INITIALIZATION

Peter Lynch, University College Dublin, Dublin, Ireland
Xiang-Yu Huang, National Center for Atmospheric Research, Colorado, USA

1 Introduction

The spectrum of atmospheric motions is vast, encompassing phenomena having periods ranging from seconds to millennia. The motions of interest to the forecaster typically have timescales of a day or longer, but the mathematical models used for numerical prediction describe a broader span of dynamical features than those of direct concern. For many purposes these higher frequency components can be regarded as *noise* contaminating the motions of meteorological interest. The elimination of this noise is achieved by adjustment of the initial fields, a process called *initialization*.

The natural oscillations of the atmosphere fall into two groups (see *e.g.* Kasahara (1976)). The solutions of meteorological interest have low frequencies and are close to geostrophic balance. They are called rotational or vortical modes, since their vorticity is greater than their divergence; if divergence is ignored, these modes reduce to the Rossby-Haurwitz waves. There are also very fast gravity-inertia wave solutions, with phase speeds of hundreds of metres per second and large divergence. For typical conditions of large scale atmospheric flow (when the Rossby and Froude numbers are small) the two types of motion are clearly separated and interactions between them are weak. The high frequency gravity-inertia waves may be locally significant in the vicinity of steep orography, where there is strong thermal forcing or where very rapid changes are occurring; but overall they are of minor importance.

A subtle and delicate state of balance exists in the atmosphere between the wind and pressure fields, ensuring that the fast gravity waves have much smaller amplitude than the slow rotational part of the flow. Observations show that the pressure and wind fields in regions not too near the equator are close to a state of geostrophic balance and the flow is quasi-nondivergent. The bulk of the energy is contained in the slow rotational motions and the amplitude of the high frequency components is small. The existence of this geostrophic balance is a perennial source of interest; it is a consequence of the forcing mechanisms and dominant modes of hydrodynamic instability and of the manner in which energy is dispersed and dissipated in the atmosphere. For a recent review of balanced flow, see McIntyre (2003). The gravity-inertia waves are instrumental in the process by which the balance is maintained, but the nature of the sources of energy ensures that the low frequency components predominate in the large scale flow. The atmospheric balance is subtle, and difficult to specify precisely. It is *delicate* in that minor perturbations may disrupt it but *robust* in that local imbalance tends to be rapidly removed through radiation of gravity-inertia waves in a process known as geostrophic adjustment.

When the basic equations are used for numerical prediction the forecast may contain spurious large amplitude high frequency oscillations. These result from anomalously large gravity-inertia waves which occur because the balance between the mass and velocity fields is not reflected faithfully in the analysed fields. High frequency oscillations of large amplitude are engendered, and these may persist for a considerable time unless strong dissipative processes are incorporated in the forecast model. It was the presence of such imbalance in the initial fields which gave rise to the totally unrealistic pressure tendency of 145 hPa/6h obtained by Lewis Fry Richardson in the first-ever objective numerical weather forecast (Richardson, 1922; Lynch, 2006).

Although they have little effect on the long-term evolution of the flow, gravity waves may profoundly influence the way it changes on shorter time-scales. Fig. 1 schematically depicts the pressure variation over a period of one day. The smooth curve represents the variation due to meteorological effects; its gentle slope (dotted line) indicates the long-term change (Phillips,

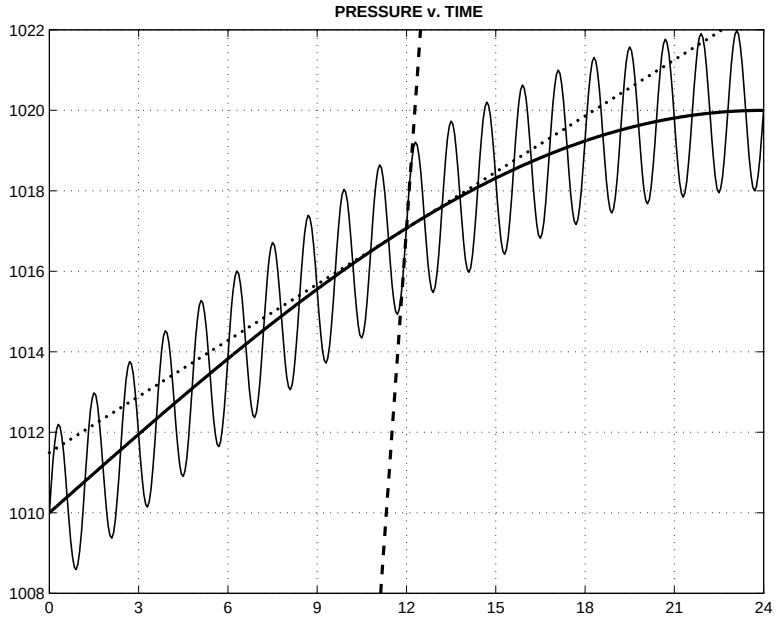

Figure 1: Schematic illustration of pressure variation over a 24 hour period. The thick line is the mean, long-term variation, the thin line is the actual pressure, with high frequency noise. The dotted line shows the rate of change, at 12 hours, of the mean pressure and the dashed line shows the corresponding rate of change of the actual pressure (after Phillips, 1973).

1973). The rapidly varying curve represents the actual pressure changes when gravity waves are superimposed on the meteorological flow: the slope of the oscillating curve (dashed line) is precipitous and, if used to determine long-range variations, yields totally misleading results. What Richardson calculated was the instantaneous rate of change in pressure for an atmospheric state having gravity wave components of large amplitude.

If the fields are not initialized, the spurious oscillations which occur in the forecast can lead to various problems. In particular, new observations are checked for accuracy against a short-range forecast. If this forecast is noisy, good observations may be rejected or erroneous ones accepted. Thus, *initialization is essential for satisfactory data assimilation.* Another problem occurs with precipitation forecasting. A noisy forecast has unrealistically large vertical velocity. This interacts with the humidity field to give hopelessly inaccurate rainfall patterns. To avoid this *spin-up*, we must control the gravity wave oscillations.

2 Early Initialization Methods

The Filtered Equations

The first computer forecast was made in 1950 by Charney, Fjørtoft and Von Neumann (1950). In order to avoid Richardson's error, they modified the prediction equations in such a way as to eliminate the high frequency solutions. This process is known as filtering. The basic filtered system is the set of quasi-geostrophic equations. These equations were used in operational forecasting for a number of years. However, they involve approximations which are not always valid, and this can result in poor forecasts. A more accurate filtering of the primitive equations leads to the *balance equations.* This system is more complicated to solve than the quasi-geostrophic system, and has not been widely used.

