
The control of gravity waves in data assimilation

27 April 1999

By **Adrian Simmons**

European Centre for Medium-Range Weather Forecasts

The need to control gravity waves in data assimilation and the general concept of initialization are discussed. The basic method of non-linear normal-mode initialization is introduced, and aspects of its implementation in models are presented, covering the vertical and horizontal modal decomposition, implicit initialization and diabatic initialization. The control of gravity waves in the ECMWF variational data assimilation system is described. An introduction to the alternative method of digital filtering is also given.

Table of contents

[1 . Introduction](#)

[2 . Non-linear normal-mode initialization](#)

[2.1 Basic method](#)

[2.2 Vertical decomposition](#)

[2.3 Horizontal decomposition](#)

[2.4 Implicit initialization](#)

[2.5 Diabatic initialization](#)

[3 . Control of gravity waves in the ECMWF variational data assimilation system](#)

[4 . Digital filtering](#)

[4.1 Adiabatic, non-recursive filtering](#)

[4.2 Diabatic, recursive filtering](#)

[4.3 Diabatic, non-recursive filtering](#)

[APPENDIX A . Definition of operators for the ECMWF vertical finite-difference scheme](#)

[APPENDIX B . The Lamb Wave](#)

[APPENDIX C . The non-recursive implementation of the recursive filter](#)

1. INTRODUCTION

The existence of *balance* is fundamental to the dynamics and predictability of the atmosphere. Atmospheric motion is dominated not by rapid fluctuations associated with sound and gravity waves, but rather by the more slowly changing weather systems. Over much of the globe these are close to the familiar state of geostrophic balance in which winds blow parallel to contours of constant geopotential height on an isobaric surface ([Fig. 1](#)).

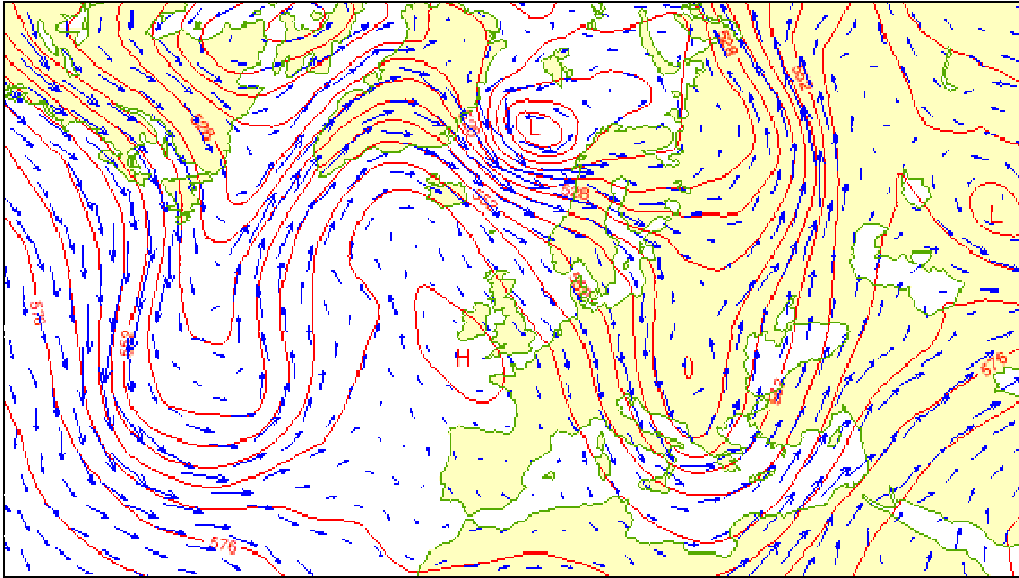


Figure 1. The operational ECMWF analysis of height and wind at 500hPa for 12 UTC 13 April 1997.

Faster wave motions are, however, excited if the atmosphere is suddenly disturbed. The massive volcanic eruption of Krakatau (between Java and Sumatra) in 1883 excited a wave whose propagation could be traced in the surface pressure field around to the antipodal point and back to the source (Fig. 2). The signal took about 35 hours to complete its outward and return trip, moving at about 320 ms^{-1} , around the speed of sound. This was what has come to be known as a Lamb wave, or in numerical weather prediction as an external gravity wave. More commonly, and on a much smaller scale, propagating internal gravity waves may be excited by intense convective systems. Fig. 3 presents observations of wave structure in mid-tropospheric temperature and wind fields over Japan apparently associated with the occurrence of exceptionally heavy rainfall.

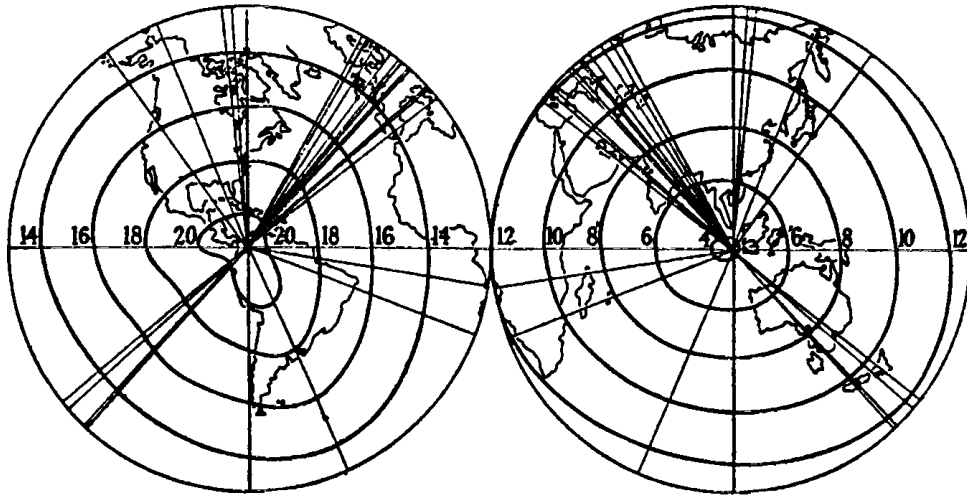


FIG. 1A.

