallel systems.

Chapter 7

Applications

7.1 Delayed mode runs

At ECMWF, a quasi-continuous run with the current HIRLAM reference system is maintained. The schedule is to run eight 3-hours cycles per day, and a 48 hours forecast from the midnight (00 GMT) analysis. The area over which the model is run is shown in the Figure. The resolution is 0.5. The number of grid points is 164 by 130, with 31 levels. The system is identified by the acronym DMR.

Although the system runs in 'delayed mode', often the results will be available in time to be valuable for evaluation of the operational forecast made with your own, national HIRLAM system. However, no guarantee of timely delivery can be given. The system has now been running since early February 2000; results became available on a quasi-regular basis from 21 February onwards. Some products are displayed on the HeXnet, see also their description.

The DMR run tries to simulate operational conditions as much as possible. A consequence of this is that it is not possible to extract the lateral boundary conditions from the ECMWF archives: it is not possible to extract the data on a rotated grid, and it is normally not allowed to access the data of the midnight run. To simulate operational conditions it was decided to use the same lateral boundary conditions as used by the corresponding forecasts and analyses of the operational forecasts at KNMI. After this decision, forced upon us by ECMWF archiving policies, it was a minor step to also use the observations from KNMI's data reception system. The advantage of this was that the runs with the DMR system can usually be delivered in time for evaluation against operational forecasts. If, on the other hand, the data were to be taken from the ECMWF archives, the system would have to run with a delay of a few days, because, again, ECMWF does not allow access to the archives of observations until after three days.

Usually, the midnight run of the DMR system uses the preceding 12 GMT ECMWF forecast for the lateral boundary conditions. So does the analysis of 3 GMT. All other analyses use the 00 GMT ECMWF forecast. If the ECMWF forecast get delayed, the preceding ECMWF forecast will be used. If more than 2 successive ECMWF forecasts are not delivered in time, the system will be interrupted, see below.

The system produces 'history files', containing the full model state, and postprocessed files for the analysis, the guess-fields, and the forecasts at 6 hours intervals. There are three sets of postprocessed files: the 'model diagnostics' files (file titles have the suffix md) with the fields used by the verification and the diagnostics packages; 'hourly diagnostics' files (suffix -s) with some single level budget fields, and 'display' files (titles have the suffix pp) with a few fields that are used for routine monitoring of the system, by display on the HeXnet. Furthermore, the data assimilation scheme produces the so-called statistics files. All these files are stored at ECMWF,

in the ecfs directory ec:/rui/hirlam/DMR. The hourly diagnostics files are stored in tar files, with suffix sf.tar.

To facilitate routine monitoring of the system, the 'display' files are sent to KNMI. The ECMWF plotting package 'MAGICS' is used to plot the following data over an area covering the HIRLAM countries fully (except Greenland and the Canary Islands):

Near-surface fields: temperature at 2 m, pressure reduced to mean sea level, and wind at 10 m. In addition, with the analysis fields the surface observations are plotted, and with forecasts the precipitation over the past 6 hours (if available).

Upper air fields: temperature at 850 hPa, height of the 500 hPa level, and wind at 300 hPa.

Data monitoring information: mean sea level pressure of SYNOP/SHIP/BUOY and 500 hPa height of TEMP reports: positions of the data used or rejected by the analysis.

If for any reason the quasi-continuous run gets interrupted, the system will try to recover with a coldstart from a first-guess that is derived from KNMI's operational system. It will try to start with the first-guess of 8 data assimilation cycles before the current time. Hence, if it succeeds, at most 8 'warming-up' cycles will have preceded the 48 hours forecast.

7.2 Operational implementations

Seven of the HIRLAM member institutes have implemented HIRLAM operationally and have been running the system routinely for many years. The HIRLAM Reference system has evolved over the years and the member institutes upgrade their operational models and Data Assimilation system to the most recent version at intervals. Due to operational stringency and constraints, this can normally not be done at the same time as the Reference System is updated. Before a new operational version is implemented it goes through a validation procedure and will have to be of acceptable quality as well as perform well on the computer system at hand. For this reason, the different HIRLAM countries often run slightly different versions, see table 7.2 below. The table (and the following one) describe the status in November 2002.

The national implementations are all using different limited areas and a number of different nesting procedures, depending on the need of each nation's weather services. All use boundary values from ECMWF forecasts, up to 2001 from ECMWF 12 UTC and 00 UTC forecasts. From 2001 the so called Boundary Conditions Optional Project was modified to provide boundary values in frames covering the boundary zones 4 times a day. For this purpose ECMWF produces forecasts from short cut-off 3D-VAR analyses.

ECMWF boundaries provide boundary values for each member's large limited area resolving the synoptic scales over Europe, North Atlantic and surrounding areas. Theses forecasts then provide boundaries for nested models with 2-5 times higher resolution covering smaller areas which resolve some meso-scale effects.

Data assimilation is more and more being carried out with 3D-VAR, but OI is still a supported option and used by some of the members. All members carry out data assimilation in their large areas, but Denmark incorporates the large scales from ECMWF analyses twice a day in their re-analysis runs. Also Spain and Norway have re-analysis cycles, where the data assimilation is re-run after the long forecast run in order to incorporate effects of additional observations having arrived after the first analysis and , if available, also later boundaries from ECMWF. These two countries also re-initialise their data assimilation by starting from ECMWF first guess once a day. Most members also carry out data assimilation in their nested areas, except for the

[1] Denmark uses a 3-6 times longer physics time step than indicated above
[2] Norway uses physics version 2.6 (2002)
[3] Sweden used physics version 2.7 (2002)

Country	Res °	Gridp	lev	Adv	Δt	DA	Ini	Phys	bdup	Fclen	Fc/d
Denmark G	0.45	202x190	31	Eul	240[1]	3V	NMI	4.6	6	60	4
N	0.15	194x210	31	Eul	100[1]	3V	NMI	4.6	1	36	2
E	0.15	272x282	31	Eul	100[1]	3V	NMI	4.6	1	54	4
D	0.05	182x170	31	Eul	36	-	NMI	4.6	1	36	2
Finland	0.4	194x140	31	Eul	180	OI	NMI	4.6	6	54	4
	0.2	194x140	31	Eul	120	OI	NMI	4.6	3	54	4
Ireland	0.147	438x284	31	SL	225	3V	DFI	5.0	3	48	4
	0.12	222x210	40	SL	150	3V	DFI	5.0	3	24	4
Netherlands	0.2	406x324	31	SL	360	OI	DFI	5.0	6	48	4
	0.1	306x290	40	SL	180	OI	DFI	5.0	3	24	8
Norway	0.5	188x152	31	SL		3V	DFI	5.0[2]	3	60	4
	0.1	224x324	31	SL		-	DFI	5.0[2]	3	48	2
	0.05	152x150	31	SL		-	DFI	5.0[2]	3	48	2
Spain	0.5	194x100	31	Eul		OI	NMI	4.6	6	48	4
	0.2	194x100	31	Eul		OI	NMI	4.6	3	24	4
Sweden	0.4	202x178	31	SL	600	3V	NMI	4.6[3]	6	48	4
	0.2	162x142	31	SL	180	3V	NMI	4.6[3]	3	36	4

Table 7.1: General properties of the operational Hirlam configurations. Shows resolution in degrees, number of gridpoints in zonal and meridional directions, number of levels, advection scheme (Eulerian or semi-Lagrangian), time step in s, Data Assimilation cycle period in h, HIRLAM physics version, boundary update cycle, i.e. time in h between subsequent boundary data files, length in h of the main long forecasts (not for DA cycle only) and number of long forecasts per day.

0.1-0.05 ° resolutions in Denmark, Norway and the Netherlands, where the initial conditions are interpolated.

The observation usage varies a little, depending on the perceived impact of some satellite data and on the availability of non-GTS data sources like scatterometer adn ATOVS data. Some members also have gridded SST analyses available based on AVHRR, in-situ observations and hand analyses (in the Baltic). Table 7.2 shows the main features of the Data Assimilation systems.

7.3 Forecast examples

The COMPARE project of the World Meteorological Organization (WMO) has been designed to improve our understanding and predictive capability at the mesoscale through numerical model intercomparison experiments. So far, there have been three different exercises in which most of the state-of-the-art mesoscale numerical models have participated. Hirlam model has been run for the last two exercises.

COMPARE II (Georgelin et al., 2000) was a typical case of flow forced by orography. Simulations at 50, 25 and 10 km. resolution were carried out. Some systematic errors appeared in all the models: winds too weak, mountain wave-amplitude too large and lee vortices poorly represented. Hirlam gave a very accurate representation of the synoptic evolution which forced the orographic flow.

COMPARE III (Nagata et al., 2001) was a case of explosive development of a tropical cyclone. The explosive deepening of typhoon Flo (9919) over a period of 72 hours was the target of this

Country	Res °	DA	cyc	Reana	Reini	satob	satem	tovs	scatt	SST
Denmark G	0.45	3V	3	Yes	Ana	No				
N	0.15	3V	3			No				
E	0.15	3V	3			No				
D	0.05	-	3			-				
Finland	0.4	OI	6			Yes	Yes			Yes
	0.2	OI	6			Yes	Yes			Yes
Ireland	0.147	3V	3	Yes	No	No				
	0.12	3V	3	Yes	No	No				
Netherlands	0.2	OI	3			Yes			Yes	Yes
	0.1	OI	3			Yes			Yes	Yes
Norway	0.5	3V	6	Yes	No	No	Yes		No	
	0.1	-	-			-				
	0.05	-	-			-				
Spain	0.5	OI	6	Yes	Fg	Yes				Yes
	0.2	OI	6			Yes				
Sweden	0.4	3V	6			No				Yes
	0.2	3V	6			No				Yes

Table 7.2: Data assimilation properties of the operational Hirlam systems. Shows resolution in degrees, Data Assimilation system, assimilation cycle (between subsequent analyses), indication whether re-analysis with late observations done or not, indication of re-initialisation by using ECMWF analysis or forecast full fields or not (once or twice per day), use of SATOB, SATEM, TOVS or scatterometer data and use of SST analysis.

case. Both track and intensity prediction were very sensitive to the choice of initial fields and cyclone bogusing. Horizontal resolution enhanced from 50 to 10 km had a large impact on the intensity prediction, probably due to a better representation of the inner core of the cyclone with higher resolution, but not in the track prediction. The reference Hirlam model using the STRACO scheme achieve very good track prediction but quite poor intensity prediction. The results improved significantly using Kain-Fritsch convection and Rasch-Kristjansson condensation in Hirlam (Nagata et al., 2001). With this updated moist physics Hirlam was among the models that could reproduce qualitatively the time evolution of the typhoon central pressure.

HIRLAM model has been used in several other case studies of extreme weather situations. de Bruijn and Brandsma (2000) studied the flodding event in Ireland connected with the Hurricane Charley. The December 1999 storms in Europe were also reanalysed and simulated.

7.4 Application of the HIRLAM-model in regional climate modeling

7.4.1 Introduction

The topics of global change and climate change have gained increasing attention during the recent years, cf. (Houghton et al., 1996, 2001). The basic tools for projecting the anthropogenic emissions of greenhouse gases and sulfur, as well as other anthropogenic effects on the climate system, such as land-use changes, are global climate models. Modern global climate models are 3-dimensional coupled ocean-atmosphere general circulation models (OAGCMs). However, as climate simulations need to be run over extended periods of time (of the order of 100-200 years)

and the complexity and dimensions of the climate system are very large indeed, global climate simulations are seldom run on resolutions allowing for much regional-to-local scale detail. This means that such details as more variable topography, complicated coastlines and regional oceans or lake systems are not represented well in global climate models and simulations. This lack is addressed with regionalization techniques, one of which is the use of regional climate models, also known as dynamical downscaling of OAGCM-simulations, e.g. Giorgi and Mearns (1991, 1999); McGregor (1997).

Dynamical downscaling is basically analogous to operational weather prediction with limited area models, such as HIRLAM. A global simulation is used to provide lateral and, as required, surface forcing conditions that are fed to the regional domain, in which the time evolution of weather, or climate, is simulated on a higher resolution than in the global model. The regional simulation might also differ from the global one in that different parameterizations are applied, such as those for convection/turbulence.