Static Initialization

Hinkelmann (1951) investigated the problem of noise in numerical integrations and concluded that if the initial winds were geostrophic, high frequency oscillations would occur but would remain small in amplitude. He later succeeded in integrating the primitive equations, using a very short timestep, with geostrophic initial winds (Hinkelmann, 1959). Forecasts made with the primitive equations were soon shown to be clearly superior to those using the quasi-geostrophic system. However, the use of geostrophic initial winds had a huge disadvantage: the valuable information contained in the observations of the wind field was completely ignored. Moreover, the remaining noise level is not tolerable in practice. Charney (1955) proposed that a better estimate of the initial wind field could be obtained by using the nonlinear balance equation. This equation — part of the balance system — is a diagnostic relationship between the pressure and wind fields. It implies that the wind is nondivergent. It was later argued by Phillips (1960) that a further improvement would result if the divergence of the initial field were set equal to that implied by quasi-geostrophic theory. Each of these steps represented some progress, but the noise problem still remained essentially unsolved.

Dynamic Initialization

Another approach, called dynamic initialization, uses the forecast model itself to define the initial fields (Miyakoda and Moyer, 1968). The dissipative processes in the model can damp out high frequency noise as the forecast procedes. We integrate the model first forward and then backward in time, keeping the dissipation active all the time. We repeat this forward-backward cycle many times until we finally obtain fields, valid at the initial time, from which the high frequency components have been damped out. The forecast starting from these fields is noise-free. However, the procedure is expensive in computer time, and damps the meteorologically significant motions as well as the gravity waves, so it is no longer popular. Digital filtering initialization, described below, is essentially a refinement of dynamic initialization. Because it used a highly selective filtering technique, is is computationally more efficient than the older method.

Variational Initialization

An elegant initialization method based on the calculus of variations was introduced by Sasaki (1958). We consider the simplest case: given an analysis of the mass and wind fields, how can they be minimally modified so as to impose geostrophic balance? This problem can be formulated as the minimization of an integral representing the deviation of the resulting fields from balance. The variation of the integral leads to the Euler-Lagrange equations, which yield diagnostic relationships for the new mass and wind fields in terms of the incoming analysis. Although the method was not widely used, the variational method is now at the centre of modern data assimilation practice. In §6 below we discuss the use of a digital filter as a weak constraint in four-dimensional variational assimilation (4DVAR; Talagrand variational assimilation chapter).

3 Atmospheric Normal Mode Oscillations

The solutions of the model equations can be separated, by a process of spectral analysis, into two sets of components or linear normal modes, slow rotational components or Rossby modes, and high frequency gravity modes. We assume that the amplitude of the motion is so small that all nonlinear terms can be neglected. The horizontal structure is then governed by a system equivalent to the linear shallow water equations which describe the small-amplitude motions of a shallow layer of incompressible fluid. These equations were first derived by Laplace in his discussion of tides in the atmosphere and ocean, and are called the Laplace Tidal Equations.

The simplest means of deriving the linear shallow water equations from the primitive equations is to assume that the vertical velocity vanishes identically.

The Laplace Tidal Equations

Let us assume that the motions under consideration can be described as small perturbations about a state of rest, in which the temperature is a constant, T_0, and the pressure $\bar{p}(z)$ and density $\bar{\rho}(z)$ vary only with height. The basic state variables satisfy the gas law and are in hydrostatic balance: $\bar{p} = \mathcal{R}\bar{\rho}T_0$ and $d\bar{p}/dz = -g\bar{\rho}$. The variations of mean pressure and density follow immediately:

$$\bar{p}(z) = p_0 \exp(-z/H)\,, \qquad \bar{\rho}(z) = \rho_0 \exp(-z/H)\,,$$

where $H = p_0/g\rho_0 = \mathcal{R}T_0/g$ is the scale-height of the atmosphere. We consider only motions for which the vertical component of velocity vanishes identically, $w \equiv 0$. Let u, v, p and ρ denote variations about the basic state, each of these being a small quantity. The horizontal momentum, continuity and thermodynamic equations, with standard notation, are (see chapters on modelling - Rood, and meteorology and dynamics - Lahoz et al.)

$$\frac{\partial \bar{\rho}u}{\partial t} - f\bar{\rho}v + \frac{\partial p}{\partial x} = 0 \tag{1}$$

$$\frac{\partial \bar{\rho}v}{\partial t} + f\bar{\rho}u + \frac{\partial p}{\partial y} = 0 \tag{2}$$

$$\frac{\partial \rho}{\partial t} + \nabla\cdot\bar{\rho}\mathbf{V} = 0 \tag{3}$$

$$\frac{1}{\gamma\bar{p}}\frac{\partial p}{\partial t} - \frac{1}{\bar{\rho}}\frac{\partial \rho}{\partial t} = 0 \tag{4}$$

Density can be eliminated from the continuity equation, Eq. 3, by means of the thermodynamic equation, Eq. 4. Now let us assume that the horizontal and vertical dependencies of the perturbation quantities are separable:

$$\left\{ \begin{array}{c} \bar{\rho}u \\ \bar{\rho}v \\ p \end{array} \right\} = \left\{ \begin{array}{c} U(x,y,t) \\ V(x,y,t) \\ P(x,y,t) \end{array} \right\} Z(z)\,. \tag{5}$$

The momentum and continuity equations can then be written

$$\frac{\partial U}{\partial t} - fV + \frac{\partial P}{\partial x} = 0 \tag{6}$$

$$\frac{\partial V}{\partial t} + fU + \frac{\partial P}{\partial y} = 0 \tag{7}$$

$$\frac{\partial P}{\partial t} + (gh)\nabla\cdot\mathbf{V} = 0 \tag{8}$$

where $\mathbf{V} = (U, V)$ is the momentum vector and $h = \gamma H = \gamma\mathcal{R}T_0/g$. This is a set of three equations for the three dependent variables U, V, and P. They are mathematically isomorphic to the Laplace Tidal Equations with a mean depth h. The quantity h is called the equivalent depth. There is no dependence in this system on the vertical coordinate z.

The vertical structure follows from the hydrostatic equation, together with the relationship $p = (\gamma g H)\rho$ implied by the thermodynamic equation. It is determined by

$$\frac{dZ}{dz} + \frac{Z}{\gamma H} = 0\,, \tag{9}$$

the solution of which is $Z = Z_0 \exp(-z/\gamma H)$, where Z_0 is the amplitude at $z = 0$. If we set $Z_0 = 1$, then U, V and P give the momentum and pressure fields at the earth's surface. These variables all decay exponentially with height. It follows from Eq. 5 that u and v actually increase with height as $\exp(\kappa z/H)$, but the kinetic energy decays.