FIG. 1.

Wave No. I—First Passage from Krakatau to the Antipodes.

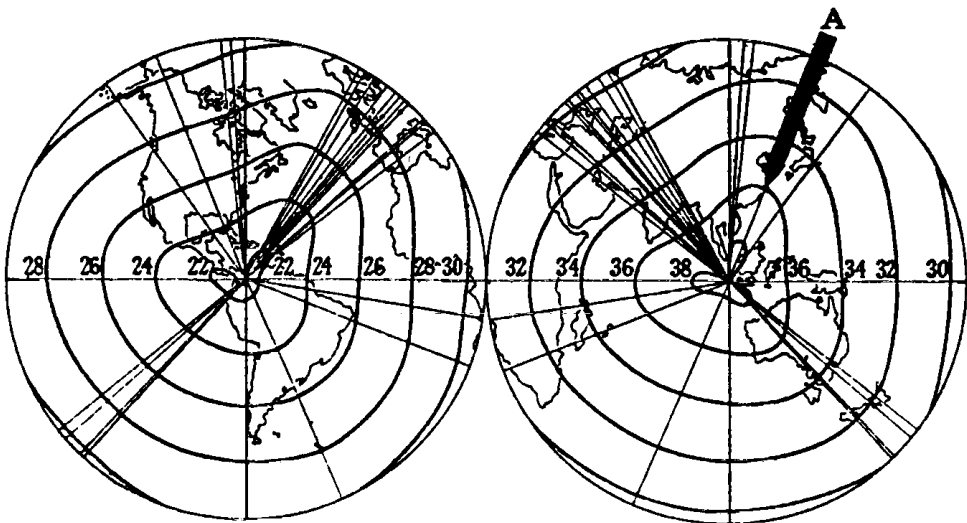


FIG. 2.

FIG. 2A.

Wave No. II—First Passage from the Antipodes back to Krakatau.

Figure 2. Position at two-hourly intervals of the surface-pressure wave excited by the volcanic eruption of Krakatau, from Taylor(1929).

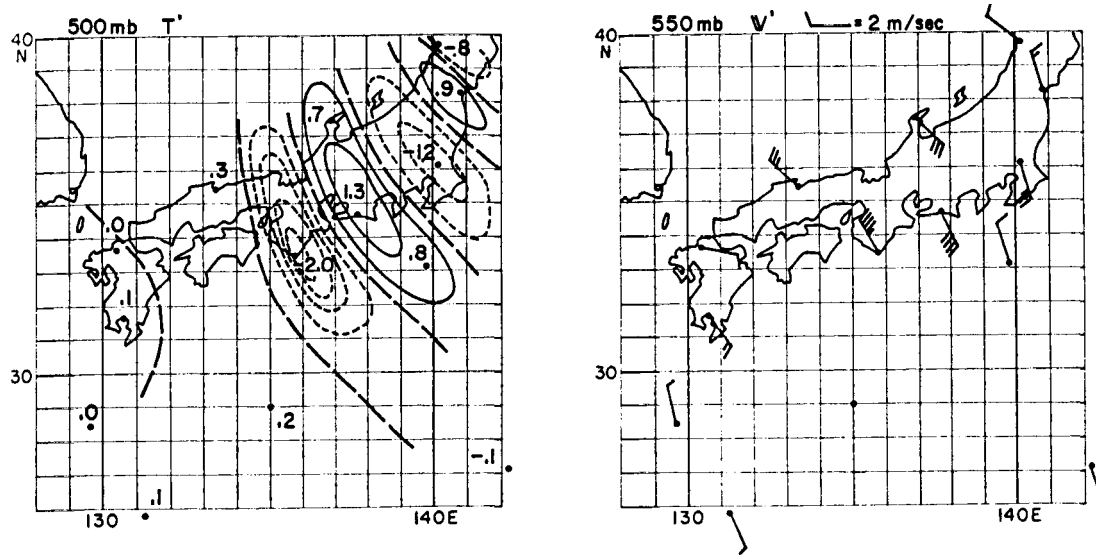


Figure 3. Fine-scale structure in temperature at 500hPa (left) and wind at 550hPa (right) over Japan at 12UTC on 1 July 1971, from Ninomiya (1983)

Numerical weather prediction models mostly assume hydrostatic balance, and thus do not support the existence of sound waves. They are generally based, however, on the primitive equations of motion, and Lamb and internal gravity waves can be excited in these models. In particular, they will be excited in data assimilation if the system uses observations in such a way as to produce analyses in which the balance present in the background state is disturbed. The waves in this case contaminate the short-range forecast that provides the background state for the next analysis, adding to the difficulty of extracting useful information from the observations.

The primary control of gravity waves in a modern data assimilation system is through the multivariate formulation of the analysis. In such a system, an observation of surface pressure or temperature that differs from the background will generally result in an analysis which differs from the background not only in surface pressure and temperature, but also in wind, thus preventing an unrealistic deviation from geostrophic balance. This aspect of the control of gravity waves is not the main concern of these lecture notes, although we shall return to the topic briefly later. Rather, we shall be concerned mostly with what is known as the process of *initialization*.