In Sweden, regional climate modeling is pursued at SMHI. In Norway, regional climate modeling is done by the RegClim-program at DNMI, and in Denmark at the Danish Climate Centre, DCC, at DMI. In each case, a regional climate model derived from HIRLAM is used, but with reworked or replaced parameterizations. At DNMI and DMI, the HIRHAM is used (HIRLAM model with "Hamburg physics", i.e. those of the AGCM called ECHAM). At SMHI, the model called RCA is used, developed from HIRLAM by means of gradually-introduced changes in the parameterizations, guided by simulations and analyses. Some discussion of the regional climate modeling at SMHI, DNMI and DMI can be found in Christensen et al. (2001). Regional climate modeling at SMHI, in SWECLIM, is discussed below, as an example of a climate application of HIRLAM.

7.4.2 The RCA regional climate model

The regional climate model at SMHI is called the Rossby Centre regional climate model. Its atmospheric component including the treatment of soil, snow and land surface is called RCA. The Rossby Centre is a research group at SMHI, established in the Swedish regional climate modeling program SWECLIM. RCA builds on the HIRLAM versions 2.5 and 2.7 (Källén, 1996; Eerola et al., 1997). The present coding structure of the RCA concurs with the HIRLAM reference system version 4.7.4.

The first version of the RCA - the RCA0 (Rummukainen et al., 1998) - featured few changes to the parent HIRLAM. So, the schemes for e.g. radiation (Savijärvi, 1990; Sass et al., 1994), convection (Kuo, 1965, 1974), large-scale cloud and precipitation microphysics (Sundqvist et al., 1989; Sundqvist, 1993) and vertical diffusion (Louis, 1979; Geleyn, 1987) were retained. Small changes were introduced in the surface/soil treatment. The seasonal ice cover on the Baltic Sea and inland lakes became based on "ice proxies" relating sea/lake ice to a deep soil temperature. It was soon assessed, however, that some aspects of the model were not ideal for the climate application. This was consequently addressed in the further development of RCA, especially moist physics, radiation, land surface parameterization including hydrological aspects and coupling of lake models and Baltic Sea models with the RCA.

A new version of the RCA model (RCA1) came in 1999 (Rummukainen et al., 2001), with updates to the treatment of the water cycle at the surface and in soil. The relaxation to prescribed deep-soil moisture was removed. In addition to the vertical diffusive transport of soil moisture as in HIRLAM, hydrological runoff generation down the soil column was included based on the hydrological HBV model (Lindström et al., 1997). I.e. excess water was no longer discarded as surface runoff when field capacity was reached. In RCA1 runoff actually started before the grid box on the average reached field capacity, as subgridscale soil moisture variability

was introduced. Evapotranspiration was drawn from both the two soil moisture layers (Noilhan and Planton, 1989). Soil freezing (Viterbo et al., 1999) was included. Furthermore, RCA1 was extended with a process-oriented, interactive Baltic Sea module. This was the 13-basin PROBE-Baltic of Omstedt and Nyberg (1996). Inland lakes were included as well, after (Ljungemyr et al., 1996; Omstedt, 1999). A modified radiation scheme allowing for varying CO2 was added into RCA1 later in 2000 (Räisänen et al., 2000).

The next version of the RCA model system (RCA2) was already being worked on while RCA1 was used for climate simulations (see next Section). Major changes were now introduced in the model's moist physics, land surface treatment as well as the treatment of the Baltic Sea. A simple lateral river routing was also included in RCA2, to transfer precipitation on land into river output of fresh water to the sea.

The moist physics in RCA2 are an interlinked package consisting of the convection scheme from Kain and Fritsch (1993) (KF), the condensation/cloud scheme from Rasch and Kristjánsson (1998) (RK) and the CBR turbulence (Cuxart, 1997; Cuxart et al., 2000). Compared to the schemes adopted earlier from the HIRLAM 2.5-2.7, these are more specific for simulations at 10-50 km horizontal resolution. The application of these in RCA2 has provided for much cooperation with the HIRLAM project. The CBR-scheme has a prognostic equation for the turbulent kinetic energy (TKE) and an analytical mixing length to close the system (Bougeault and Lacarrère, 1989). In RCA2, the scheme uses dry variables. The KF scheme has removal of grid-column CAPE as closure for deep convection. The shallow convection closure is based on TKE. In RK, conversion between water vapor and condensate follows Sundqvist et al. (1989). In the conversion from condensate to precipitation, a more explicit determination of the physical processes is made. Cloud liquid water is prognostic. Cloud ice is still a diagnostic quantity. Cloud cover is diagnostic. Two methods to diagnose the cloud fraction have been studied, one based on relative humidity (Slingo, 1987) and another based on both cloud water and relative humidity (Xu and Randall, 1996).

The RCA2 land surface treatment is further updated, cf. Bringfelt et al. (2001) with additions like rainfall interception on vegetation, surface (stomatal) resistance to transpiration with seasonal effects and a snow melt parameterization with subgridscale effects.

As an alternative to the process-based Baltic Sea model implemented in RCA1, RCA2 comes with a coupling to the high-resolution 3-D Rossby Centre Ocean model (RCO, Meier et al. (1999)). For flexibility, the RCA and RCO are retained as separate modules, connected by a coupler-program OASIS (Valcke et al., 2000). The coupler performs interpolation and extrapolation between grids and can do subgridscale redistribution of fluxes. At present, the coupled RCA-RCO system (called the RCAO) is flux-coupled and does not require flux adjustments even in multi-year simulations. Most of the atmosphere-ocean fluxes that represent the physical processes of energy and water exchange are calculated in the atmosphere, based on the sea surface temperature and sea ice. One exception is the sea surface emission of longwave radiation that is in RCAO provided by the ocean model. As the radiation emission depends on the state of the emitting surface, a more detailed calculation can be made in the RCO-model that typically has a higher horizontal resolution (2-6 nautical miles) and thus more detail than the RCA model (a few tens of kilometers). A schematic of the coupling cycle is shown in Figure 7.1.

The domain for RCA is set with a spherical, rotated latitude/longitude Arakawa C grid, the hybrid vertical coordinate and a model top at 10 hPa, as in HIRLAM. RCA0 and RCA1 feature 19 vertical levels whereas in RCA2 a denser vertical discretization is in use. The 8-point lateral boundary relaxation follows Davies (1976) as in HIRLAM. RCA0. Most of the simulations are so far done with a rather typical regional climate model resolution of 40-50 km. Other horizontal resolutions so far in use are 88 km, and 20 km. The typical model domain of RCA0 and RCA1

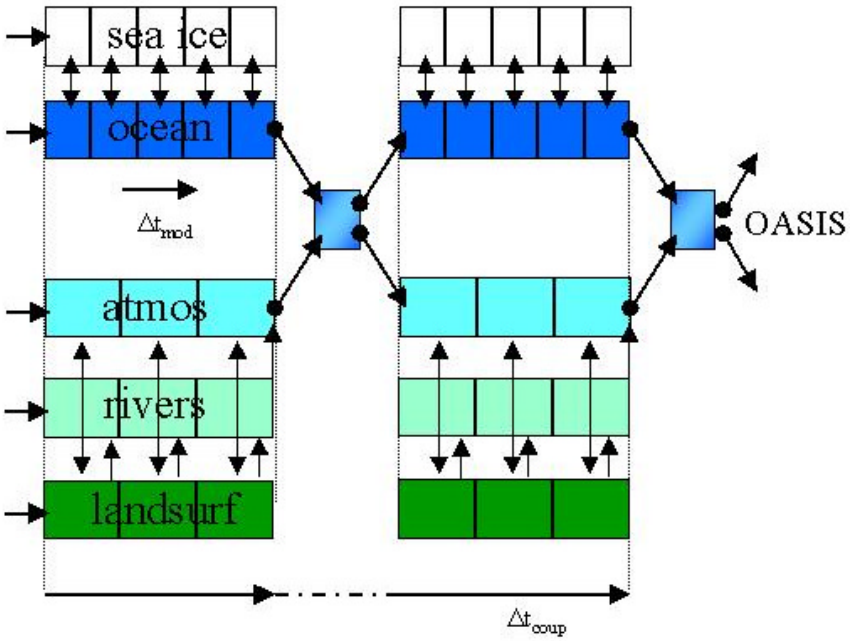


Figure 7.1: The coupled SWECLIM regional (climate) modeling system in 2001-2002. The atmospheric and the land surface/soil components are based on a version of the HIRLAM model, incremented also by models for inland lakes and lake systems (not shown) and lateral river runoff. The regional ocean model RCO hosts also a sea ice model. The coupling between the atmosphere-lake-land surface-rivers and ocean-sea ice is implemented with the OASIS coupler program. The coupling time step is three hours.

is shown in Figure 7.2. These versions are run with the Eulerian transport scheme.

In RCA2, the HIRLAM semi-Lagrangian transport is implemented, resulting in a considerable speed-up of the simulations. RCA2 can be integrated with a timestep up to 4-5 times longer than RCA1 where the recurring grid-point storms posed a more serious limit on the time step length. RCA2, run at 40-50 km resolution, allows the use of a 30-minute time step. Even longer time steps are computationally feasible, but they start affecting the simulated results.

7.4.3 RCA-simulations

RCA is used to downscale coupled global climate simulations, global analyses and reanalyses from ECMWF and reanalyses from NCEP.

Downscaling of global climate simulations aims at generating regional climate scenarios with added detail. Typically, the regional model is run over two so-called time slices (or windows) from a longer global simulation. The first window should represent a control or a present-day climate and the second window some future climate. The differences between the time slices represent a measure of climate change as conditioned by the global model experiment.

The RCA0 and RCA1 were forced with six-hourly global model data corresponding to 10-year time slices. Data from both the UKMO HadCM2 (Johns et al., 1997; Mitchell and Johns, 1997) and from the ECHAM4/OPYC3 of the Max Planck Institute for Meteorology in Hamburg (Roeckner et al., 1999) were used. These two global models had appr. horizontal resolution of 300 km. Their regionalization with RCA to 44 km implied therefore a considerable increase

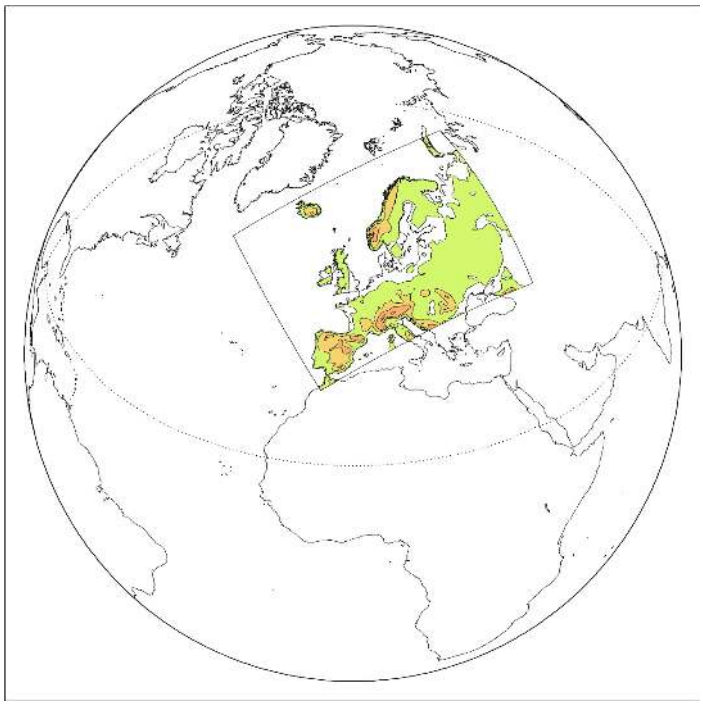


Figure 7.2: Typical RCA simulation domain applied during 1998-2000.

in resolution. In total, RCA0 and RCA1 were run for about 120 years of such simulations ((Rummukainen et al., 1998; Räisänen et al., 1999; Rummukainen et al., 2001; Räisänen et al., 2001; Bergström et al., 2001; Graham et al., 2001; Räisänen and Joelsson, 2001)). Overview of some of the simulated climate changes in screen temperature, precipitation and 10-meter wind for are shown in Figure 7.3. For the Nordic land area, the scenario changes are robust across different versions of the RCA, different horizontal resolutions and even the choice between the two choices of global scenarios. Above and around the Baltic Sea, however, differences between RCA0 and RCA1 are obvious. It seems to be the simplistic treatment of the ice climate that led in RCA0 to locally unrealistic climate change responses giving rise to large temperature and wind increases over the Baltic Sea. Also precipitation increases were the largest in RCA0. The very short time steps that had to be used in RCA1-22H, due to its high horizontal resolution and the tendency of grid-point storms, exposed an unphysical dependency of snowmelt on the time step length. Finally, in the 44 km simulations of RCA1-44H and RCA1-44E, the variable CO2 content by that time introduced in the radiation scheme explains a few tenths of the higher temperature change compared to the earlier 88 km runs that were otherwise done almost with RCA1.