Vorticity and Divergence

We examine the solutions of the Laplace Tidal Equations in some enlightening limiting cases. Holton (1992) gives a more extensive analysis, including treatments of the equatorial and mid-latitude β-plane approximations. By means of the Helmholtz Theorem, a general horizontal wind field $\mathbf{V}$ may be partitioned into rotational and divergent components

$$\mathbf{V} = \mathbf{V}_\psi + \mathbf{V}_\chi = \mathbf{k} \times \nabla\psi + \nabla\chi \,.$$

The stream function ψ and velocity potential χ are related to the vorticity and divergence by the Poisson equations $\nabla^2\psi = \zeta$ and $\nabla^2\chi = \delta$. It is straightforward to derive equations for the vorticity and divergence tendencies. Together with the continuity equation, they are

$$\frac{\partial\zeta}{\partial t} + f\delta + \beta v = 0 \tag{10}$$

$$\frac{\partial\delta}{\partial t} - f\zeta + \beta u + \nabla^2 P = 0 \tag{11}$$

$$\frac{\partial P}{\partial t} + gh\delta = 0 \,. \tag{12}$$

These equations are completely equivalent to Eqs. 6–8; no additional approximations have yet been made. However, the vorticity and divergence forms enable us to examine various simple approximate solutions.

Rossby-Haurwitz Modes

If we suppose that the solution is quasi-nondivergent, *i.e.*, we assume $|\delta| \ll |\zeta|$, the wind is given approximately in terms of the stream function $(u, v) \approx (-\psi_y, \psi_x)$, the vorticity equation becomes

$$\nabla^2\psi_t + \beta\psi_x = O(\delta) \,, \tag{13}$$

and we can ignore the right-hand side. Assuming the stream function has the wave-like structure of a spherical harmonic, $Y_n^m(\lambda, \phi) = P_n^m(\sin\phi)\exp(im\lambda)$, we substitute the expression $\psi = \psi_0 Y_n^m(\lambda, \phi)\exp(-i\nu t)$ in the vorticity equation and immediately deduce an expression for the frequency:

$$\nu = \nu_R \equiv -\frac{2\Omega m}{n(n+1)} \,. \tag{14}$$

This is the celebrated dispersion relation for Rossby-Haurwitz waves (Haurwitz, 1940). If we ignore sphericity (the β-plane approximation) and assume harmonic dependence $\psi(x, y, t) = \psi_0 \exp[i(kx + \ell y - \nu t)]$, then (13) has the dispersion relation

$$c = \frac{\nu}{k} = -\frac{\beta}{k^2 + \ell^2} \,,$$

which is the expression for phase-speed found by Rossby (1939). The Rossby or Rossby-Haurwitz waves are, to the first approximation, non-divergent waves which travel westward, the phase speed being greatest for the waves of largest scale. They are of relatively low frequency — Eq. 14 implies that $|\nu| \leq \Omega$ — and the frequency decreases as the spatial scale decreases.

To the same degree of approximation, we may write the divergence equation, Eq. 11, as

$$\nabla^2 P - f\zeta - \beta\psi_y = O(\delta) \,. \tag{15}$$

Ignoring the right-hand side, we get the *linear balance equation*

$$\nabla^2 P = \nabla\cdot f\nabla\psi \,, \tag{16}$$

a diagnostic relationship between the geopotential and the stream function. This also follows immediately from the assumption that the wind is both non-divergent ($\mathbf{V} = \mathbf{k} \times \nabla\psi$) and geostrophic ($f\mathbf{V} = \mathbf{k} \times \nabla P$). If variations of f are ignored, we can assume $P = f\psi$. The wind and pressure are in approximate geostrophic balance for Rossby-Haurwitz waves.

Gravity Wave Modes

If we assume now that the solution is quasi-irrotational, *i.e.* that $|\zeta| \ll |\delta|$, then the wind is given approximately by $(u, v) \approx (\chi_x, \chi_y)$ and the divergence equation becomes

$$\nabla^2 \chi_t + \beta \chi_x + \nabla^2 P = O(\zeta)$$

with the right-hand side negligible. Using the continuity equation to eliminate P, we get

$$\nabla^2 \chi_{tt} + \beta \chi_{xt} - gh\nabla^4 \chi = 0\,.$$

Seeking a solution $\chi = \chi_0 Y_n^m(\lambda, \phi) \exp(-i\nu t)$, we find that

$$\nu^2 + \left(-\frac{2\Omega m}{n(n+1)} \right) \nu - \frac{n(n+1)gh}{a^2} = 0\,. \tag{17}$$

The coefficient of the second term is just the Rossby-Haurwitz frequency ν_R found in Eq. 14 above, so that

$$\nu = \pm\sqrt{\nu_G^2 + (\tfrac{1}{2}\nu_R)^2} - \tfrac{1}{2}\nu_R\,, \qquad \text{where} \qquad \nu_G \equiv \sqrt{\tfrac{n(n+1)gh}{a^2}}\,,$$

Noting that $|\nu_G| \gg |\nu_R|$, it follows that

$$\nu_\pm \approx \pm\nu_G\,,$$

the frequency of pure gravity waves. There are then two solutions, representing waves travelling eastward and westward with equal speeds. The frequency increases approximately linearly with the total wavenumber n.

4 Normal Mode Initialization

The model equations can be written schematically in the form

$$\dot{\mathbf{X}} + i\mathbf{L}\mathbf{X} + \mathcal{N}(\mathbf{X}) = \mathbf{0} \tag{18}$$

with $\mathbf{X}$ the state vector, $\mathbf{L}$ a matrix and N a nonlinear vector function. If $\mathbf{L}$ is diagonalized, the system separates into two subsystems, for the low and high frequency components (LF and HF, respectively):

$$\dot{\mathbf{Y}} + i\mathbf{\Lambda_Y}\mathbf{Y} + \mathcal{N}(\mathbf{Y}, \mathbf{Z}) = \mathbf{0} \tag{19}$$

$$\dot{\mathbf{Z}} + i\mathbf{\Lambda_Z}\mathbf{Z} + \mathcal{N}(\mathbf{Y}, \mathbf{Z}) = \mathbf{0} \tag{20}$$

where $\mathbf{Y}$ and $\mathbf{Z}$ are the coefficients of the LF and HF components of the flow, referred to colloquially as the *slow* and *fast* components respectively, and $\mathbf{\Lambda_Y}$ and $\mathbf{\Lambda_Z}$ are diagonal matrices of eigenfrequencies for the two types of modes.