The term initialization is usually used to refer to the adjustment of analysed initial model conditions to prevent excessive gravity-wave activity in the subsequent forecast. This *dynamical* initialization is needed if the multivariate formulation of the analysis is unable to provide an adequate balance throughout the model domain. This is the type of initialization that will be discussed here. More widely, initialization may be defined as an adjustment of the initial conditions either to maintain dynamical or physical balance or to satisfy some physical constraint. Thus it could include the adjustment of fields to ensure that relative humidities lie in the range from 0 to 100%, or the adjustment of fields to ensure that initial rainfall rates match estimates from observations. This is referred to as *physical* initialization.

Although the principal use of dynamical initialization is to remove imbalance introduced during the analysis step in data assimilation, it has additional applications in numerical weather prediction. It is often used to remove imbalance in an initial state in which the model resolution or orography has been changed, for example in:

- Incremental variational data assimilation (as discussed later);
- Creating initial conditions for ensemble prediction based on a model with resolution lower than that of the model used in the data assimilation;

- Testing of a new higher resolution version of a model or new specification of orography.

Initialization may also be used to remove imbalance when running forecasts from analyses produced by other forecasting centres. This is done occasionally to investigate forecasting failures and is under investigation in the context of ensemble prediction.

Some further characteristics of atmospheric motion must be taken into account in considering the topic of initialization. A rich structure of largely stationary gravity-wave motion may be excited by flow over orography. Fig. 4 shows a classical example. These waves can have an important direct effect on the tropospheric circulation, and together with the convectively-excited waves play a critical role in aspects of the circulation of the stratosphere and mesosphere. Their effects are in general parametrized in large-scale models, but they are coming increasingly to be resolved directly in limited-area and finer-resolution global models, and are a target for short-range mesoscale prediction. They should not be unduly suppressed by initialization.

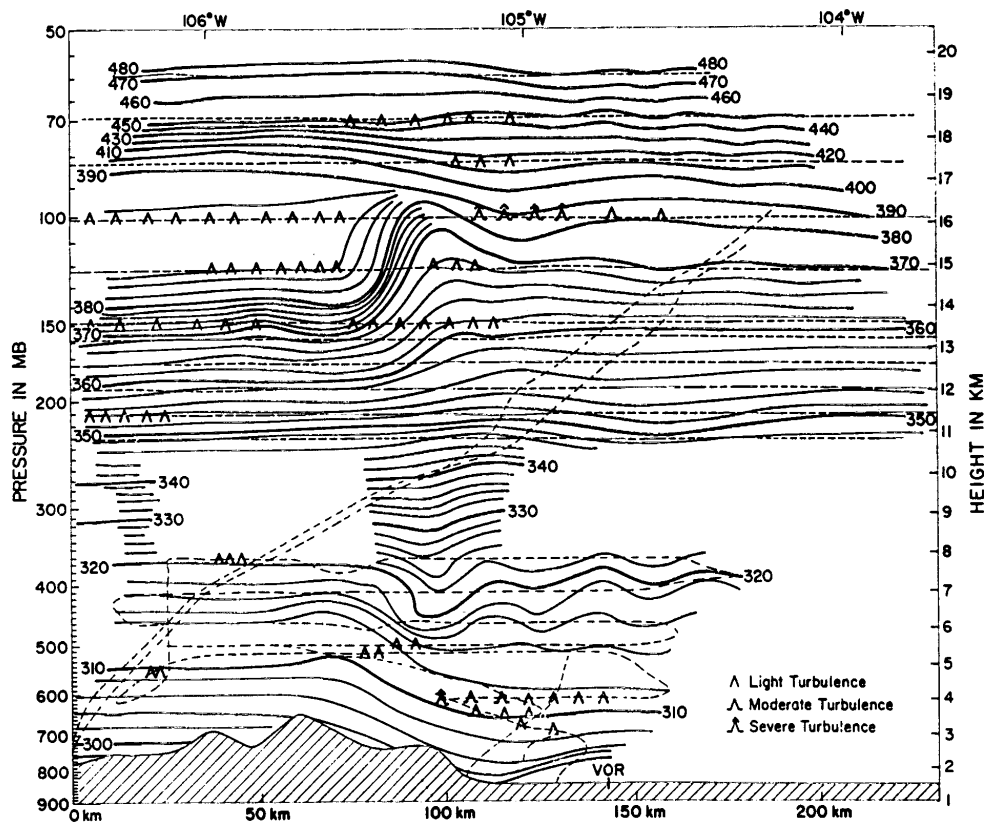


Figure 4. Section of potential temperature crossing the Rocky Mountains in Colorado, for 17 February 1970, from Lilly and Kennedy (1973).

Tidal motion in the atmosphere is a form of forced gravity-wave motion which also should not be suppressed by initialization. The strong semi-diurnal component at the surface is forced by the daily variation in solar heating of the stratosphere, dominating the more strongly-forced diurnal component due to differences in downward propagation. It is seen in quite pronounced oscillations in surface pressure in the tropics and subtropics (Fig. 5 and Fig. 17) and needs to be handled well if other signals in the observations are to be interpreted effectively. Moreover, the quasi-steady tropical circulation systems (as illustrated schematically in Fig. 6) characteristically involve a balance between diabatic heating (or cooling) and the adiabatic cooling (or heating) associated with vertical motion.

Care is needed to ensure that initialization does not produce analyses in which these circulations are weakened.