The regional climate simulations of SWECLIM have received considerable attention from a range of users. They have been intended as concretization examples on how the global-scale climate change might project to national and sub-national scales that the general public is most familiar with. Regionalization can also help in better understanding possible needs of local-to-regional adaptation and mitigation measures. Per se regionalization can not be assumed either to increase or decrease the quality or certainty of global climate scenarios. Regionalization does lead to a better representation of the physical system and feedback on scales not represented in global models, and the information can be made more suited to impact modeling.

In addition to the 120 years of regional scenario simulations, about a comparable amount of testing was conducted with RCA0 and RCA1. Additional runs were done using ECMWF ERA-

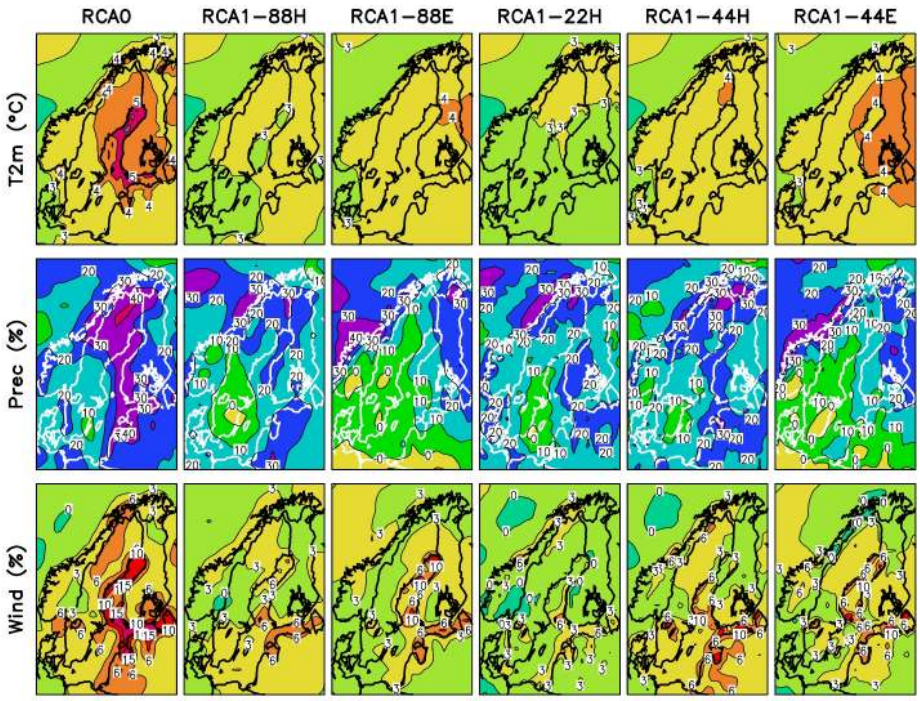


Figure 7.3: Changes in annual mean temperature (C), precipitation (%) and wind speed (%) in six regional climate scenarios with the RCA in 1998-2000. RCA0 and RCA1 have been used. RCA0 was run on a 44 km regional resolution. The resolution in the RCA1-simulations is indicated at the figure top (always one of 88 km, 22 km or 44 km). 'H' = OAGCM-forcing from HadCM2, 'E' = ECHAM4/OPYC3.

15. The latter, also known as "perfect boundary conditions", cf. Sass and Christensen (1995), were performed in earnest during 2000-2001 in the development and testing of the next model version, the RCA2. First, a series of simulations were performed over a 5-year period from 1988 to 1993, to test, evaluate and develop the coupled RCAO system. The RCA2, without a Baltic Sea component was later run over a period of 17 years from 1984 to 2001 with a combination of ERA-15 and operational analyses as boundary and initial conditions. These perfect boundary simulations with the RCAO and RCA2 systems have been compared to existing observations and analyses of these, e.g. for screen temperature, precipitation, radiation, cloudiness, Baltic Sea temperatures, ice cover and salinity variations.

The typical set-up of the RCA2 in the long perfect boundary simulations featured 44 km resolution, 24 levels, semi-Lagrangian dynamics and time step of the order of 30 minutes. Results from this integration will be used to document the quality of RCA2 in simulating the present observed climate variability over Europe. A few key results are presented here (Figure 7.4 – 7.5).

Figure 7.4 shows the climatological mean annual cycle of precipitation, as a mean for the Nordic land region, for the period 1985-1999. Two observational data sets are included for comparison. The blue curve shows the RCA2-modeled result. A clear maximum is seen in the precipitation in the summer to fall period. The green curve shows the annual cycle according to CRU-data. These data are uncorrected for e.g. undercatch. RCA2 overestimates precipitation, relative to the CRU data, by as much as 20% in fall and in winter. However, it is well known that uncorrected precipitation data sets underestimate, in particular, true wintertime precipitation.

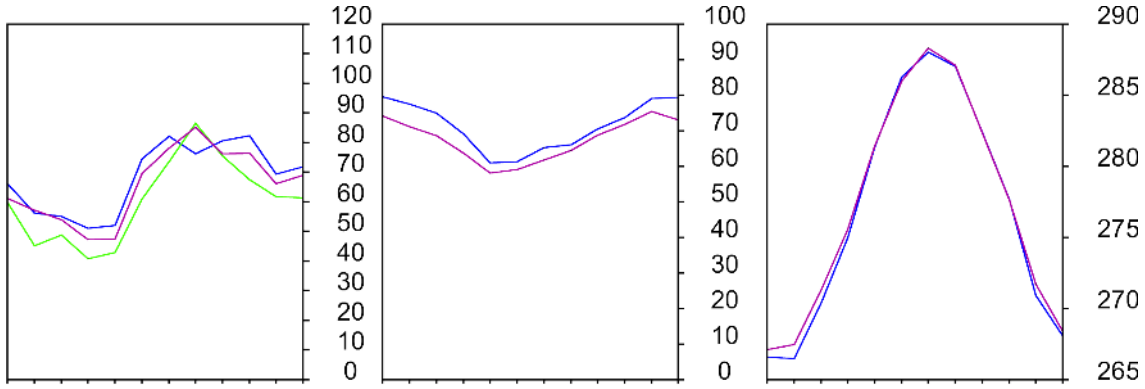


Figure 7.4: The mean annual cycle of (left) precipitation in mm/month, (middle) total cloud cover in % and (right) screen temperature in K for the Nordic land area, in 1985-1999. Comparisons are made to various observations-based data (see text). The blue curves are for the RCA2-modeled results, the model being forced by ECMWF analyses.

The red curve is for a corrected data set, the GPCP (Global Precipitation Climatology Project) version 2 ((Huffman and Bolvin, 2001)). These data are global and combine both satellite and station estimates. Local corrections have been applied to account for precipitation undercatch and wind effects ((Huffman and Bolvin, 2001; Rubel and Hantel, 2001)). It is concluded that at these spatial scales, RCA2 has an extremely realistic simulation of precipitation. Furthermore, care must be taken in the use of the various available observations-based data when evaluating model performance.

Figure 7.4 shows also the mean annual cycle of total cloudiness for the Nordic land region, the observations being from the CRU data. The annual cycle is well simulated by RCA2 with a minimum in the summer period. There is a slight overestimate of total cloud cover, relative to the CRU data. Some of this can be explained by the known negative bias in station-based observations during periods of darkness. Hahn et al. (1995) indicate climatological wintertime observations of cloudiness can be up to 2-3% underestimated. Nevertheless, there appears to be a positive bias in the RCA2 cloud field. A part of this problem may lie in the fact the RCA2 cloudiness we plot here is derived from a maximum-random overlap of clouds in the vertical (the technique believed to most closely mimic reality). The cloud field used in the radiation code uses a maximum overlap technique. As a result the radiatively active cloud field in the model is always less than the diagnostic cloud field. An effect of this is that, with respect to longwave radiation, a positive bias in the total cloud field is required to prevent a negative bias in the surface radiation budget. Also this effect can be seen in Figure 7.4, which shows the mean annual cycle of 2-meter temperature derived from the CRU observations and from the RCA2 simulation. Note in particular the small cold bias in the winter period. The main control on the Nordic wintertime near surface temperature is by longwave radiation. Periods of cloudiness, in the winter, are associated with increased atmospheric longwave emission to the surface and thus warmer temperatures. There is some indication that due to deficiencies in the representation of longwave radiation in the model, a positive bias in the cloud field is required to simulate accurately the surface longwave radiation budget and prevent a large wintertime near surface cold bias developing. Such an aspect is of course unsatisfactory. It will be studied in the further development of the RCA-model.

Figure 7.5 illustrates the monthly fields, starting from 1985, upon which the climatological mean annual cycles were based. After 1992 the CRU cloud data becomes unusable due to the lack of actual observations. Cloud observations are therefore only plotted out to the end of

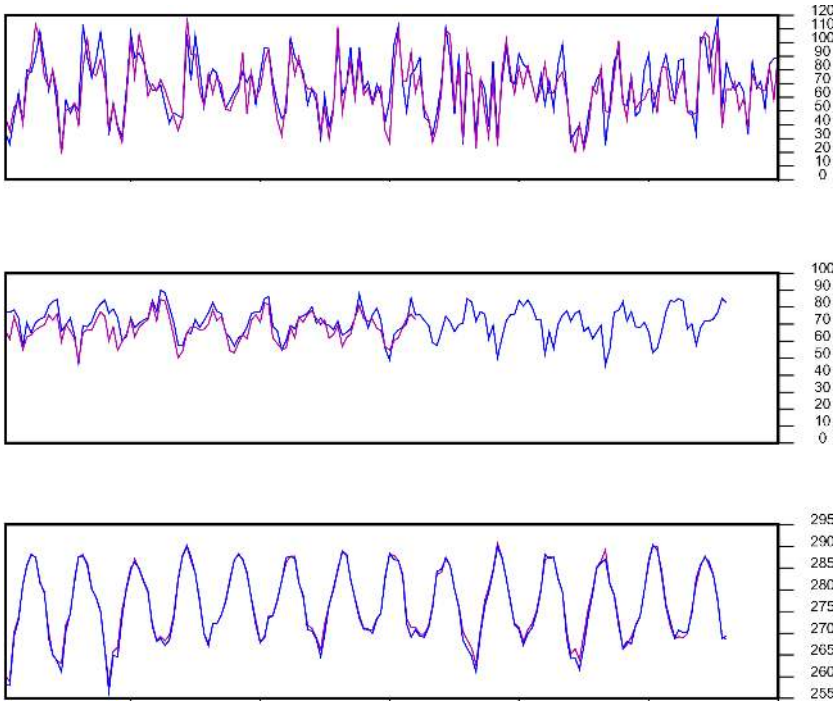


Figure 7.5: The monthly course of (top) precipitation in mm/month, (middle) total cloud cover in % and (bottom) screen temperature in K for the Nordic land area, in 1985-1999. Comparisons are made to various observations-based data (see text). The blue curves are for the RCA2-modeled results, the model being forced by ECMWF analyses.

1992. Similarly the temperature data set is only presently available to the end of 1998. The precipitation observations are available for the entire period. The intra-seasonal variability is well represented in all 3 parameters. The strong seasonal variability in the precipitation field is very well simulated. With respect to the cloud field it is worth noting that the positive bias in the modeled cloud field occurs mainly in the first three winters, these being relatively cold. This may indicate a problem in representing clouds in cold conditions, although it is felt more likely that the problem lies in the longwave clear sky portion of the radiation code. An over-transparent clear sky to longwave radiation will lead to a cold bias developing in the atmosphere and an associated positive relative humidity bias. As a direct result too many clouds will form and prevent the excessive (longwave) cooling and thus control the surface and atmospheric (cold) temperature errors through a positive bias in the cloud field. This issue illustrates well the tight coupling between atmospheric and surface radiation and temperature fields and the model cloud field.

The RCA has also been used for simulations for other regions of the world, e.g. continental U.S., eastern Pacific Ocean, Africa and the Arctic region (see Figure 7.6). Such simulations have been undertaken as participation in intercomparisons of regional models and in research projects targeting improved physical parameterizations in operational and climate models, both regional and global alike.