Let us suppose that the initial fields are separated into slow and fast parts, and that the latter are removed so as to leave only the Rossby waves. It might be hoped that this process of "linear normal mode initialization", imposing the condition

$$\mathbf{Z} = \mathbf{0} \quad \text{at} \quad t = 0$$

would ensure a noise-free forecast. However, the results of the technique are disappointing: the noise is reduced initially, but soon reappears; the forecasting equations are nonlinear, and the slow components interact nonlinearly in such a way as to generate gravity waves. The problem

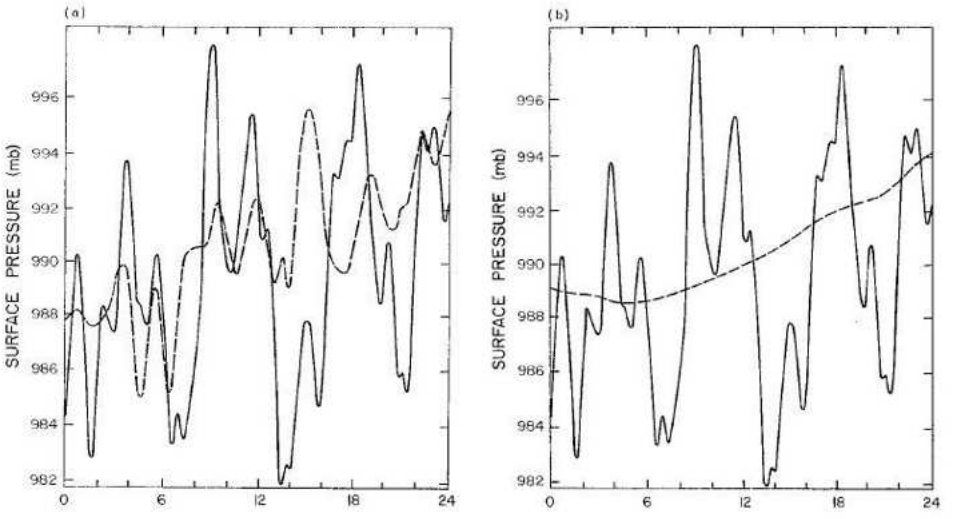

Figure 2: Pressure variation over a 24 hour period for forecasts from uninitialized data (solid lines, both panels), LNMI data (dashed line, left panel) and NNMI data (dashed line, right panel). LNMI = linearly initialized data; NNMI = nonlinearly initialized data. From Williamson and Temperton (1981).

of noise remains: the gravity waves are small to begin with, but they grow rapidly (see Daley, 1991, Ch. 9).

Machenhauer (1977) examined gravity wave dynamics in simple systems and found that the amplitude of the high-frequency components is quasi-stationary. To control the growth of HF components, he proposed setting their initial rate-of-change to zero, in the hope that they would remain small throughout the forecast. Baer (1977) proposed a somewhat more general method, using a two-timing perturbation technique. The forecast, starting from initial fields modified so that $\dot{\mathbf{Z}} = \mathbf{0}$ at $t = 0$ is very smooth and the spurious gravity wave oscillations are almost completely removed. The method takes account of the nonlinear nature of the equations, and is referred to as nonlinear normal mode initialization:

$$\dot{\mathbf{Z}} = \mathbf{0} \quad \text{at} \quad t = 0\,.$$

The method is comprehensively reviewed in Daley (1991).

In Fig. 2, we show the evolution of surface pressure for three 24 hour forecasts (Williamson and Temperton, 1981). The solid lines (in both panels) are the pressure variation for forecasts from uninitialized data. Forecasts from linearly initialized data (LNMI) are shown by the dashed line in the left panel. Forecasts from data that is nonlinearly initialized (NNMI) are shown by the dashed line in the right panel. It is clear that LNMI is ineffective in removing spurious oscillations. NNMI is excellent in this regard.

5 Digital Filter Initialization

Normal mode initialization, or NMI, has been used in many NWP centres, and has performed satisfactorily. Its most natural context is for global models, for which the horizontal structure of the normal modes corresponds to the Hough functions, the eigenmodes of the Laplace Tidal Equations. For limited area models, normal modes can also be derived, but the lateral boundaries force the introduction of simplifying assumptions. An alternative method of initialization, called digital filter initialization (DFI), was introduced by Lynch and Huang (1992). It was generalised

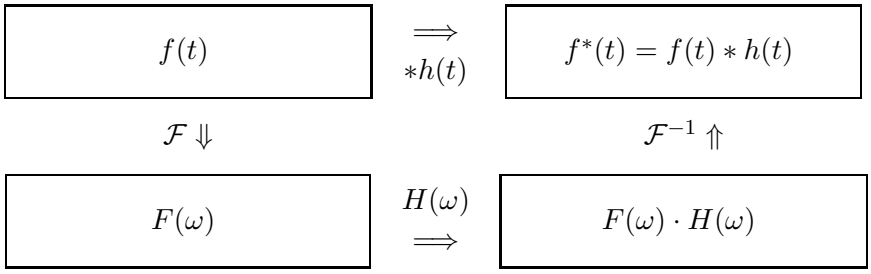

Figure 3: Schematic representation of the equivalence between convolution and filtering in Fourier space.

to allow for diabatic effects by Huang and Lynch (1993). The latter paper also discussed the use of an optimal filter. A much simpler filter, the Dolph-Chebyshev filter, which is a special case of the optimal filter, was applied to the initialization problem by Lynch (1997). A more efficient formulation of DFI was presented by Lynch, Giard and Ivanovici (1997).

Digital filter initialization (DFI) uses filters similar to those arising in signal processing. The selection principle for these is generally based on the frequency of the signal components. There are a number of ideal types — lowpass, highpass, bandpass and bandstop — corresponding to the range of frequencies which pass through the filter and those which are rejected. In many cases the input consists of a low-frequency (LF) signal contaminated by high-frequency (HF) noise, and the information in the signal can be isolated by using a lowpass filter which rejects the noise. Such a situation is typical for the application to meteorology discussed below.

The method of digital filter initialization has significant advantages over alternative methods, and is now in use operationally at several major weather prediction centres (see Numerical Weather Prediction chapter - Swinbank). In DFI there is no need to compute or store normal modes; this advantage becomes more pronounced as the number of degrees of freedom of the model increases. There is no need to separate the vertical modes; NMI requires the introduction of an auxiliary geopotential variable, and partitioning of its changes between the temperature and surface pressure involves an *ad hoc* assumption. DFI is free from this problem. There is complete compatibility with model discretization, eliminating discretization errors due to grid disparities. DFI is applicable to exotic grids on arbitrary domains, facilitating its use with stretched or irregular model grids. There is no iterative numerical procedure which may diverge; therefore, all vertical modes can be initialized effectively. The simplicity of the method makes it easy to implement and maintain. The method is applicable to all prognostic model variables; thus, DFI produces initial fields for these variables which are compatible with the basic dynamical fields. Last but not least, DFI filters the additional prognostic variables in non-hydrostatic models in a manner identical to the basic variables. The DFI method is thus immediately suitable for non-hydrostatic models (Bubnová *et al.*, 1995; Chen and Huang, 2006).