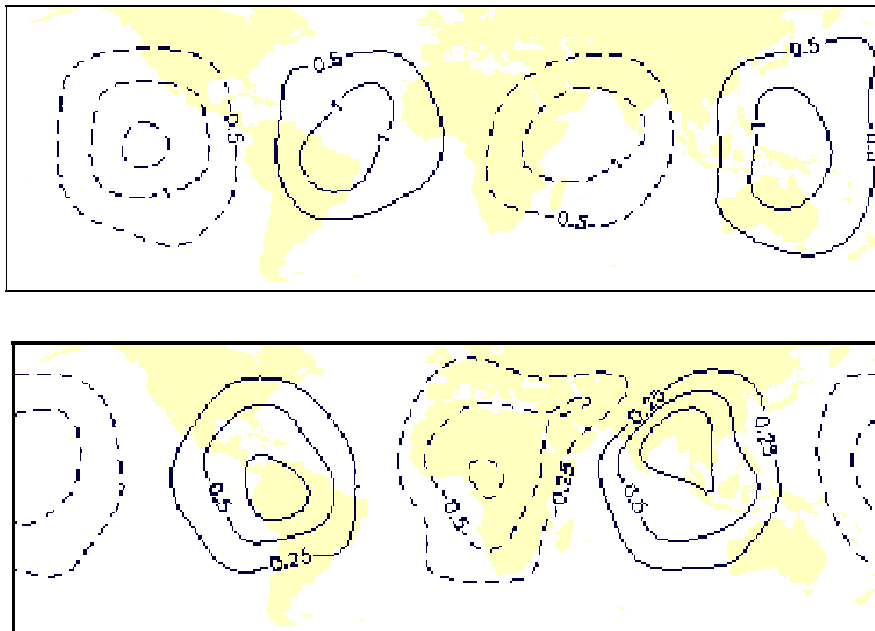


Figure 5. The semi-diurnal tide. The upper plot shows surface pressure (hPa) and the lower plot the tendency of pressure (hPa/hour) at the lowest model level. The quantity plotted is $(1/4)(A_{00} - A_{06} + A_{12} - A_{18})$ where A_{xx} denotes the mean ECMWF analysis at xx UTC for January 1997, truncated spectrally at T10 to remove local orographic and station-specific features. Solid contours denote positive values and dashed contours negative values. The nature of atmospheric tides is described in [Chapman and Lindzen \(1970\)](#).

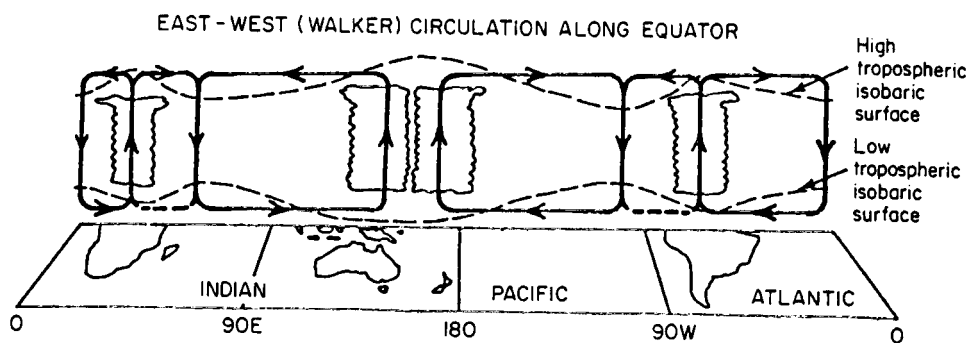


Figure 6. Schematic view of the mean tropospheric circulation in the equatorial plane, from [Webster \(1983\)](#).

The method of non-linear normal-mode initialization that has been used extensively at ECMWF to initialize its global models is described in the following section. The way gravity waves are controlled in the Centre's variational data assimilation system is described and illustrated in section 3. An alternative approach to initialization, that of digital filtering, has been developed in recent years, particularly for application in limited-area models. An introduction to this method is given in section 4.

2. NON-LINEAR NORMAL-MODE INITIALIZATION

2.1 Basic method

Consider a forecast model linearized about a state of rest which is statically stable. The governing equations of the model can be written in the following general form:

$$\frac{d}{dt}(\mathbf{x}) = i\mathbf{L}\mathbf{x} \quad (1)$$

where \mathbf{x} is a column vector of suitably-scaled model variables, of dimension K say, and \mathbf{L} is a symmetric, real $K \times K$ matrix. Explicit examples of sub-elements of \mathbf{L} will be given in the following subsections. The eigenvalues of \mathbf{L} are the frequencies of the small-amplitude wave motions that the resting state can support, and the corresponding eigenfunctions describe the structures of the waves. In limiting cases these waves can be separated into a set of “meteorological” or “Rossby” waves which move relatively slowly westward, and sets of eastward- and westward-moving gravity modes, the graver of which move much faster than the Rossby waves. More generally, the categorization is complicated in the tropics by the existence of the so-called mixed Rossby-gravity and Kelvin waves. Comprehensive studies of the modes of the shallow-water equations (known as the “Hough modes”) have been made by [Languet-Higgins\(1968\)](#) and [Kasahara\(1976\)](#).

The general non-linear model equations can be written in the form:

$$\frac{d}{dt}(\mathbf{x}) = i\mathbf{L}\mathbf{x} + \mathbf{N}(\mathbf{x}) \quad (2)$$

where the term $\mathbf{N}(\mathbf{x})$ represents both the non-linear terms and the residual linear terms that arise from differences between the actual state of rest and that used to compute \mathbf{L} .