As an example, the international "Project to Intercompare Regional Climate Simulations" (PIRCS) is briefly mentioned. PIRCS is organized under the auspices of the International Institute of Theoretical and Applied Physics at Iowa State University. PIRCS 1b-experiment runs were conducted for a 2 month long simulation of the summer of 1993 over the continental United States. This was a period with extreme flood events in the Upper Mississippi basin.

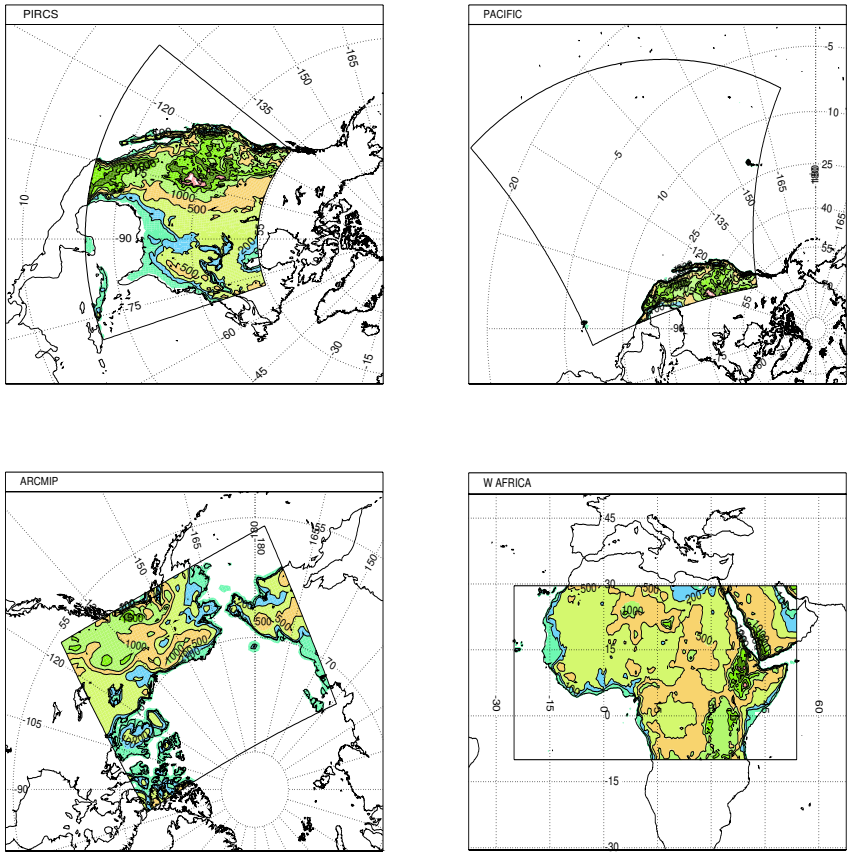


Figure 7.6: Different regional domains that the RCA2-model has been used for. Clockwise from top left: Continental U.S. (PIRCS-study), Eastern Pacific (EUROCS), Central Africa (EUROCS) and part of the Arctic region (ARC-MIP).

The participating models were run at 50 km resolution. RCA was forced with both ECMWF and NCEP reanalyses. The RCA very successfully reproduced the overall flood event and the intra-seasonal and diurnal variability of rainfall. The latter is illustrated in Figure 7.7 .

Akin to the other models, RCA produced too much frontal rainfall. Of the observed rainfall, a large proportion fell from isolated, intense Mesoscale Convective Complexes, likely requiring an even higher resolution to be properly captured in models.

7.4.4 Future outlook

The development of RCA continues, in cooperation with the international HIRLAM project as well as other R&D-constellations. Some of the foci are the description of exchanges between the atmosphere and the land and ocean surface, incorporation of hydrology, and further progress towards a more comprehensive, "earth system" -type of a model system. Work on computational aspects, as well as the coupled issue of moist physics and radiation is also targeted.

During the next few years, new regional simulations will be performed. New regional climate change scenarios will be constructed from several 30-year long time slices, based on the global models at the Hadley Centre and the Max-Planck-Institute for Meteorology in Hamburg and the recent IPCC SRES emission scenarios. Compared to the earlier regional climate simulations with the RCA-model, the 30-year simulations will provide better sampling of simulated natural

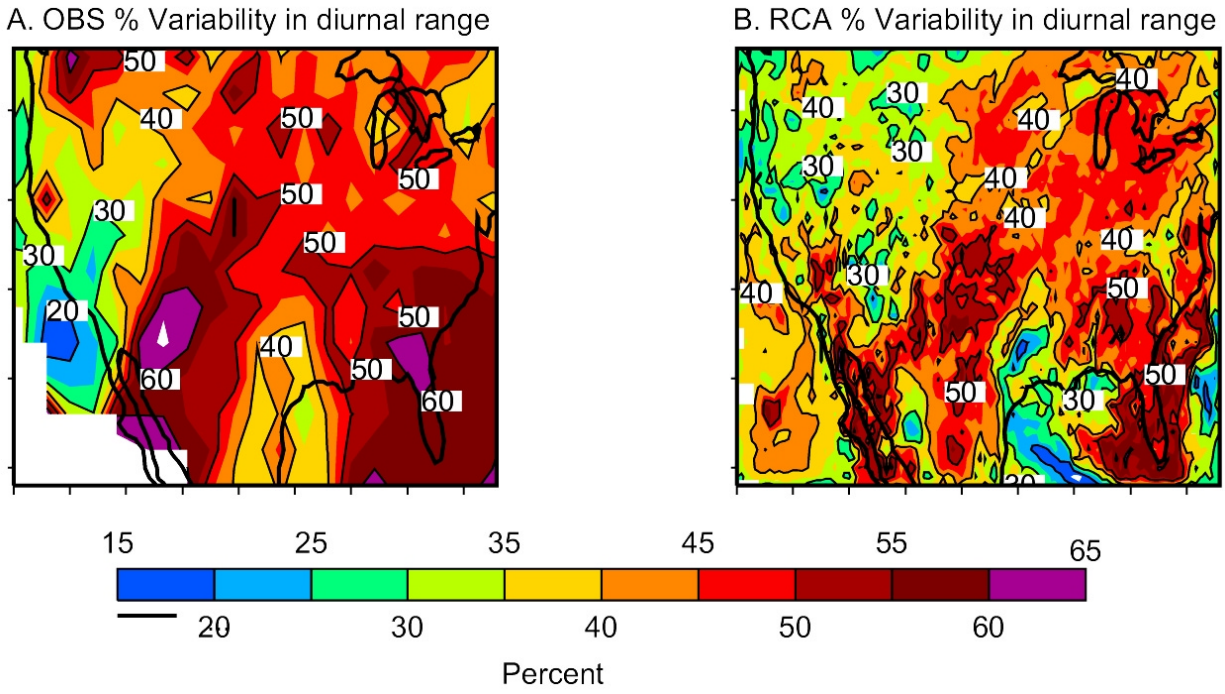


Figure 7.7: The June-July 1993 percentage of precipitation variability with higher frequency than diurnal. The observed data are from (Higgins et al., 1996). The RCA was run at 50 km resolution forced by ECMWF analyses.

variability and extreme statistics and so more detailed scenarios for possible changes in these due to climate change. Long simulations with different RCAO-model system configurations will also be pursued, utilizing the reanalyses at ECMWF. Whenever possible, the RCAO-system will be tested and applied for different parts of the world, both to do climate system research and to pursue testing of the model system and its parameterizations.

Increased cooperation between the Nordic climate applications of HIRLAM will also be explored. One potential goal is a "unified model system" that would suit for both the operational and the climate applications alike.

Bibliography

Alexander, R. C., and R. L. Mobley, 1974: Monthly average sea-surface temperatures and ice-pack limits for 1° global grid. Technical report, RAND Rep. R-1310-ARPA, 1–30.

Andersson, E., and H. Järvinen, 1999: Variational Quality Control. *Quart. J. Roy. Met. Soc.*, **125**, 697–722.

Avissar, R., and R. A. Pielke, 1989: A parameterization of heterogeneous land surfaces for atmospheric numerical models and its impact on regional meteorology. *Mon. Wea. Rev.*, **117**, 2113–2136.

Ayuso, J. J., 1995: Implementation of ASSISBA in the HIRLAM system. *Proceedings of Hirlam 3 Workshop on soil processes and soil/surface data assimilation*, 46-50. Available at SMHI, S-60176 Norrköping, Sweden.

Beljaars, A. C. M., and P. Viterbo, 1998: The role of the boundary layer in a numerical weather prediction model. *Clear and Cloudy Boundary Layers*, A. A. M. Holtslag and P. G. Duynkerke, Eds. North Holland Publishers, 287–304.

Bergström, S., B. Carlsson, M. Gardelin, G. Lindström, A. Pettersson, and M. Rummukainen, 2001: Climate change impacts on runoff in Sweden - assessments by global climate models, dynamical downscaling and hydrological modelling. *Clim. Res.*, **16**, 101–112.

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

Bjørge, D., and R. Skålin, 1995: Parlam, the parallel HIRLAM version at DNMI. Research Report. Technical Report 27, Norwegian Meteorological Institute, 46.

Blyth, E. M., A. J. Dolman, and N. Wood, 1993: Effective resistance to sensible and latent heat flux in heterogeneous terrain. *Quart. J. Roy. Met. Soc.*, **119**, 423–442.

Boer, G. J., 1983: Homogeneous and isotropic turbulence on the sphere. *J. Atm. Sci.*, **40**, 154–163.

Bougeault, P., and P. Lacarrère, 1989: Parameterisation of orography induced turbulence in a meso-beta scale model. *Mon. Wea. Rev.*, **117**, 1872–1890.

Bouttier, F., J.-F. Mahfouf, and J. Noilhan, 1993: Sequential Assimilation of Soil Moisture from Atmospheric Low-Level Parameters. Part I: Sensitivity and Calibration Studies. *J. Appl. Meteor.*, **32**, 1335–1351.

Brancovic, C., and J. van Maanen, 1985: The ECMWF Climate System. Technical Report 109, ECMWF Research Department Tech. Memo, 51 + figures.

Braud, I., J. Noilhan, P. Bessemoulin, P. Mascart, R. Haverkamp, and M. Vauclin, 1993: Bare-ground surface heat and water exchanges under dry conditions: observations and parameterization. *Bound. Lay. Met.*, **66**, 173–200.

Bretherton, F. P., 1969: Momentum transport by gravity waves. *Quart. J. Roy. Met. Soc.*, **95**, 213–243.

Brier, G. W., and R. A. Allen, 1951: Verification of Weather Forecasts. *Compendium of Meteorology*, T. F. Malone, Ed. American Meteorological Society, 841–848.

Bringfelt, B., 1996: Test of a new land-surface treatment in HIRLAM. Technical Report 23, HIRLAM. Available at SMHI, S-60176 Norrköping, Sweden.

Bringfelt, B., N. Gustafsson, P. Vilmusenaho, and S. Järvenoja, 1995: Updating of the HIRLAM physiography and climate data base. Technical Report 19, HIRLAM, 42. Available at SMHI, S-60176 Norrköping, Sweden.

Bringfelt, B., M. Heikinheimo, N. Gustafsson, V. Perov, and A. Lindroth, 1999: A new land surface treatment for HIRLAM - Comparisons with NOPEX measurements. *J. of Agr. and Forest Met.*, **98–99**, 239–256.

Bringfelt, B., J. Räisänen, S. Gollvik, G. Lindström, , L. P. Graham, and A. Ullerstig, 2001: The land surface treatment for the Rossby Centre regional atmospheric climate model - version 2 (RCA2). Technical Report 98, SMHI Reports Meteorology and Climatology (RMK), 40. Available at SMHI, S-60176 Norrköping, Sweden.

Brutsaert, W., 1975: The roughness length for water vapor, sensible heat and other scalars. *J. Atm. Sci.*, **32**, 2028–2031.

Cameron, D. R., and S. F. Milton, 2001: Diagnosing Error Sources in Extratropical Cyclone Lifecycles. Numerical Weather Prediction Technical Report 356, Met Office.

Charnock, H., 1955: Wind stress on a water surface. *Quart. J. Roy. Met. Soc.*, **81**, 639–640.

Chen, C., and W. R. Cotton, 1987: The physics of the marine stratocumulus-capped mixed layer. *J. Atm. Sci.*, **44**, 2951–2977.

Chou, M.-D., 1986: Atmospheric solar heating in the water vapour bands. *J. Appl. Meteor.*, **25**, 1532–1542.