5.1 *Design of Nonrecursive Filters*

Consider a function of time, $f(t)$, with low and high frequency components. To filter out the high frequencies one may proceed as follows:

[1] Calculate the Fourier transform $F(\omega)$ of $f(t)$;
[2] Set the coefficients of the high frequencies to zero;
[3] Calculate the inverse transform.

(See Fig. 3). Step [2] may be performed by multiplying $F(\omega)$ by an appropriate weighting function $H(\omega)$.

Suppose that f is known only at discrete moments $t_n = n\Delta t$, so that the sequence $\{\cdots, f_{-2}, f_{-1}, f_0, f_1, f_2, \cdots\}$ is given. For example, f_n could be the value of some model variable at a particular grid point at time t_n. The shortest period component that can be represented with a time step Δt is $\tau_N = 2\Delta t$, corresponding to a maximum frequency, the so-called Nyquist frequency, $\omega_N = \pi/\Delta t$. The sequence $\{f_n\}$ may be regarded as the Fourier coefficients of a function $F(\theta)$:

$$F(\theta) = \sum_{n=-\infty}^{\infty} f_n e^{-in\theta}, \tag{21}$$

where $\theta = \omega\Delta t$ is the *digital frequency* and $F(\theta)$ is periodic, $F(\theta) = F(\theta + 2\pi)$. High frequency components of the sequence may be eliminated by multiplying $F(\theta)$ by a function $H(\theta)$ defined by

$$H(\theta) = \begin{cases} 1, & |\theta| \le |\theta_c|; \\ 0, & |\theta| > |\theta_c|, \end{cases} \tag{22}$$

where the cutoff frequency $\theta_c = \omega_c \Delta t$ is assumed to fall in the Nyquist range $(-\pi, \pi)$ and $H(\theta)$ has period 2π. This function may be expanded:

$$H(\theta) = \sum_{n=-\infty}^{\infty} h_n e^{-in\theta} \qquad ; \qquad h_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} H(\theta) e^{in\theta} d\theta. \tag{23}$$

The values of the coefficients h_n follow immediately from Eqs. 22 and 23:

$$h_n = \frac{\sin n\theta_c}{n\pi}. \tag{24}$$

Let $\{f_n^\star\}$ denote the low frequency part of $\{f_n\}$, from which all components with frequency greater than θ_c have been removed. Clearly,

$$H(\theta) \cdot F(\theta) = \sum_{n=-\infty}^{\infty} f_n^\star e^{-in\theta}.$$

The convolution theorem for Fourier series now implies that $H(\theta) \cdot F(\theta)$ is the transform of the convolution of $\{h_n\}$ with $\{f_n\}$:

$$f_n^\star = (h * f)_n = \sum_{k=-\infty}^{\infty} h_k f_{n-k}. \tag{25}$$

This enables the filtering to be performed directly on the given sequence $\{f_n\}$. In practice the summation must be truncated at some finite value of k. Thus, an approximation to the low frequency part of $\{f_n\}$ is given by

$$f_n^\star = \sum_{k=-N}^{N} h_k f_{n-k}. \tag{26}$$

A more sophisticated method uses the Chebyshev alternation theorem to obtain a filter whose maximum error in the pass- and stop-bands is minimized. This method yields a filter meeting required specifications with fewer coefficients that the other methods. The design of nonrecursive and recursive filters is outlined in Hamming (1989), where several methods are described, and fuller treatments may be found in Oppenheim and Schafer (1989).

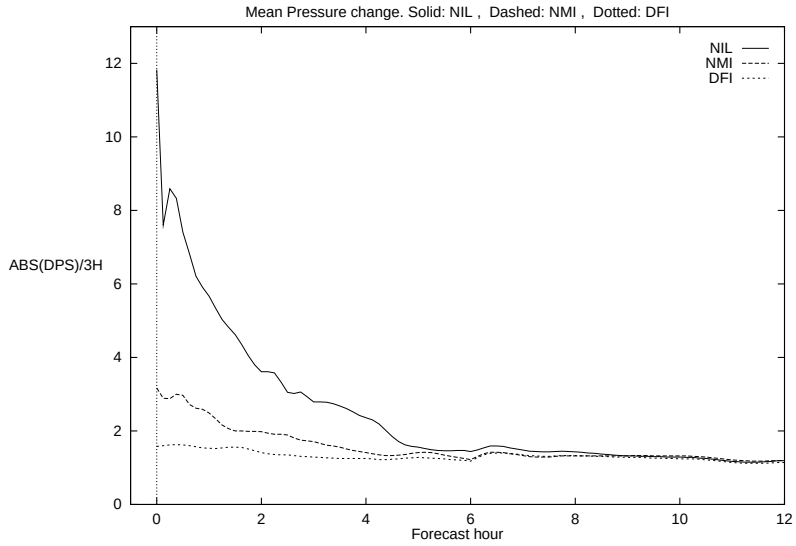

Figure 4: Mean absolute surface pressure tendency for three forecasts. Solid: uninitialized analysis (NIL). Dashed: Normal mode initialization (NMI). Dotted: Digital filter initialization (DFI). Units are hPa/3 hours.

5.2 *Application of a Nonrecursive Digital Filter to Initialization*

An initialization scheme using a nonrecursive digital filter has been developed by Lynch and Huang (1992) for the HIRLAM (High Resolution Limited Area Model) model. The uninitialized fields of surface pressure, temperature, humidity and winds were first integrated forward for three hours, and running sums of the form

$$f_F^\star(0) = \frac{1}{2}h_0 f_0 + \sum_{n=1}^{N} h_{-n} f_n, \tag{27}$$

where $f_n = f(n\Delta t)$, were calculated for each field at each gridpoint and on each model level. These were stored at the end of the three hour forecast. The original fields were then used to make a three hour 'hindcast', during which running sums of the form

$$f_B^\star(0) = \frac{1}{2}h_0 f_0 + \sum_{n=-1}^{-N} h_{-n} f_n \tag{28}$$

were accumulated for each field, and stored as before. The two sums were then combined to form the required summations:

$$f^\star(0) = f_F^\star(0) + f_B^\star(0). \tag{29}$$

These fields correspond to the application of the digital filter Eq. 26 to the original data, and will be referred to as the filtered data.