Consider a subset comprising J orthonormal eigenvectors, $(\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_J)$, of \mathbf{L} with large eigenvalues $(\nu_1, \nu_2, \dots, \nu_J)$. These represent the gravity waves that are to be initialized. A general model state \mathbf{x} can be written in the form:

$$\mathbf{x} = \mathbf{x}_R + \sum_{j=1}^J \alpha_j \mathbf{g}_j \quad (3)$$

Under idealized conditions in which the J eigenvectors represent a complete set of gravity waves, \mathbf{x}_R represents the component of the atmospheric state comprising solely Rossby waves. More generally, \mathbf{x}_R represents the residual state orthogonal to the set of J gravity modes, which comprises Rossby waves, the slower gravity waves that will not be initialized, and mixed waves. We denote the component that will be initialized by \mathbf{x}_G :

$$\mathbf{x}_G = \sum_{j=1}^J \alpha_j \mathbf{g}_j \quad (4)$$

The α_j are determined by taking the scalar products of the two sides of (3) with the \mathbf{g}_j , and using orthogonality, $\mathbf{x}_R \cdot \mathbf{g}_j = 0$, and orthonormality, $\mathbf{g}_i \cdot \mathbf{g}_j = \delta_{ij}$, to obtain

$$\alpha_j = \mathbf{g}_j \cdot \mathbf{x} \quad (5)$$

or

$$\mathbf{x}_G = \mathbf{E}_G \mathbf{E}_G^T \mathbf{x} \quad (6)$$

where \mathbf{E}_G is the $K \times J$ matrix whose columns are the \mathbf{g}_j , for $j = 1, 2, \dots, J$.

Substituting (3) into (2) and projecting onto the J modes to be initialized gives:

$$\frac{d}{dt}(\alpha_j) = i\nu_j \alpha_j + N_j(\mathbf{x}_R, \alpha_1, \alpha_2, \dots, \alpha_J) \quad (7)$$

where $N_j = (\mathbf{E}_G^T \mathbf{N})_j$.

2.1 (a) *Linear normal-mode initialization*: Linear normal-mode initialization comprises simply setting to zero the amplitudes α_j of the J modes to be initialized. It thus involves replacing the uninitialized state \mathbf{x}_U by \mathbf{x}_I , where

$$\mathbf{x}_I = (\mathbf{I} - \mathbf{E}_G \mathbf{E}_G^T) \mathbf{x}_U \quad (8)$$

and \mathbf{I} is the identity matrix.

Fig. 7 presents an example of the impact of linear normal-mode initialization on the evolution of surface pressure at two points, one in the Great Plains of North America and one in the Himalayas. It is taken from the first study of normal-mode initialization for the original ECMWF finite-difference model (Temperton and Williamson, 1981; Williamson and Temperton, 1981), and presents an extreme case in that the initial data had been taken from an analysis produced using a different model. Imbalance in the interpolated initial conditions for the ECMWF model resulted in large-amplitude oscillations in the surface pressure, as indicated by the solid lines in the figure. The amplitude of these oscillations was clearly reduced by applying linear normal-mode initialization (dashed lines), but significant fluctuations remained. Also worthy of note is the large, rapid fall in surface pressure at the Himalayan point at the start of the integration from the initialized analysis.

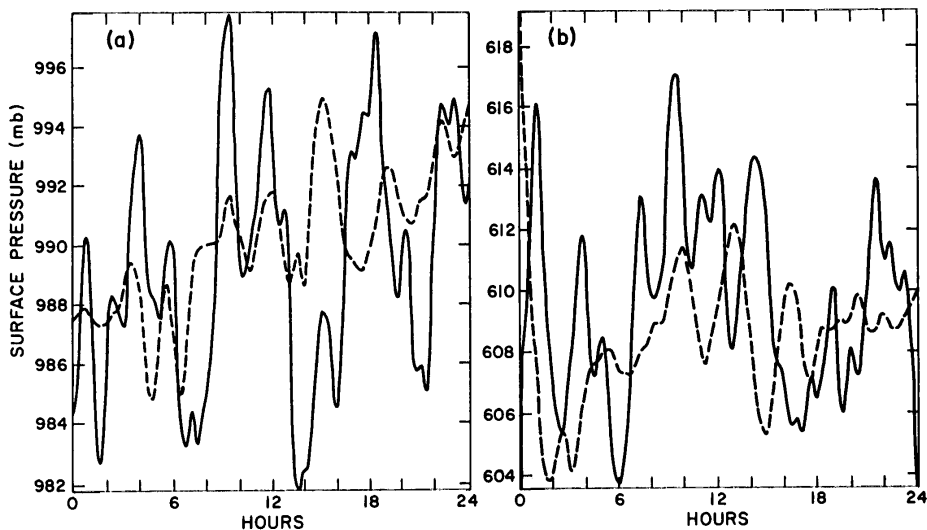


Figure 7. Surface pressure (hPa) as a function of time before (solid) and after (dashed) linear normal-mode initialization at $40^\circ\text{N } 90^\circ\text{W}$ (left) and $30^\circ\text{N } 90^\circ\text{E}$ (right), from Temperton and Williamson(1981).