Christensen, J.-H., J. Räisänen, T. Iversen, D. Bjørge, O. B. Christensen, and M. Rummukainen, 2001: A synthesis of regional climate change simulations. A Scandinavian perspective. *Geophys. Res. Lett.*, **28**, 1003–1006.

Clark, T. L., and W. R. Peltier, 1984: Critical reflection and the resonant growth of nonlinear mountain waves. *J. Atm. Sci.*, **41**, 3122–3134.

Claussen, M., 1991: Estimation of areally-averaged surface fluxes. *Bound. Lay. Met.*, **54**, 387–410.

Claussen, M., 1995: Flux aggregation at large scales. On the limits of validity of the concept of blending height. *J. Hydrology*, **166**, 371–382.

Clayson, C. A., C. W. Fairall, and J. A. Curry, 1996: Evaluation of turbulent fluxes at the ocean surface using surface renewal theory. *J. Geoph. Res.*, **101**(C12), 28503–28513.

Cordeneanu, E., and J.-F. Geleyn, 1998: Application to local circulations above the Carpathian-Black sea area of a NEP-type meso-scale model. *Contr. Atm. Phys.*, **71**, 191–212.

Courtier, P., E. Andersson, W. Heckley, J. Pailleux, V. D, M. Hamrud, A. Hollingsworth, F. Rabier, and M. Fisher, 1998: The ECMWF implemenentation of the three dimensional variational assimilation (3D-Var). Part I. Formulation. *Quart. J. Roy. Met. Soc.*, **124**, 1783–1808.

Courtier, P., J.-N. Thépaut, and A. Hollingsworth, 1994: A strategy for operational implementation of 4D-Var using an incremental approach. *Quart. J. Roy. Met. Soc.*, **120**, 1367–1388.

Cressman, G., 1959: An operational objective analysis system. *Mon. Wea. Rev.*, **88**, 327–342.

Crutcher, H. L., and J. M. Meserve, 1970: *Selected level heighs, temperatures and dew points for the Northern Hemisphere.* Government Printing Office, Washington D.C., 132 pp.

Cuxart, J., 1997: *PBL simulations: from LES to GCM models. PhD thesis.* (Univ. Barcelona), 205 pp + reprints pp.

Cuxart, J., P. Bougeault, and J. L. Redelsberger, 2000: A turbulence scheme allowing for mesoscale and large-eddy simulations. *Quart. J. Roy. Met. Soc.*, **126**, 1–30.

Daley, R., 1985: The analysis of synoptic scale divergence by a statistical interpolation scheme. *Mon. Wea. Rev.*, **113**, 1066–1079.

Daley, R., 1991: *Atmospheric data analysis.* Cambridge University Press.

Davies, H. C., 1976: A lateral boundary formulation for multi-level, prediction models. *Quart. J. Roy. Met. Soc.*, **102**, 405–418.

de Bruijn, E. I. F., and T. Brandsma, 2000: Rainfall prediction for a flodding event in Ireland caused by the remnants of Hurricane Charley. *J. Hydrol.*, **239**, 148–161.

DeCosmo, J., K. B. Katsaros, S. D. Smith, R. J. Anderson, W. A. Oost, K. Bumke, and H. Chadwick, 1996: Air-sea exchange of water vapour and sensible heat: The Humidity Exchange Over the Sea (HEXOS) results. *J. Geoph. Res.*, **101**(C5), 12001–12016.

Delage, Y., 1997: Parameterising sub-grid scale vertical transport in atmosperic models under statically stable conditions. *Bound. Lay. Met.*, **82**, 23–48.

Douville, H., J. F. Royer, and J.-F. Mahfouf, 1995: A new snow parameterization for the Météo-France climate model. Part I: Validation in stand-alone experiments. *Clim. Dyn.*, **12**, 21–35.

Duynkerke, P. G., and A. G. M. Driedonks, 1987: A model for the turbulent structure of the stratocumulus-toppped atmospheric boundary layer. *J. Atm. Sci.*, **44**, 43–46.

Dyer, A. J., 1974: A review of flux-profile relationships. *Bound. Lay. Met.*, **7**, 363–372.

Ebert, E. E., and J. L. McBride, 2000: Verification of Precipitation in Weather Systems: Determination of Systematic Errors. *Journal of Hydrology*, **239**, 179–202.

Eerola, K., 1995: Experiences with the analysis of sea surface temperature, ice coverage and snow depth. *Proceedings of Hirlam 3 Workshop on soil processes and soil/surface data assimilation*, 33-37. Available at SMHI, S-60176 Norrköping, Sweden.

Eerola, K., D. Salmond, N. Gustafsson, J. A. Garcia-Moya, P. Lönnberg, and S. Järvenoja, 1997: A parallel version of the HIRLAM forecast model: Strategy and results. In: Making Its Mark, Proceedings of the Seventh ECMWF Workshop on the Use of Parallel Processors in Meteorology (Hoffmann and Kreitz, ed.). Technical report, World Scientific, Singapore, 134–143.

Engelen, R. A. V., I. Heitlager, L. Wolters, and G. Cats, 1997: Incorporating application dependent information in an automatic code generation environment. *Proceedings of the 11th ACM International Conference on Supercomputing, New York, ACM Press*, 180-187.

FAO-UNESCO, 1987: Soils of the World. Food and Agriculture Organization and United Nations Educational, Scientific and Cultural Organization. Technical report, Elsevier Science Publishing Co. Inc., New York.

Fritsch, J. M., and C. F. Chappell, 1980: Numerical Prediction of Convectively Driven Mesoscale Pressure Systems. Part I: Convective parameterization. *J. Atm. Sci.*, **37**, 1722–1733.

Garrat, J. R., 1977: Review of drag coefficients over oceans and continents. *Mon. Wea. Rev.*, **105**, 915–929.

Garrat, J. R., 1992: *The Atmospheric Boundary Layer*. Cambridge University Press, 316 pp.

Geleyn, J.-F., 1987: Use of a modified Richardson number for parameterizing the effect of shallow convection. *J. Meteorol. Soc. Japan*, **65**, 141–149.

Geleyn, J.-F., 1988: Interpolation of wind, temperature and humidity values from the model levels to the height of measurement. *Tellus*, **40A**, 347–351.

Geleyn, J.-F., 2000: Arpege technical documentation. Technical report, Available at Meteo France, Toulouse. (http://www.cnrm.meteo.fr/aladin/MODELES/EXT/Physics/).

Geleyn, J.-F., and H. J. Preuss, 1983: A new data set of, satelite-derived surface albedo values for operational use at ECMWF. *Arch. Meteor. Geophys. Bioclim.*, **A 32**, 353–359.

Georgelin, M., P. Bougeault, T. Black, N. Brzovic, A. Buzzi, J. Calvo, V. Cassé, M. Desgagné, R. El-Khatib, J.-F. Geleyn, T. Holt, S.-Y. Hong, T. Kato, J. Katzfey, K. Kurihara, B. Lacroix, and et.al, 2000: The Second COMPARE exercise: A model intercomparison using a case of a typical mesoscale orographic flow, the PYREX IOP3. *Quart. J. Roy. Met. Soc.*, **126**, 991–1029.

Giard, D., and E. Bazile, 2000: Implementation of a new assimilation scheme for soil and surface variables in a global NWP model. *Mon. Wea. Rev.*, **128**, 997–1015.

Giorgi, F., and L. Mearns, 1991: Approaches to the simulation of regional climate change: A review. *Rev. Geophys.*, **29**, 191–216.

Giorgi, F., and L. Mearns, 1999: Introduction to special section: Regional climate modeling revisited. *J. Geoph. Res.*, **104**, 6335–6352.

Grachev, A. A., C. W. Fairall, and S. E. Larsen, 1998: On the determination of the neutral drag coefficient in the convective boundary layer. *Bound. Lay. Met.*, **86**, 257–278.

Graham, L. P., M. Rummukainen, M. Gardelin, and S. Bergström, 2001: Modelling Climate Change Impacts on Water Resources in the Swedish Regional Climate Modelling Programme. *Detecting and Modelling Regional Climate Change and Associated Impacts.*, M. Brunet and D. Lopez, Eds. Springer-Verlag, Berlin/Heidelberg/New York,, 567–580.

Gustafsson, N., 1985: Development of mesoscale analysis scheme for nowcasting and very short range forecasting. *Proceedings of Workshop on High Resolution Analysis*, 24-26.

Gustafsson, N., L. Berre, S. Hörnquist, X.-Y. Huang, M. Lindskog, B. Navascués, K. S. Mogensen, and S. Thorsteinsson, 2001: Three-dimensional variational data assimilation for a limited area model. Part I: General formulation and the background error constraint. *Tellus*, **53A**, 425–446.

Hahn, D. J., S. G. Warren, and J. London, 1995: The effect of moonlight on observation of cloud cover at might, and application to cloud climatology. *JCl*, **8**, 1429–1446.

Hansen, C., and S. E. Larsen, 1997: Further work on the Kitaigorodskii roughness length model: A new derivation using Lettau's expression on steep waves. *Geophysica*, **33**(2), 29–44.

Higgins, R. W., J. E. Janowiak, and Y. Yao, 1996: A gridded hourly precipitation database for the united states (1963-1993). *NCEP/Climate Prediction Center Atlas*, **1**, 47.

Hollingsworth, A., and P. Lönnberg, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. *Tellus*, **38A**, 111–136.

Holtslag, A. A. M., and A. Beljaars, 1989: Surface flux parameterization schemes: developments and experiences at KNMI. Proceedings of the ECMWF Workshop on Parameterization of Fluxes over Land Surfaces. 24-26 October 1989. Technical report, ECMWF, 121–147.

Houghton, J. T., Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell, and C. A. J. (eds.)., 2001: *(Intergovernmental Panel of Climate Change). Climate Change 2001: The scientific basis.* Cambridge University Press, Cambridge, U. K., 881 pp. pp.

Houghton, J. T., L. G. Meira-Filho, B. A. Callander, N. Harris, A. Kattenberg, and K. M. (eds.)., 1996: (Intergovernmental Panel of Climate Change). Climate Change 1995: The science of climate change. Technical report, Cambridge University Press, Cambridge, U. K. 570 pp.

Hu, Y.-X., and K. Stamnes, 1993: An accurate parameterization of the radiative properties of water clouds suitable for use in climate models. *J. Climatol.*, **6**, 728–742.

Huang, X.-Y., and P. Lynch, 1993: Diabatic digital filtering initialization: application to the HIRLAM model. *Mon. Wea. Rev.*, **121**, 589–603.

Huffman, G. J., and D. T. Bolvin, 2001: Gpcp version 2 combined precipitation data set documentation. Technical report, Available from ftp.ncdc.noaa.gov, at pub/data/gpcp/v2/documentation/V2doc.

Hunt, C. R., and W. H. Snyder, 1980: Experiments on stably stratified flow over a model three-dimensional hill. *J. Fluid Mech.*, **96**, 671–704.

Jaeger, L., 1976: Monatskarten des Niederschlags für die ganze Erde. Technical Report 139, Berichte des Deutschen Wetterdienstes, 38.

Johns, T. C., R. E. Carnell, J. F. Crossley, J. M. Gregory, J. F. B. Mitchell, C. A. Senior, S. F. B. Tett, and R. A. Wood, 1997: The second Hadley Centre coupled ocean-atmosphere GCM: model description, spinup and validation. *Climate Dynamics*, **13**, 103–134.

Joly, A., D. Jorgensen, M. A. Shapiro, A. Thorpe, P. Bessemoulin, K. A. Browning, J.-P. Cammas, J.-P. Chalon, S. A. Clough, K. A. Emanuel, L. Eymard, R. Gall, P. H. Hildebrand, R. H. Langland, Y. Lemaitre, P. Lynch, J. A. Moore, P. O. G. Persson, C. Snyder, and M. R. Wakimoto, 1997: The Fronts and Atlantic Storm-Track Experiment (FASTEX). Scientific Objectives and Experimental Design. *Bull. Am. Met. Soc.*, **78**, 1917–1940.

Joseph, D., 1980: Navy 10' global elevation values. Technical report, NCAR notes on the FNWC terrain data set, 3.

Kain, J. S., and J. M. Fritsch, 1990: A One-Dimensional Entraining/Detraining Plume Model and Its Application in Convective Parameterization. *J. Atm. Sci.*, **47**, 2784–2802.