Complete technical details of the original implementation of DFI in the HIRLAM model may be found in Lynch, *et al.*, (1999). A reformulation of the implementation, with further testing and evaluation, is presented in Huang and Yang (2002).

5.3 *Initialization Example*

A detailed case study based on the implementation in HIRLAM was carried out to check the effect of the initialization on the initial fields and on the forecast, and to examine the

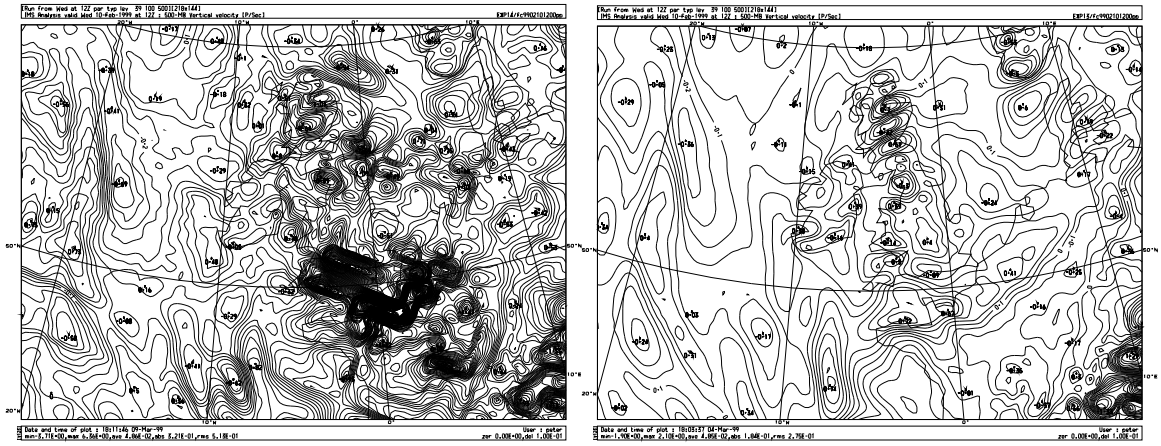

Figure 5: Vertical velocity at 500 hPa over western Europe and the eastern North Atlantic. (Left) Uninitialized analysis (NIL); (Right) after digital filtering (DFI).

efficacy of DFI in eliminating high frequency noise. The digital filter initialization was compared to the reference implicit normal mode initialization (NMI) scheme, and to forecasts with no initialization (NIL). Forecasts starting from the analysis valid at 1200 UTC on 10 February, 1999 were compared.

We first checked the effect of DFI on the analysis and forecast fields. The maximum change in surface pressure due to initialization was 2.2 hPa, with an *rms* change of about 0.5 hPa. The changes to the other analysed variables were in general comparable in size to analysis errors, and considerably smaller in magnitude than typical changes brought about by the analysis itself: the *rms* change in surface pressure from first-guess to analysis was about 1 hPa. The *rms* and maximum differences between the uninitialized 24 hour forecast (NIL) and the filtered forecast (DFI) for all prognostic variables were examined. When we compare these values to the differences at the initial time they were seen to be generally smaller. The changes made by DFI are to the high frequency components; since these are selectively damped during the course of the forecast, the two forecasts were very similar. After 24 hours the maximum difference in surface pressure was less than 1 hPa and the *rms* difference was only 0.1 hPa.

The basic measure of noise is the mean absolute value of the surface pressure tendency

$$N_1 = \left(\frac{1}{\mathrm{N}}\right) \sum_{\mathrm{n}=1}^{\mathrm{N}} \left| \frac{\partial p_s}{\partial t} \right|_{\mathrm{n}} .$$

For well-balanced fields this quantity has a value of about 1 hPa/3h. For uninitialized fields it can be an order of magnitude larger. In Fig. 4 we plot the value of N_1 for three forecasts. The solid line represents the forecast from uninitialized data: we see that the value of N_1 at the beginning of the forecast is about 12 hPa/3h. This large value reflects the lack of an effective multivariate balance in the analysis. It takes about six hours to fall to a reasonable value. The dashed line is for a forecast starting from data initialized using the implicit normal mode method (NMI). The starting value is about 3 hPa/3h, falling to about 1.5 hPa/3h after twelve hours. The final graph (the dotted line) is for the digitally filtered data (DFI). The initial value of N_1 is now about 1.5, and remains more-or-less constant throughout the forecast. It is clear from this measure that DFI is more effective in removing high frequency noise than NMI.

The measure N_1 indicates the noise in the vertically integrated divergence field. However, even when this is small, there may be significant activity in the internal gravity wave modes. To see this, we look at the vertical velocity field at 500 hPa for the NIL and DFI analyses. The left panel in Fig. 5 shows the uninitialized vertical velocity field, zoomed in over western

Europe and the eastern North Atlantic. There is clearly substantial gravity wave noise in this field. In fact, the field is physically quite unrealistic. The right panel shows the DFI vertical velocity. It is much smoother; the spurious features have been eliminated and the large values with small horizontal scales which remain are clearly associated with the Scottish Highlands, the Norwegian Mountains and the Alps. Comparison with the NMI method (see Lynch *et al.* (1999), for details) indicates that DFI is more effective than NMI in dealing with internal gravity wave noise. It is noteworthy that stationary mountain waves are unaffected by digital filtering, since they have zero frequency. This is a desirable characteristic of the DFI scheme.

5.4 *Benefits for the Data Assimilation Cycle*

In Lynch *et al.* (1999), a parallel test of data for one of the FASTEX (Fronts and Atlantic Storm Track EXperiment) intensive observing periods showed that the DFI method resulted in slightly improved scores compared to NMI. As it is not usual for an initialization scheme to yield significant improvements in forecast accuracy, some discussion is merited. We cannot demonstrate beyond question the reason for this improvement. However, the comparative results showed up some definite defects in the implicit normal mode initialization as implemented in the reference HIRLAM model. It was clear that the NMI scheme did not eliminate imbalance at lower model levels. Moreover, although the noise level indicated by the parameter N_1 fell to a reasonable level in six hours, there was still internal gravity wave noise, not measured by this parameter. Any noise in the six hour forecast will be carried through to the next analysis cycle, and will affect the quality control and assimilation of new observational data. It is believed that the DFI scheme, with its superior ability to establish atmospheric balance, results in improved assimilation of data and consequently in a reduction of forecast errors.