2.1 (b) *Non-linear normal-mode initialization:* Linear normal-mode does not prevent gravity waves from growing to significant amplitude because the non-linear or residual linear terms represented by $\mathbf{N}(\mathbf{x})$ are not negligible. In *non-linear* normal-mode initialization (Machenhauer, 1977), the initialized gravity-wave amplitudes α_j are set to non-zero values which are chosen to make the gravity-wave tendencies small. An approximate solution of the equation

$$\frac{d}{dt}(\alpha_j) = 0 \quad (9)$$

is sought. Equation (7) is used to define an iterative solution of (9):

$$\alpha_j^{(n)} = -\left(\frac{1}{i\nu_j}\right)N_j(\mathbf{x}_R, \alpha_1^{(n-1)}, \alpha_2^{(n-1)}, \dots, \alpha_J^{(n-1)}). \quad (10)$$

Two iterations of the procedure are usually found to be sufficient. The starting values could be taken from a linear normal-mode initialization ($\alpha_j^{(0)} = 0$), but in practice it is found to be sufficient to start from the uninitialized analysis ($\alpha_j^{(0)} = \mathbf{g}_j \cdot \mathbf{x}_U$). This is the most straightforward starting point when the procedure is implemented in terms of updates applied directly to model variables, as outlined below in equations (12) to (15). It is, moreover, the necessary starting point for an effective full-field diabatic initialization using the approach described in 2.5 (a).

An example of the working of non-linear normal-mode initialization is given in Fig. 8. This is for the same case as presented in Fig. 7 for linear normal-mode initialization. Fig. 8 also contains information related to the vertical decomposition of normal modes which is discussed in the following subsection. For the moment, attention should be directed first towards the upper two panels and the curves labelled “0” which show the forecast from uninitialized initial conditions and “5” which show the forecast from the conditions most fully initialized by the non-linear normal-mode technique. These two panels can be compared directly with Fig. 7 which shows the forecast from linear normal-mode initialization. It is clear that the non-linear technique is much more successful than the linear technique. It not only prevents high-frequency oscillations in surface pressure over the 24-hour integration period, but also avoids the large change to the initial surface pressure at the Himalayan point, a change that is rejected by the model in the first few time steps of the forecast from linear initialization. Comparison of the lower-right and upper-left panels of Fig. 8 shows how the first iteration is sufficient to remove most of the short-period oscillations present when running from the uninitialized analysis, with the second iteration removing most of what is left after the first iteration.

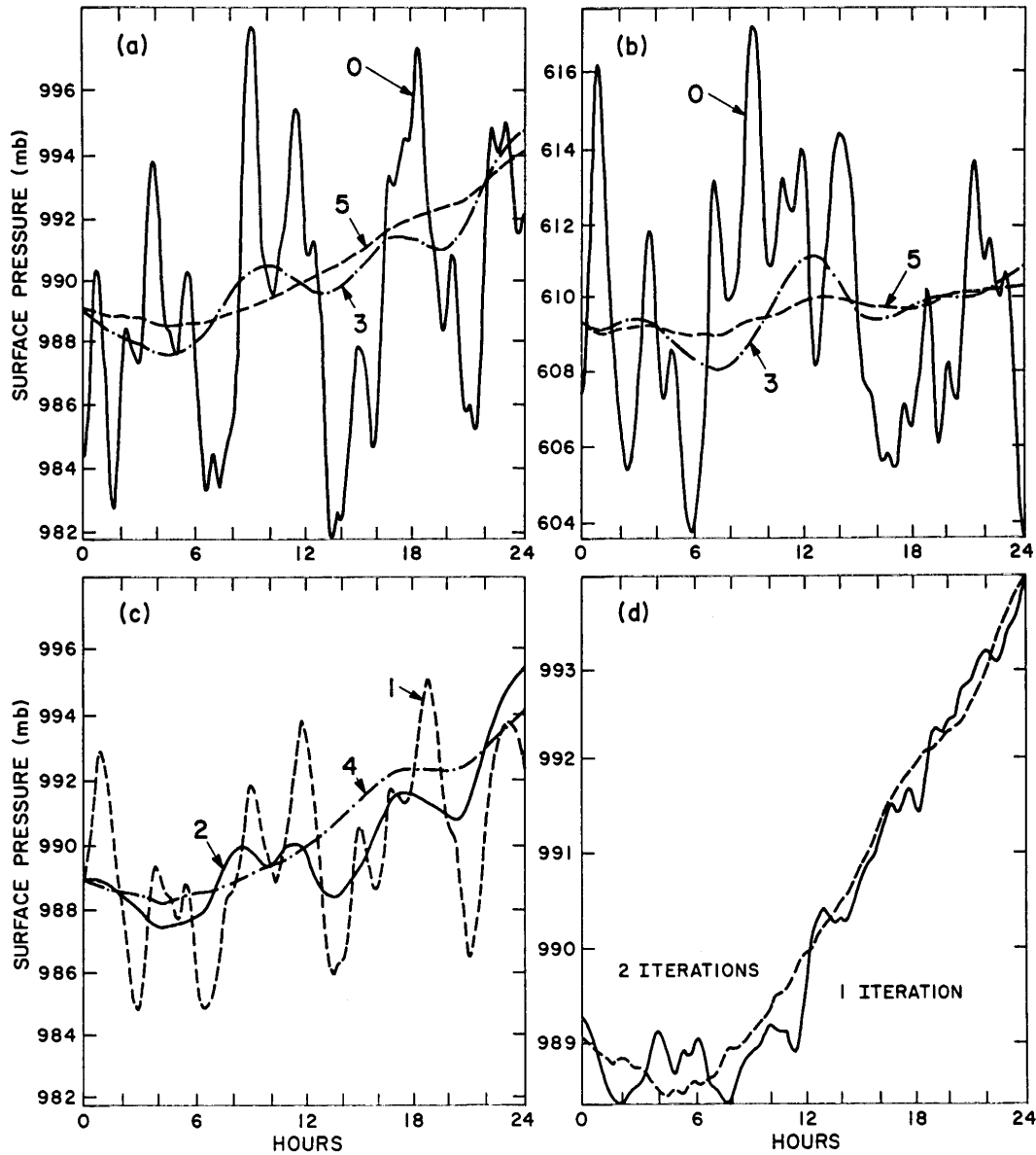


Figure 8. Surface pressure (hPa) as a function of time at $40^{\circ}\text{N } 90^{\circ}\text{W}$ (upper left) and $30^{\circ}\text{N } 90^{\circ}\text{E}$ (upper right) with no initialization (solid) and after non-linear normal-mode initialization of the first three and five vertical modes with two iterations. Surface pressure is also shown at $40^{\circ}\text{N } 90^{\circ}\text{W}$ with one, two and four modes initialized with two iterations (lower left), and with five modes initialized with one and two iterations (lower right). From Williamson and Temperton(1981).