Kain, J. S., and J. M. Fritsch, 1993: Convective Parameterisation for Mesoscale Models: The Kain-Fritsch scheme. In: The representation of cumulus convection in numerical models. Eds: K. A. Emanuel and D. J. Raymond. *AMS Monograph*, **46**, 246p.

Källén, E., 1996: HIRLAM documentation manual. System 2.5. Technical report, HIRLAM, 178 pp. + 55 pp. appendix. Available from SMHI, S-60176 Norrköping, Sweden.

Kessler, E., 1969: On the Distribution and Continuity of Water Substance in Atmospheric Circulations. *Meteor. Monogr.*, **32**, 84.

Kim, J., and L. Mahrt, 1992: Simple formulation of turbulent mixing in the stable free atmosphere and nocturnal boundary layer. *Tellus*, **44A**, 381–394.

Koster, R. D., and M. J. Suarez, 1992: A comparative analysis of two land surface heterogeneity representations. *J. Climatol.*, **5**, 1379–1390.

Kuo, H. L., 1965: On the formation and intensification of tropical cyclone through latent heat release by cumulus convection. *J. Atm. Sci.*, **22**, 40–63.

Kuo, H. L., 1974: Further Studies of the Parameterization of the Influence of Cumulus Convection on Large-Scale Flow. *J. Atm. Sci.*, **31**, 1232–1240.

Lenderink, G., 2002: An integral mixing length formulation for a TKE-l turbulence closure in atmospheric models. in prep.

Lenderink, G., and W. de Rooy, 2000: A robust mixing length formulation for a TKE-l turbulence scheme. *Hirlam Newsletter*, **36**, 25–29.

Lenderink, G., and A. A. M. Holtslag, 2000: Evaluation of of the kinetic energy approach for modelling turbulent fluxes in stratocumulus. *Mon. Wea. Rev.*, **128**, 244–258.

Lin, B., R. R. Farley, and H. D. Orville, 1983: Bulk parameterization of the snow field in a cloud model. *J. Climate Appl. Meteor.*, **22**, 1065–1092.

Lindskog, M., N. Gustafsson, B. Navascués, K. S. Mogensen, X.-Y. Huang, X. Yang, U. Andræ, L. Berre, S. Thorsteinsson, and J. Rantakokko, 2001: Three-dimensional variational data assimilation for a limited area model. Part II: Observation handling and assimilation experiments. *Tellus*, **53A**, 447–468.

Lindström, G., M. Gardelin, B. Johansson, and S. Bergström, 1997: Development and test of the distributed HBV-96 hydrological model. *J. Hydrol.*, **201**, 272–288.

Lindzen, R. S., 1981: Turbulence and stress due to gravity wave and tidal breakdown. *J. Geoph. Res.*, **86**, 9707–9714.

Liu, W. T., K. B. Katsaros, and J. A. Businger, 1979: Bulk parameterization of air-sea interaction of heat and water vapor including the molecular constraints at the interface. *J. Atm. Sci.*, **36**, 1722–1735.

Ljungemyr, P., N. Gustafsson, and A. Omstedt, 1996: Parameterization of lake thermodynamics in a high resolution weather forecasting model. *Tellus*, **48A**, 608–621.

Lönnberg, P., 1988: Developments in the ECMWF analysis system. Proceedings of the ECMWF Seminar on Data Assimilation and the Use of Satellite Data. Technical report, ECMWF, Reading UK, 75–119.

Lönnberg, P., and A. Hollingsworth, 1986: The statistical structure of short-range forecast errors as determined from radiosonde data. Part II: The covariance of height and wind errors. *Tellus*, **38A**, 137–161.

Lorenc, A., 1986: Analysis methods for numerical weather prediction. *Quart. J. Roy. Met. Soc.*, **112**, 1177–1194.

Lorenc, A., 1988: Optimal nonlinear objective analysis. *Quart. J. Roy. Met. Soc.*, **114**, 205–240.

Lorenc, A., and O. Hammon, 1988: Objective quality control of observations using Bayesian methods. Theory and practical implementation. *Quart. J. Roy. Met. Soc.*, **114**, 515–543.

Lorenc, A. C., 1981: A global three-dimensional multivariate statistical interpolation schme. *Mon. Wea. Rev.*, **109**, 701–721.

Lott, F., and M. J. Miller, 1997: A new subgrid-scale orographic drag parametrization: Its formulation and testing. *Quart. J. Roy. Met. Soc.*, **123**, 101–127.

Louis, J. F., 1979: A parametric model of vertical eddy fluxes in the atmosphere. *Bound. Lay. Met.*, **17**, 187–202.

Louis, J. F., M. Tiedtke, and J.-F. Geleyn, 1982: A short story of the PBL parameterization at ECMWF. Proceedings of the ECMWF Workshop on Planetary Boundary Layer Parameterization, 25-27 November 1981. Technical report, ECMWF, 59–80.

Lynch, P., 1997: The Dolph-Chebyshev Window: A Simple Optimal Filter. *Mon. Wea. Rev.*, **125**, 655–660.

Lynch, P., D. Giard, and V. Ivanovici, 1997: Improving the efficiency of a digital filtering scheme. *Mon. Wea. Rev.*, **125**, 1976–1982.

Lynch, P., and X.-Y. Huang, 1992: Initialization of the HIRLAM model using a digital filter. *Mon. Wea. Rev.*, **120**, 1019–1034.

Lynch, P., R. McGrath, and A. McDonald, 1999: Digital Filter Initialization for HIRLAM. Technical Report 42, HIRLAM Technical Report. Available at SMHI, S-60176 Norrköping, Sweden.

Maat, N., C. Kraan, and W. A. Oost, 1991: The roughness of wind waves. *Bound. Lay. Met.*, **54**, 89–103.

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

Mahfouf, J.-F., 1991: Analysis of Soil Moisture fron Near-Surface Parameters: A Feasibility Study. *J. Appl. Meteor.*, **30**, 1534–1547.

Mahfouf, J.-F., and J. Noilhan, 1996: Inclusion of gravitational drainage in a land surface scheme based on the force restore method. *J. Appl. Meteor.*, **35**, 987–992.

Männik, A., and R. Rõõm, 2001: Nonhydrostatic adiabatic kernel for HIRLAM. Part II: Anelastic, hybrid-coordinate, explicit-Eulerian model. Technical Report 49, HIRLAM Technical Report, 53. Available at SMHI, S-60176 Norrköping, Sweden.

Martin, G. M., D. W. Johnson, and A. Spice, 1994: The measurement and parameterization of effective radius of droplets in warm stratocumulus clouds. *J. Atm. Sci.*, **51**, 1823–1842.

Mason, P. J., 1988: The formation of areally-averaged roughness lengths. *Quart. J. Roy. Met. Soc.*, **114**, 399–420.

McDonald, A., 1999: An examination of alternative extrapolations to find the departure point position in a "two-time-level" semi-lagrangian integration. *Mon. Wea. Rev.*, **127**, 1985–1993.

McGregor, J. L., 1997: Regional climate modeling. *Meteorol. Atmos. Phys.*, **63**, 105–117.

Meier, H. E. M., R. Döscher, A. C. Coward, J. Nycander, and K. Döös, 1999: RCO - Rossby Centre regional ocean climate model: model description (version 1.0) and first results from the hindcast period 1992/93. Technical Report 26, Reports Oceanography, SMHI, Norrköping, Sweden., 102.

Mintz, Y., and Y. Serafini, 1981: Global fields of soil moisture and land-surface evapotranspiration. Technical Report 1980/81, NASA Goddard Space Flight Center Tech. Memo 83907, 178–180.

Mitchell, J. F. B., and T. C. Johns, 1997: On modification of global warming by sulphate aerosols. *J. Climatol.*, **10**, 245–267.

Murphy, A. H., 1997: Forecast Verification. *Economic Value of Weather and Climate Forecasts*, R. W. Katz and A. H. Murphy, Eds. Cambridge University Press.

Nagata, M., L. Leslie, H. Kamahori, R. Nomura, H. Mino, Y. Kurihara, E. Rogers, R. Elsberry, B. Basu, A. Buzzi, J. Calvo, M. Desgagné, M. D'Isidoro, S.-Y. Hong, J. Katzfey, D. Majewski, P. Malguzzi, J. McGregor, A. Murata, J. Nachamkin, M. Roch, and C. Wilson, 2001: A mesoscale model intercomparison in a case of explosive development of a tropical cyclone (COMPARE III). *J. Meteor. Soc. Japan*, **79**, 999–1033.

Nagata, M., L. Leslie, Y. Kurihara, R. Elsberry, M. Yamasaki, H. Kamahori, R. A. Jr, K. Besso, J. Calvo, J. Chan, P. Clark, M. Desgagne, S.-Y. Hong, D. Majewski, P. Malguzzi, J. McGregor, H. Mino, A. Murata, J. Nachamkin, M. Roch, and C. Wilson, 2001: Third COMPARE Workshop: A Model Intercomparison Experiment of Tropical Cyclone Intensity and Track Prediction. *Bull. Am. Met. Soc.*, **89**, 2007–2020.

Navascues, B., 1997: Analysis of 2 m temperature and relative humidity. Technical Report 28, HIRLAM. Available from SMHI, S-60176 Norrköping, Sweden.

NCSA, 1998: HDF User's Guide, Version 4.1r3. Technical report, NCSA Scientific Data Technologies, http://hdf.ncsa.uiuc.edu.

Nielsen, N. W., 1999: A revised formulation of surface fluxes over sea. *HIRLAM 4 Workshop on Physical Parameterization, Madrid, 11-13 November 1998*, 112–120.

Nieuwstadt, F. T. M., 1984: The turbulent structure of the stable, nocturnal boundary layer. *J. Atm. Sci.*, **41**, 2202–2216.

Noilhan, J., and J.-F. Mahfouf, 1996: The ISBA land surface parameterization scheme. *Global and Planetary Change*, **13**, 145–149.

Noilhan, J., and S. Planton, 1989: A simple parameterization of land surface processes for meteorological models. *Mon. Wea. Rev.*, **117**, 536–549.

Ogura, Y., and H.-R. Cho, 1973: Diagnostic determination of cumulus cloud populations from observed large-scale variables. *J. Atm. Sci.*, **30**, 1276–1286.

Olson, J. S., 1994a: Global ecosystem framework-definitions. Technical report, USGS EROS, Data Center Internal Report., Sioux Falls, SD, 37.

Olson, J. S., 1994b: Global ecosystem framework-translation strategy. Technical report, USGS EROS Data Center Internal Report., Sioux Falls, SD, 39.

Omstedt, A., 1999: Forecasting ice on lakes, estuaries and shelf seas. In: Ice Physics in the Natural and Endangered Environment (ed: J S Wettlaufer, J G Dash and N Untersteiner). Technical report, NATO ASI I, Springer-Verlag, Berlin, Heidelberg, Germany, 185–208.

Omstedt, A., and L. Nyberg, 1996: Response of Baltic Sea ice to seasonal, interannual forcing and climate change. *Tellus*, **48A**, 644–662.

Oost, W. A., 1998: The KNMI HEXMAX stress data–a reanalysis. *Bound. Lay. Met.*, **86**, 447–468.

Ou, S.-C., and K. N. Liou, 1995: Ice microphysics and climatic temperature feedback. *Atmos. Res.*, **35**, 127–138.

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

Peltier, W. R., and T. L. Clark, 1979: Evolution and stability of finite-amplitude mountain waves, Pt. 2, Surface wave drag and severe downslope windstorms. *J. Atm. Sci.*, **36**, 1498–1529.

Queney, P., 1948: The problem of air flow over mountains: A summary of theoretical studies. *Bull. Am. Met. Soc.*, **29**, 16–25.

Rabier, F., A. McNally, E. Andersson, P. Courtier, P. Undén, J. Eyre, A. Hollingsworth, and F. Bouttier, 1998: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). Part II: Structure functions. *Quart. J. Roy. Met. Soc.*, **124**, 1809–1830.

Räisänen, J., and R. Joelsson, 2001: Changes in average and extreme precipitation in two regional climate model experiments. *Tellus*, **53A**, 547–566.

Räisänen, J., M. Rummukainen, and A. Ullerstig, 2001: Downscaling of greenhouse gas induced climate change in two GCMs with the Rossby Centre regional climate model for northern europe. *Tellus*, **53A**, 168–191.