6 Constraints in 4DVAR

We conclude with a discussion on the application of a digital filter as a weak constraint in four-dimensional variational assimilation (4DVAR; see chapter on variational data assimilation - Talagrand). The idea is that if the state of the system is noise-free at a particular time, *i.e.*, is close to the slow manifold, it will remain noise-free, since the slow manifold is an invariant subset of phase-space (Leith, 1980). We consider a sequence of values $\{x_0, x_1, x_2, \cdots x_N\}$ and form the filtered value

$$\bar{x} = \sum_{n=0}^{N} h_n x_n. \tag{30}$$

The evolution is constrained, so that the value at the mid-point in time is close to this filtered value, by addition of a term

$$J_c = \tfrac{1}{2}\gamma ||x_{N/2} - \bar{x}||^2$$

to the cost function to be minimized (γ is an adjustable parameter). It is straightforward to derive the adjoint of the filter operator (Gustafsson, 1992). Gauthier and Thépaut (2001) applied such a constraint to the 4DVAR system of Météo-France. They found that a digital filter weak constraint imposed on the low-resolution increments efficiently controlled the emergence of fast oscillations while maintaining a close fit to the observations. As the values required for input to the filter are already available, there is essentially no computational overhead in applying this procedure. The dynamical imbalance was significantly less in 4DVAR than in 3DVAR. Wee and Kuo (2004) included a J_c term in the MM5 4DVAR. They found that the weak contraint not only reduces the dynamic imbalance in the 4DVAR solution, but also improves the quality of the analysis and forecast significantly.

To illustrate the impact of the J_c constraint, experiments are carried out using the WRF (Weather Research and Forecasting) 4DVAR system (Huang *et al.*, 2009), (a) without and (b)

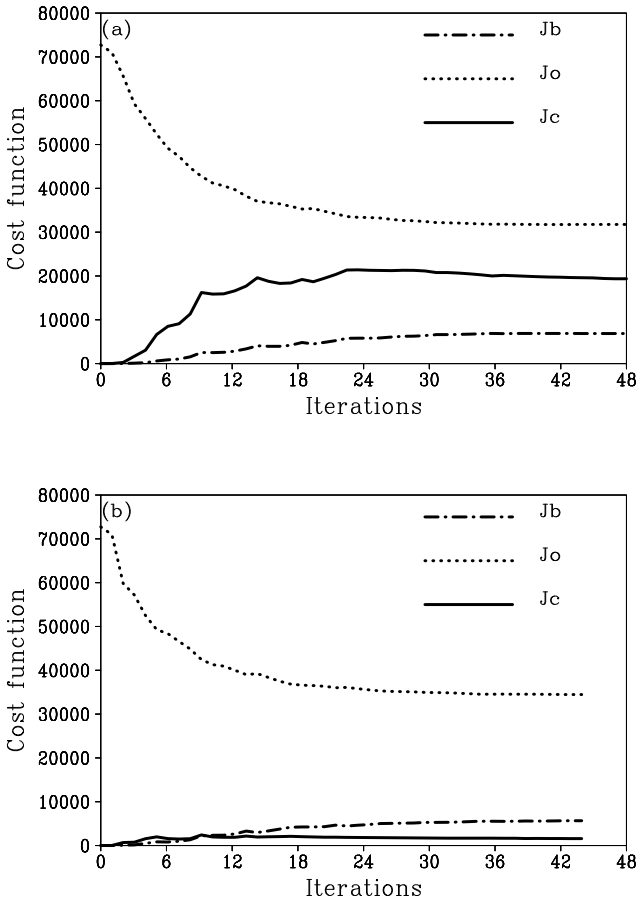

Figure 6: Cost functions for experiment (a) without J_c and (b) with J_c in the mininmization.

with the penalty term in the minimization. The cost functions (J_o, J_b and J_c) are shown in Fig. 6. In both panels, J_c is computed with $\gamma = 0.1$. It is clear that the unconstrained 4DVAR analysis contains a significant amount of noise, with J_c large, and the weak constraint J_c is able to control the noise level. In most of our experiments, J_c also helps the convergence of the miminimization.

To further demonstrate the noise control effect of J_c, we computed N_1 during the forecasts from WRF 4DVAR analyses using different γ. The results from five experiments are shown in Fig. 7. NoJcDF: forecast start from a WRF 4DVAR analysis without J_c or $\gamma = 0$. JcDF(0.1): forecast start from a WRF 4DVAR analysis with J_c and $\gamma = 0.1$. JcDF(1): forecast start from a WRF 4DVAR analysis with J_c and $\gamma = 1$. JcDF(10): forecast start from a WRF 4DVAR analysis with J_c and $\gamma = 10$. FGS: forecast start from firstguess, which is a 3 h forecast from prevous analysis cycle and can be considered as a noise free forecast. The larger the weight we assign to J_c in the 4DVAR minimization, the lower the noise level becomes in the subsequent forecast. However, a larger weight in J_c may compromise the fit to observations (a larger J_o at the end of minimization). The tuning of γ is necessary.

7 Conclusion

We have described several methods of eliminating noise from the forecast by removal of spuriously large-amplitude gravity-wave components from the initial data. This is essential for practical

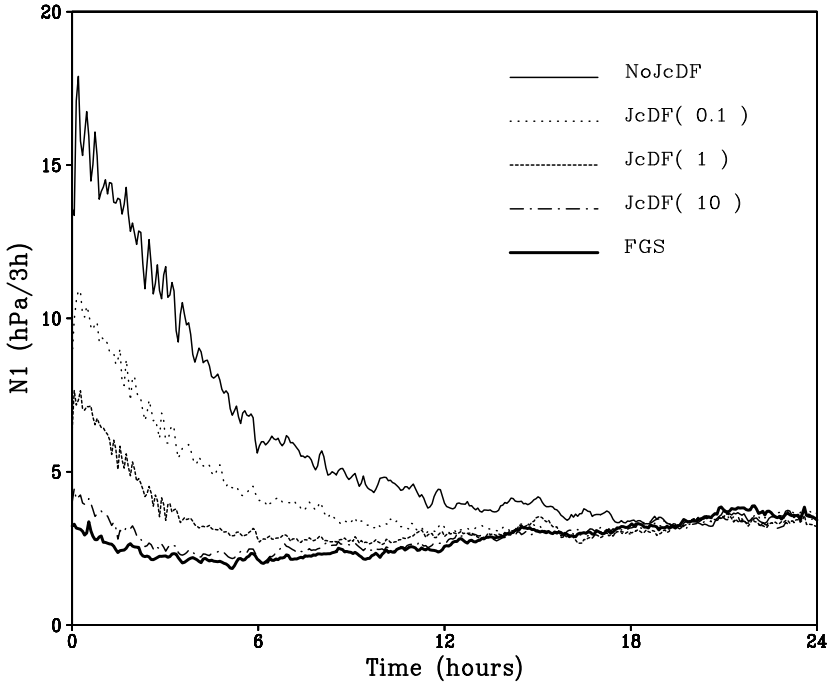

Figure 7: Mean absolute surface pressure tendency for 5 forecasts.

reasons and, in particular, for avoidance of problems in the assimilation cycle. The benefits of initialization are clear. However, it is noteworthy that modern variational assimilation methods are capable of producing fields in good balance, so that a separate initialization stage is less important now. Constraints to ensure good balance can be incorporated directly into variational assimilation schemes. The digital filter method is particularly attractive in this respect, and is a natural choice for variational analysis.