Equation (10) is not normally solved directly. (7) and (10) are combined to express the iterative non-linear normal-mode initialization procedure in terms of the modifications made to the gravity-wave amplitudes at each iteration:

$$(\Delta\alpha_j)^{(n+1)} = \alpha_j^{(n+1)} - \alpha_j^{(n)} = -\left(\frac{1}{i\nu_j}\right)\left(\frac{d\alpha_j}{dt}\right)^{(n)} \quad (11)$$

Equation (11) is then converted into the initialization increment applied to model variables. Starting the iteration



from the uninitialized analysis, the non-linear initialization procedure becomes:

$$\mathbf{x}^{(0)} = \mathbf{x}_U \quad (12)$$

$$\Delta \mathbf{x}^{(n+1)} = i \mathbf{E}_G \Lambda_G^{-1} \mathbf{E}_G^T \left(\frac{d\mathbf{x}}{dt}(\mathbf{x}) \right)^{(n)} \quad (13)$$

where Λ_G is the $J \times J$ diagonal matrix of eigenvalues ν_j for $j = 1, 2, \dots, J$. The non-linear model tendency $(d\mathbf{x}/dt)^{(n)}$ is computed using the latest approximation, $\mathbf{x}^{(n)}$, to the initialized state:

$$\mathbf{x}^{(n)} = \mathbf{x}^{(0)} + \sum_{i=1}^n \Delta \mathbf{x}^{(i)} \quad (14)$$

The final initialized state, \mathbf{x}_I , is given by:

$$\mathbf{x}_I = \mathbf{x}^{(0)} + \sum_{n=1}^N \Delta \mathbf{x}^{(n)} \quad (15)$$

where N is the number of iterations.

This “explicit” form of non-linear normal-mode initialization thus involves a repeated set of operations comprising:

- One timestep of the model to compute the tendencies in physical space, starting from the latest approximation to the linearized state;
- A projection of these tendencies onto the gravity modes to be initialized;
- Division of the gravity-mode tendencies by the frequencies of the modes;
- Projection back to physical space;
- Addition of the increment to form a new approximation to the initialized state.

Equation (13) may be written in the alternative form:

$$\mathbf{L} \Delta \mathbf{x}^{(n+1)} = i \mathbf{E}_G \mathbf{E}_G^T \left(\frac{d\mathbf{x}}{dt} \right)^{(n)} = i \left(\frac{d\mathbf{x}}{dt} \right)_G^{(n)} \quad (16)$$

Equation (16) is used as the starting point for the derivation of “implicit” non-linear normal-mode initialization. As discussed in subsection 2.4, this approach uses a modified form of the governing, linearized equations which enables increments to be calculated in physical space without use of explicit projections to and from gravity-wave space.

2.2 Vertical decomposition

Normal-mode initialization is a practical method because it is possible to separate the vertical and horizontal dependence of the gravity-wave modes. Furthermore, it has to be applied only to a limited number of modes in the vertical.

We consider small-amplitude perturbations about a state of rest with temperature T_r and surface pressure p_{sr} , with eastward and northward wind perturbations u' and v' , and temperature perturbation T' . The equations are set out in the form appropriate for a model (such as that of ECMWF) in which the logarithm of surface pressure is a basic prognostic variable, with perturbation given by $(\ln p_s)'$. We restrict the presentation here to the case in

which the reference temperature T_r is isothermal, rather than varying with pressure. This simplifies the equations, and is not a serious restriction in practice. Effects of unrepresentivity of the reference state of rest are taken into account in non-linear normal-mode initialization through the term $N(\mathbf{x})$. Moreover, gravity-wave phase speeds and vertical mode structures are not strongly dependent on the choice of reference state, as illustrated in [Appendix A](#).

The dry, linearized, primitive equations for a terrain-following vertical coordinate are then:

$$\frac{\partial u'}{\partial t} = 2\Omega \sin\theta v' - \frac{1}{a \cos\theta} \frac{\partial}{\partial \lambda} (\phi' + RT_r (\ln p_s)') \quad (17)$$

$$\frac{\partial v'}{\partial t} = -2\Omega \sin\theta u' - \frac{1}{a} \frac{\partial}{\partial \theta} (\phi' + RT_r (\ln p_s)') \quad (18)$$

$$\frac{\partial T'}{\partial t} = \frac{\kappa T_r}{p} \omega' \quad (19)$$

$$\frac{\partial}{\partial t} (\ln p_s)' = -\frac{1}{p_{sr}} \int_0^{p_{sr}} D' dp \quad (20)$$

Here D' is the divergence:

$$D' = \frac{1}{a \cos\theta} \left(\frac{\partial u'}{\partial \lambda} + \frac{\partial}{\partial \theta} (v' \cos\theta) \right) \quad (21)$$

ω' is the pressure-coordinate vertical velocity:

$$\omega' = -\int_0^p D' dp \quad (22)$$

and ϕ' is the temperature-dependent part of the perturbation geopotential:

$$\phi' = \int_p^{p_{sr}} \frac{RT'}{p} dp \quad (23)$$

Ω is the planetary rotation rate, a is the planetary radius, λ is longitude, θ is latitude and R is the gas constant of dry air. $\kappa = R/c_p$, where c_p is the specific heat of dry air at constant pressure. In evaluating the pressure on a coordinate surface in equations (19), (20), (22) and (23), the surface pressure is taken to be the reference value p_{sr} .