Räisänen, J., M. Rummukainen, A. Ullerstig, B. Bringfelt, U. Hansson, and U. Willén, 1999: The first Rossby Centre regional climate scenario - Dynamical downscaling of CO2-induced climate change in the HadCM2 GCM. Technical Report 85, SMHI Reports Meteorology and Climatology (RMK), 56.

Räisänen, P., M. Rummukainen, and J. Räisänen, 2000: Modification of the Hirlam radiation scheme for use in the Rossby Centre regional atmospheric climate model. Technical Report 49, Department of Meteorology, University of Helsinki, 71.

Rasch, P. J., and J. E. Kristjánsson, 1998: A comparison of the CCM3 model climate using diagnosed and predicted condensate parameterizations. *J. Climatol.*, **11**, 1587–1614.

Raymond, W. H., 1988: High-order low-pass implicit tangent filters for use in finite area Calculations. *Mon. Wea. Rev.*, **116**, 2132–2141.

Redelsperger, J. L., and G. Sommeria, 1986: Three-Dimensional Simulation of a Convective Storm: Sensitivity Studies on Subgrid Parameterization and Spatial Resolution. *J. Atm. Sci.*, **43**, 2619–2635.

Ritchie, H., and M. Tanguay, 1996: A comparison of spatially averaged Eulerian and semi-Lagrangian treatments of mountains. *Mon. Wea. Rev.*, **124**, 167–181.

Rodriguez-Camino, E., and R. Avissar, 1998: Comparison of three land-surface schemes with the Fourier Amplitude Sensitivity Test (FAST). *Tellus*, **50A**, 313–332.

Rodriguez-Camino, E., and R. Avissar, 1999: Effective parameters to heat surface fluxes in heterogeneous terrain. *Tellus*, **51A**, 387–399.

Roeckner, E., L. Bengtsson, J. Feichter, J. Lelieveld, and H. Rodhe, 1999: Transient climate change simulations with a coupled atmosphere-ocean GCM including the tropospheric sulfur cycle. *J. Climatol.*, **12**, 3004–3032.

Rogers, R. R., and M. K. Yau, 1989: *A short course in cloud physics. Third edition.* Pergamon Press, 293 pp. pp.

Rontu, L., K. Sattler, and R. Sigg, 2002: Parametrization of subgrid-scale orography effects in HIRLAM. Technical Report 56, HIRLAM, 46. Available at SMHI, S-60176 Norrköping, Sweden.

Rõõm, R., 2001: Nonhydrostatic adiabatic kernel for HIRLAM. Part I: Fundamentals of nonhydrostatic dynamics in pressure-related coodinates. Technical Report 48, HIRLAM Technical Report, 25. Available at SMHI, S-60176 Norrköping, Sweden.

Rõõm, R., and A. Männik, 2002: Nonhydrostatic adiabatic kernel for HIRLAM. Part III: Semi-implicit Eulerian scheme. Technical Report 55, HIRLAM Technical Report, 26. Available at SMHI, S-60176 Norrköping, Sweden.

Rubel, F., and M. Hantel, 2001: Baltex 1/6-degree daily precipitation climatology 1996-1998. *Meteorol. Atmos. Phys.*, **77**, 155–166.

Rummukainen, M., J. Räisänen, B. Bringfelt, A. Ullerstig, A. Omstedt, U. Willén, U. Hansson, and C. Jones, 2001: A regional climate model for northern Europe - model description and results from the downscaling of two GCM control simulations. *Climate Dynamics*, **17**, 339–359.

Rummukainen, M., J. Räisänen, A. Ullerstig, B. Bringfelt, U. Hansson, P. Graham, and U. Willén, 1998: RCA - Rossby Centre regional Atmospheric climate model: model description and results from the first multi-year simulation. Technical Report 83, SMHI Reports Meteorology and Climatology (RMK), 76.

Sass, B. H., and J. H. Christensen, 1995: A simple framework for testing the quality of atmospheric limited-area models. *Mon. Wea. Rev.*, **123**, 444–459.

Sass, B. H., L. Rontu, and P. Räisänen, 1994: HIRLAM-2 Radiation Scheme: Documentation and Tests. Technical Report 16, The HIRLAM-3 Project, SMHI, S-60176 Norrköping, Sweden.

Sattler, K., 1999: New high resolution physiographic data and climate generation for the HIRLAM forecasting system. *DMI Tech. Rep.*, **99-11**.

Savijärvi, H., 1990: Fast radiation parameterization schemes for mesoscale and short-range forecast models. *J. Appl. Meteor.*, **29**, 437–447.

Savijärvi, H., A. Arola, and P. Räisänen, 1997: Shortwave optical properties of precipitating waterclouds. *Quart. J. Roy. Met. Soc.*, **123**, 883–899.

Savijärvi, H., and P. Räisänen, 1998: Longwave optical properties of water clouds and rain. *Tellus*, **50A**, 1–11.

Schumann, U., and T. Gertz, 1995: Turbulent mixing in stably stratified shear flows. *J. Appl. Meteor.*, **34**, 33–48.

Scinocca, J. F., and N. A. McFarlane, 2000: The parametrization of drag induced by stratified flow over anisotropic orography. *Quart. J. Roy. Met. Soc.*, **126**, 2353–2393.

Scorer, R. S., 1949: Theory of waves in the lee of mountains. *Quart. J. Roy. Met. Soc.*, **75**, 41–56.

Shaw, D. B., P. Lönnberg, A. Hollingsworth, and P. Undén, 1987: Data assimilation: The 1984/85 revisions of the ECMWF mass and wind analysis. *Quart. J. Roy. Met. Soc.*, **113**, 533–566.

Simpson, J., 1983: Cumulus clouds: interactions between laboratory experiments and observations as foundations for models. *Mesoscale meteorology.* Reidel, 399–412.

Slingo, J. M., 1987: The development and verification of a cloud prediction scheme for the ECMWF model. *Quart. J. Roy. Met. Soc.*, **113**, 899–927.

Stössel, A., and M. Claussen, 1993: On the momentum forcing of a large-scale sea-ice model. *Climate Dyn.*, **9**, 71–80.

Stull, R. B., 1988: *An introduction to boundary layer meteorology.* Kluwer Academic Publishers, London & High Wycombe, 666 pp.

Sundqvist, H., 1988: Parameterization of condensation and associated clouds in models for weather prediction and general circulation simulation. Physically Based Modelling and Simulation of Climate and Climatic Change, M E Schlesinger, Ed. *Kluwer Academic*, **1**, 433–461.

Sundqvist, H., 1993: Inclusion of Ice Phase of Hydrometeors in Cloud Parameterization for Mesoscale and Largescale Models. *Contr. Atm. Phys.*, **66**, 137–147.

Sundqvist, H., E. Berge, and J. E. Kristjánsson, 1989: Condensation and Cloud Parameterization Studies with a Mesoscale Numerical Weather Prediction Model. *Mon. Wea. Rev.*, **117**, 1641–1657.

Taljaard, J. J., H. van Loon, H. J. Crutcher, and R. L. Jenne, 1969: *Climate of the upper air, Part 1 – Southern Hemisphere; Temperatures, dew points and heights at selected pressure levels. NAVAIR Atlas 50-1C-55.* Government Printing Office, Washington D.C., 135 pp.

Tanguay, M., E. Yakimiw, H. Ritchie, and A. Robert, 1992: Advantages of spatial averaging in semi-implicit and semi-Lagrangian Schemes. *Mon. Wea. Rev.*, **120**, 113–123.

Temperton, C., 1988: Implicit normal mode initialization. *Mon. Wea. Rev.*, **116**, 1013–1031.

Tibaldi, S., and J.-F. Geleyn, 1981: The Production of a new Orography, Land-Sea Mask and associated Climatological Surface Fields for Operational Purposes. Technical Report 40, ECMWF Research Department Tech. Memorandum, 13 + Appendices and Figures.

Tiedtke, M., 1993: Representation of Clouds in Large-Scale Models. *Mon. Wea. Rev.*, **121**, 3040–3061.

Trenberth, K. E., 1997: Using atmospheric budgets as a constraint on surface fluxes. *J. Climatol.*, **10**, 2796–2809.

Tripoli, G. J., and W. R. Cotton, 1980: A numerical investigation of several factors contributing to the observed variable intensity of deep convection over south Florida. *J. Appl. Meteor.*, **19**, 1037–1063.

Undén, P., 1989: Tropical data assimilation and analysis of divergence. *Mon. Wea. Rev.*, **117**, 2495–2517.

Undén, P., 1999: Analysis Modifications of box sixes, moving platform check and observation errors. *HIRLAM Newsletter*, **34**, 15–32, Available at SMHI, S-60176 Norrköping, Sweden.

USGS, 1997: The Global Land Characteristics Data Base. Technical report, U.S. Geological Survey, http://edcdaac.usgs.gov/glcc/glcc.html.

USGS, 1998: GTOPO30, Global 30 Arc Second Elevation data set. Technical report, U.S. Geological Survey, http://edcdaac.usgs.gov/gtopo30/gtopo30.html.

Valcke, S., L. Terray, and A. Piacentini, 2000: Oasis 2.4, Ocean atmosphere sea ice soil: user's guide. Technical report, Technical Report TR/CMGC/00/10, CERFACS, Toulouse, France.

Viterbo, P., A. Beljaars, J.-F. Mahfouf, and J. Teixeira, 1999: The representation of soil moisture freezing and its impact on the stable boundary layer. *Quart. J. Roy. Met. Soc.*, **125**, 2401–2426.

Wilson, M. F., and A. Henderson-Sellers, 1985: A global archive of land cover and soils data sets for use in General Circulation Models. *J. Climatol.*, **5**, 119–143.

Wyser, K., L. Rontu, and H. Savijärvi, 1999: Introducing the effective radius into a fast radiation scheme of a mesoscale model. *Contr. Atm. Phys.*, **72**, 205–218.

Xu, K.-M., and D. A. Randall, 1996: A Semiempirical cloudiness parameterization for use in climate models. *J. Atm. Sci.*, **53**, 3084–3101.

Yelland, M., and T. K. Taylor, 1996: Wind stress measurements from the open ocean. *J. Phys. Oceanog.*, **26**, 541–558.

Appendices

List of variables and constants

a	:	Radius of the earth (6.371×10^6 m)
c_{pd}	:	Specific heat of dry air (1004.64 J kg^{-1}K^{-1})
c_{pv}	:	Specific heat of moist air (1869.46 J kg^{-1}K^{-1})
f	:	Coriolis parameter (dynamics) (s^{-1})
	:	Cloud fraction (0-1)
g	:	Acceleration due to gravity (9.80665 m s^{-2})
G	:	Linearised geopotential (g $\times$ height) (m^2 s^{-2})
K	:	Horizontal diffusion coefficient
L	:	Latent heat of vaporization/sublimation (J kg^{-1})
m	:	Cloud water mixing ratio (kg m^{-2})
N	:	Number of levels in vertical
p	:	Pressure (Pa)
p^r	:	Reference pressure for semi-implicit scheme (Pa)
p_s	:	Surface pressure (Pa)
q	:	Specific humidity (kg kg^{-1})
q_c	:	Cloud water and ice specific humidity (kg kg^{-1})
q_t	:	Total specific humidity $q_t = q + q_c$ (kg kg^{-1})
R_d	:	Gas constant for dry air (287.04 J kg^{-1} K^{-1})
R_v	:	Gas constant for water vapour (451.51 J kg^{-1} K^{-1})
$\dot{s}$	:	Vertical velocity
t	:	Time (s)
T	:	Temperature (K)
T^0	:	Constant temperature for semi-implicit scheme(K)
T_v	:	Virtual temperature (K)
u	:	Zonal velocity (m s^{-1})
v	:	Meridional velocity (m s^{-1})
δ	:	c_{pv}/c_{pd}
ϵ	:	R_d/R_v
ϵ_g	:	De-centered damping coefficient
ϵ_N	:	Filter coefficient
η	:	Vertical coordinate
$\dot{\eta}$	:	η -vertical velocity
κ	:	R_d/c_{pd}
λ	:	Longitude
Ω	:	Angular speed of the earth (7.292×10^{-5} s^{-1})
Φ	:	Geopotential (g $\times$ height) (m^2 s^{-2})
Φ_s	:	Surface geopotential (m^2 s^{-2})
θ	:	Latitude
	:	Potential temperature (K)