References

Baer, F. 1977. Adjustment of initial conditions required to suppress gravity oscillations in nonlinear flows. *Beitr. Phys. Atmos.*, **50**, 350–366.

Bubnová, R., Hello, G., Bénard, P. and Geleyn, J.-F. 1995. Integration of the fully elastic equations cast in the hydrostatic pressure terrain-following coordinate in the framework of the ARPEGE/Aladin NWP system. *Mon. Wea. Rev.*, **123**, 515–535.

Charney, J.G. 1955. The use of the primitive equations of motion in numerical prediction. *Tellus*, **7**, 22–26.

Charney, J.G., Fjørtoft, R. and von Neumann, J. 1950. Numerical integration of the barotropic vorticity equation. *Tellus*, **2**, 237–254.

Chen, M. and Huang, X.-Y. 2006. Digital filter initialization for MM5. *Mon. Wea. Rev.*, **134**, 1222–1236.

Daley, R. 1991. *Atmospheric Data Assimilation.* Cambridge University Press. 457pp.

Gauthier, P. and Thépaut, J.-N. 2001. Impact of the digital filter as a weak constraint in the pre-operational 4D-Var assimilation system of Météo-France. *Mon. Wea. Rev.*, **129**, 2089–2102.

Gustafsson, N. 1992. Use of a digital filter as weak constraint in variational data assimilation. *Pages 327–338 of: Proceedings of the ECMWF Workshop on Variational Assimilation, with special emphasis on three-dimensional aspects.*

Hamming, R.W. 1989. *Digital Filters.* Prentice-Hall International, Inc. 284pp.

Haurwitz, B. 1940. The motion of atmospheric disturbances on the spherical earth. *J. Marine Res.*, **3**, 254–267.

Hinkelmann, K. 1951. Der Mechanismus des meteorologischen Lärmes. *Tellus*, **3**, 285–296.

Hinkelmann, K. 1959. *Ein numerisches Experiment mit den primitiven Gleichungen.* Vol. Rossby Memorial Volume. Rockerfeller lnstitute Press. Pages 486–500.

Holton, J.R. 1992. *An Introduction to Dynamic Meteorology.* 3rd edn. International Geophysics Series, vol. 48. Academic Press. Chap. 7.

Huang, X.-Y. and Lynch, P. 1993. Diabatic digital filter initialization: Application to the HIRLAM model. *Mon. Wea. Rev.*, **121**, 589–603.

Huang, X.-Y. and Yang, X. 2002. *A new implementation of digital filtering initialization schemes for HIRLAM.* Technical Report 53, 36pp. Available from HIRLAM-5, c/o Per Undén, SMHI, S-60176 Norrköping, Sweden.

Huang, X.-Y., Xiao, Q., Barker, D. M., Zhang, Xin, Michalakes, J., Huang, W., Henderson, T., Bray, J., Chen, Y., Ma, Z., Dudhia, J., Guo, Y., Zhang, Xiaoyan, Won, D.-J., Lin, H.-C. and Kuo, Y.-H. 2009. Four-Dimensional Variational Data Assimilation for WRF: Formulation and Preliminary Results. *Mon. Wea. Rev.*, **137**, in press.

Kasahara, A. 1976. Normal modes of ultralong waves in the atmosphere. *Mon. Wea. Rev.*, **104**, 669–690.

Leith, C.E. 1980. Nonlinear normal mode initialization and quasi-geostrophic theory. *J. Atmos. Sci.*, **37**, 958–968.

Lynch, P. 1997. The Dolph-Chebyshev window: a simple optimal filter. *Mon. Wea. Rev.*, **125**, 655–660.

Lynch, P. 2006. *The Emergence of Numerical Weather Prediction: Richardson's Dream.* Cambridge University Press. 279pp.

Lynch, P. and Huang, X.-Y. 1992. Initialization of the HIRLAM model using a digital filter. *Mon. Wea. Rev.*, **120**, 1019–1034.

Lynch, P., Giard, D. and Ivanovici, V. 1997. Improving the efficiency of a digital filtering scheme. *Mon. Wea. Rev.*, **125**, 1976–1982.

Lynch, P., McGrath, R. and McDonald, A. 1999. *Digital filter initialization for HIRLAM.* HIRLAM Tech. Rep. 42, 22pp. Available from HIRLAM-5, c/o Per Undén, SMHI, S-60176 Norrköping, Sweden.

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

McIntyre, M. E. 2003. *Balanced Flow.* Vol. 1, Encyclopedia of Atmospheric Sciences, Ed. J R Holton, J Pyle and J A Curry. 6 Vols. ISBN 0-12-227090-8. Academic Press.

Miyakoda, K. and Moyer, R.W. 1968. A method for initialization for dynamic weather forecasting. *Tellus*, **20**, 115–128.

Oppenheim, A.V. and Schafer, R.W. 1989. *Discrete-Time Signal Processing.* Prentice-Hall International, Inc. 879pp.

Phillips, N. A. 1960. On the problem of initial data for the primitive equations. *Tellus*, **12**, 121–126.

Phillips, N. A. 1973. *Principles of large scale numerical weather prediction.* Vol. Dynamic Meteorology, Ed. P Morel. D. Reidel. Pages 1–96.

Richardson, L.F. 1922. *Weather Prediction by Numerical Process.* Cambridge University Press. 236 *pp.* Reprinted by Dover Publications, New York, 1965.

Rossby, C.G. 1939. Relations between variations in the intensity of the zonal circulation of the atmosphere and the displacements of the semipermanent centers of action. *J. Marine Res.*, **2**, 38–55.

Sasaki, Y. 1958. An objective method based on the variational method. *J. Met. Soc. Japan*, **36**, 77–88.

Wee, T.-K. and Kuo, Y.-H. 2004. Impact of a digital filter as a weak constraint in MM5 4DVAR: An observing system simulation experiment. *Mon. Wea. Rev.*, **132**, 543–559.

Williamson, D. and Temperton, C. 1981. Normal mode initialization for a multilevel gridpoint model. Part II: Nonlinear aspects. *Mon. Wea. Rev.*, **109**, 745–757.