Definition of the matrix operators

$$(\gamma \mathbf{T})_k = R_d(\alpha^r T)_k + R_d \sum_{j=k+1}^{N} (T \Delta lnp^r)_j$$

$$(\tau \mathbf{D})_k = \kappa T^0 \left[\left(\frac{\Delta lnp}{\Delta p} \right)_k^r \sum_{j=1}^{k-1} D_j \Delta p_j^r + \alpha_k^r D_k \right]$$

$$\nu.\mathbf{D} = \sum_{j=1}^{N} D_j \frac{\Delta p_j}{p^r}$$

The Specified Observation Errors of HIRLAM 3D-Var

Upper air data

Pressure (hPa)	TEMP u/v (m/s)	SATOB u/v (m/s)	AIREP u/v (m/s)	PILOT u/v (m/s)	TEMP T (K)	AIREP T (K)
1000	2.1	2.0	2.5	2.3	1.1	1.4
850	2.0	2.0	2.5	2.3	0.9	1.3
700	2.0	2.0	3.0	2.5	0.8	1.2
500	2.5	3.5	3.5	3.0	0.8	1.2
400	3.1	4.3	4.0	3.5	0.8	1.2
300	3.6	5.0	4.0	3.7	1.0	1.3
250	3.8	5.0	4.0	3.5	1.1	1.3
200	3.5	5.0	4.0	3.5	1.1	1.4
150	3.1	5.0	4.0	3.4	1.1	1.4
100	2.9	5.0	4.0	3.3	1.0	1.4
70	2.5	5.0	4.0	3.2	0.9	1.5
50	2.3	5.0	4.0	3.2	1.0	1.6
30	2.3	5.0	4.0	3.3	1.1	1.8
20	2.2	5.0	4.0	3.6	1.1	2.0
10	2.2	5.7	4.0	4.5	1.1	2.2

Surface data

Pressure (hPa)	DRIBU Z (m)	SYNOP Z (m)	SHIP Z (m)	DRIBU u/v (m/s)	SYNOP u/v (m/s)	SHIP u/v (m/s)	DRIBU T (K)	SYNOP T (K)	SHIP T (K)
1000	11.5	7.0	14.0	2.4	3.0	3.0	1.8	2.0	1.8
850		8.0			3.0			1.5	
700		8.6			3.0			1.3	
500		12.1			3.4			1.2	
400		14.9			3.6			1.3	
300		18.8			3.8			1.5	

Output variables of HIRLAM

Quantities in boldface are included in the history files used by model.

State variables and diagnostics on model surfaces

Variable	unit	GRIB-code IPAR	GRIB-code LTYP	note
wind: x-component	ms^{-1}	033	109	relative to the rotated coordinates, given at u-points
wind: y-component	ms^{-1}	034	109	relative to the rotated coordinates, given at v-points
temperature	K	011	109	
specific humidity	kg $kg^{-}1$	051	109	
specific cloud condensate	kg $kg^{-}1$	076	109	
kinetic energy of turbulence	m^2s^{-2}	200	109	given at model half-levels
geopotential	m^2s^{-2}	006	109	
potential temperature	K	013	109	
relative humidity	fraction	052	109	
hybrid coordinate vertical velocity	$s^{-}1$	038	109	code definition in violation of the GRIB standard
pressure coordinate vertical velocity	Pa $s^{-}1$	039	109	
total cloud cover	fraction	071	109	
convective cloud cover	fraction	072	109	
flux of precipitating water	$kgm^{-}2$	061	109	accumulated

State variables and diagnostics on constant pressure surfaces

Variable	unit	GRIB-code IPAR	GRIB-code LTYP	note
geopotential	m^2s^{-2}	006	100	
temperature	K	011	100	
potential temperature	K	013	100	
dew point temperature	K	017	100	
u-component of wind	ms^{-1}	033	100	relative to the rotated coordinates, given at u-points
v-component of wind	ms^{-1}	034	100	relative to the rotated coordinates, given at v-points
pressure coordinate vertical velocity	Pa s^{-1}	039	100	
relative humidity	fraction	052	100	
specific cloud condensate	kg $kg^{-}1$	076	100	
total cloud cover	fraction	071	100	

Surface and soil variables

Variable	unit	GRIB-code IPAR	LTYP	LEV	note
surface pressure	Pa	001	105	0	not the mean sea level pressure
surface pressure tendency	$Pa\ s^{-1}$	003	105	0	
surface geopotential	m^2s^{-2}	006	105	0	geopotential relative to mean sea level
fraction of land	fraction	081	105	0	fraction of grid-square covered by land
fraction of ice	fraction	091	102	0	fractional ice coverage in non-land part of grid-square
sea surface temperature	K	011	102	0	
surface temperature	K	011	105	0	average over all subtypes
mean surface temperature	K	011	105	999	average over all subtypes
deep soil temperature	K	011	105	998	climatological value
surface soil water content	m^3m^{-3}	086	105	0	average over all subtypes
total soil water content	m^3m^{-3}	086	105	999	average over all subtypes
deep soil wetness	m	086	105	998	climatological value
water equivalent of snowpack	m	066	105	0	
background albedo	m	084	105	0	not used by the present forecast model
actual albedo	m	169	105	0	grid-square average with solar height correction
orographic roughness length	m	083	105	0	
roughness length over sea	m	083	102	0	analysis files contain a background-value
standard deviation of mesoscale orography	m	204	105	0	to be implemented
anisotropy of mesoscale orography	0...1	205	105	0	to be implemented
direction of mesoscale orography	rad	206	105	0	to be implemented

Near-surface diagnostics

		GRIB-code			
Variable	unit	IPAR	LTYP	LEV	note
mean sea level pressure	Pa	002	105	0	
(mean sea level pressure)	Pa	001	103	0	obsole definition in violation of the GRIB standard
screen level temperature	K	011	105	2	2 m above ground
screen level dew point temperature	K	017	105	2	2 m above ground
x-component of wind at 10 m height	ms^{-1}	033	105	10	relative to the rotated coordinates, given at mass-points
y-component of wind at 10 m height	ms^{-1}	034	105	10	relative to the rotated coordinates, given at mass-points
screen level specific humidity	kg $kg^{-}1$	051	105	2	2 m above ground
screen level relative humidity	$fraction^{-}1$	052	105	2	2 m above ground
depth of the boundary layer	m	067	105	0	
total cloud cover	fraction	071	105	0	
low level cloud cover	fraction	073	105	0	
medium level cloud cover	fraction	074	105	0	
high level cloud cover	fraction	075	105	0	

Accumulated surface fluxes

Variable	unit	GRIB-code IPAR	LTYP	LEV	note
total precipitation	kg m^{-2}	061	105	0	
grid-scale precipitation	kg m^{-2}	062	105	0	
sub grid-scale precipitation	kg m^{-2}	063	105	0	
total snowfall	kg m^{-2}	065	105	0	
grid-scale snowfall	kg m^{-2}	079	105	0	
sub grid-scale snowfall	kg m^{-2}	078	105	0	
evaporation	kg m^{-2}	057	105	0	to be implemented
surface water run-off	kg m^{-2}	090	105	0	to be implemented
net short-wave radiation	J m^{-2}	111	105	0	
net long-wave radiation	J m^{-2}	112	105	0	
downwelling long-wave radiation	J m^{-2}	115	105	0	
downwelling (global) short-wave radiation	J m^{-2}	117	105	0	
momentum flux due to mesoscale orography(u-component)	N m^{-2}s	119	105	0	relative to the rotated grid, given at mass-points
momentum flux due to mesoscale orography (v-component)	N m^{-2}s	120	105	0	relative to the rotatated grid, given at mass-points
latent heat flux	J m^{-2}	121	105	0	
sensible heat flux	J m^{-2}	122	105	0	
turbulent momentum flux (u-component)	N m^{-2}s	124	105	0	relative to the rotatated grid, given at mass-points
turbulent momentum flux (v-component)	N m^{-2}s	125	105	0	relative to the rotatated grid, given at mass-points

Accumulated fluxes at the top of the atmosphere

Variable	unit	GRIB-code IPAR	LTYP	LEV	note
net short-wave radiation	J m^{-2}	113	008	0	
net long-wave radiation	J m^{-2}	114	008	0	
downwelling short-wave radiation	J m^{-2}	117	008	0	

Vertically-integrated quantities

		GRIB-code			
Variable	unit	IPAR	LTYP	LEV	note
precipitable water	kg m^{-2}	054	105	0	
cloud condensate	kg m^{-2}	076	105	0	
change of temperature due to parameterized processes	K	210	200	0	accumulated
change of temperature due to turbulent mixing	K	211	200	0	accumulated
change of temperature due to radiation	K	212	200	0	accumulated
change of specific humidity due to parameterized processes	kg $kg^{-}1$	220	200	0	accumulated
change of specific humidity due to turbulent mixing	kg $kg^{-}1$	221	200	0	accumulated
change of specific cloud condensate due to parameterized processes	kg $kg^{-}1$	230	200	0	accumulated
change of specific cloud condensate due to turbulent mixing	kg $kg^{-}1$	231	200	0	accumulated
change of u-momentum due to parameterized processes	$ms^{-}1$	240	200	0	accumulated
change of u-momentum due to turbulent mixing	$ms^{-}1$	241	200	0	accumulated
change of v-momentum due to parameterized processes	$ms^{-}1$	245	200	0	accumulated
change of v-momentum due to turbulent mixing	$ms^{-}1$	246	200	0	accumulated

Change of state variables on model surfaces

Variable	unit	GRIB-code IPAR	LTYP	note
change of temperature due to parameterized processes	K	210	109	accumulated
change of temperature due to turbulent mixing	K	211	109	accumulated
change of temperature due to radiation	K	212	109	accumulated
change of specific humidity due to parameterized processes	kg kg^{-}1	220	109	accumulated
change of specific humidity due to turbulent mixing	kg kg^{-}1	221	109	accumulated
change of specific cloud condensate due to parameterized processes	kg kg^{-}1	230	109	accumulated
change of specific cloud condensate due to turbulent mixing	kg kg^{-}1	231	109	accumulated
change of u-momentum due to parameterized processes	ms^{-}1	240	109	accumulated
change of u-momentum due to turbulent mixing	ms^{-}1	241	109	accumulated
change of v-momentum due to parameterized processes	ms^{-}1	245	109	accumulated
change of v-momentum due to turbulent mixing	ms^{-}1	246	109	accumulated

Surface subtypes

A coding system violating WMO GRIB standard is presently used to define numbers for near-surface subtype variables. WMO code numbers are used to denote the variables, with the level type set to 105 for all of them. To define the actual level type the WMO *level height* parameter is used. There are presently five subtypes of surface variables:

JTYPE	subtype	note
1	sea or inland water	
2	sea or inland ice	
3	bare land	
4	low vegetation	
5	forest	
6	snow	to be implemented

Surface sublevel types

Three indexes are used to denote level type of surface variables:

LTYPE	level type	note
800	diagnostic levels	used for ten-metre wind, two-metre temperature and humidity
900	surface	
950	in-soil level (mean values)	

Near-surface state and diagnostic variables for subtypes

The variable written as *level height* parameter is obtained by combination of subtype and sublevel parameters: LEV = LTYPE + JTYPE. E.g. the two-metre temperature over forest is now denoted by combination of parameters IPAR = 11, LTYP = 105, LEV = 805. The following variables may have subtype values:

		GRIB-code			
Variable	unit	IPAR	LTYP	LEV	note
screen level temperature	K	011	105	801-805	2 m above ground
screen level specific humidity	kg kg^{-}1	051	105	801-805	2 m above ground
x-component of wind at 10 m height	ms^{-1}	033	105	801-805	relative to the rotated coordinates, given at mass-points
y-component of wind at 10 m height	ms^{-1}	034	105	801-805	relative to the rotated coordinates, given at mass-points
surface temperature	K	011	105	901-905	901 = 011-102-0
mean surface temperature	K	011	105	951-955	
surface specific humidity	kg kg^{-}1	051	105	901-905	
surface soil water content	m^{3}m^{-3}	086	105	901-905	
total soil water content	m^{3}m^{-3}	086	105	951-955	
water equivalent of snowpack	m	066	105	901-905	
sensible heat flux	J m^{-2}	122	105	901-905	
latent heat flux	J m^{-2}	121	105	901-905	
scalar momentum flux	N m^{-2}s	128	105	901-905	given at mass-points
canopy water	kgm^{-2}	192	105	901-905	
fraction of surface type	fraction	194	105	901-905	fractions of water, ice, bare soil, low vegetation, forest
soil type		195	105	901-905	code numbers 1...11
vegetation type		199	105	901-905	