We consider a model with L levels, define \mathbf{u} to be the column vector comprising the values of u' at the L levels, and define \mathbf{v} , \mathbf{D} and \mathbf{T} similarly. The column vector with each element equal to RT_r is denoted by $\boldsymbol{\mu}$, and we use matrices $\boldsymbol{\gamma}$ and $\boldsymbol{\tau}$, and vector $\boldsymbol{\delta}$ to represent a general finite-difference scheme:

$$\frac{1}{p_{sr}} \int_0^{p_{sr}} D dp \rightarrow \boldsymbol{\delta}^T \mathbf{D}$$

$$\frac{\kappa T_r}{p} \omega \rightarrow -\tau \mathbf{D}$$

$$\phi' \rightarrow \gamma \mathbf{T}$$

$$(\mathbf{R}T_r)(\ln p_s)' \rightarrow \mu(\ln p_s)'$$

Explicit expressions for γ , τ and δ are given in [Appendix A](#) for the vertical discretization scheme developed for the ECMWF model by [Simmons and Burridge\(1981\)](#).

We further define:

$$\mathbf{P} = \gamma \mathbf{T} + \mu(\ln p_s)' \quad (24)$$

Then, using (17), (18) and the combination of (19) and (20), the linearized equations become:

$$\frac{\partial \mathbf{u}}{\partial t} = 2\Omega \sin \theta \mathbf{v} - \frac{1}{a \cos \theta} \frac{\partial \mathbf{P}}{\partial \lambda} \quad (25)$$

$$\frac{\partial \mathbf{v}}{\partial t} = -2\Omega \sin \theta \mathbf{u} - \frac{1}{a} \frac{\partial \mathbf{P}}{\partial \theta} \quad (26)$$

$$\frac{\partial \mathbf{P}}{\partial t} = -\mathbf{G} \mathbf{D} \quad (27)$$

where

$$\mathbf{G} = \gamma \tau + \mu \delta^T \quad (28)$$

The normal-mode initialization procedure determines an increment to the “mass” variable \mathbf{P} . To obtain increments in temperature and the logarithm of surface pressure, equation (27) is used together with the vector forms of the equations for \mathbf{T} and $(\ln p_s)'$:

$$\frac{\partial \mathbf{T}}{\partial t} = -\tau \mathbf{D} \quad (29)$$

$$\frac{\partial}{\partial t}(\ln p_s)' = -\delta^T \mathbf{D} \quad (30)$$

Given initialization increment $\Delta \mathbf{P}$, the corresponding increments to temperature and the logarithm of surface pressure, $\Delta \mathbf{T}$ and $\Delta(\ln p_s)'$, are given by

$$\Delta \mathbf{T} = \tau \mathbf{G}^{-1} \Delta \mathbf{P} \quad (31)$$

and

$$\Delta(\ln p_s)' = \delta^T \mathbf{G}^{-1} \Delta \mathbf{P} \quad (32)$$

Now let Φ denote the diagonal matrix of eigenvalues Φ_l of \mathbf{G} , for $l = 1, 2, \dots, L$, and let \mathbf{N} denote the matrix whose columns are the eigenvectors of \mathbf{G} . \mathbf{G} can be written: $\mathbf{G} = \mathbf{N}\Phi\mathbf{N}^{-1}$.

If the prognostic variables are transformed in the following way:

$$\hat{u}_l = (\mathbf{N}^{-1}\mathbf{u})_l$$

$$\hat{v}_l = (\mathbf{N}^{-1}\mathbf{v})_l$$

and

$$\hat{P}_l = (\mathbf{N}^{-1}\mathbf{P})_l$$

equations (25), (26) and (27) become a set of L uncoupled shallow water equations with “equivalent” depths Φ_l/g :

$$\frac{\partial \hat{u}_l}{\partial t} = 2\Omega \sin \theta \hat{v}_l - \frac{1}{a \cos \theta} \frac{\partial \hat{P}_l}{\partial \lambda} \quad (33)$$

$$\frac{\partial \hat{v}_l}{\partial t} = -2\Omega \sin \theta \hat{u}_l - \frac{1}{a} \frac{\partial \hat{P}_l}{\partial \theta} \quad (34)$$

$$\frac{\partial \hat{P}_l}{\partial t} = -\Phi_l \hat{D}_l \quad (35)$$

Here \hat{D}_l is the divergence associated with velocities \hat{u}_l and \hat{v}_l as defined by (21):

$$\hat{D}_l = \frac{1}{a \cos \theta} \left(\frac{\partial \hat{u}_l}{\partial \lambda} + \frac{\partial}{\partial \theta} (\hat{v}_l \cos \theta) \right) \quad (36)$$

The governing equation for pure gravity waves is obtained by setting the rotation rate, Ω , to zero, taking the time derivative of (35), and using (33) and (34) to eliminate the rate of change of divergence \hat{D}_l . This gives:

$$\frac{\partial^2 \hat{P}_l}{\partial t^2} - \Phi_l \nabla^2 \hat{P}_l = 0 \quad (37)$$

For scales small enough for a local plane geometry to be valid, the phase speed of the pure gravity waves is $\sqrt{\Phi_l}$. On the sphere the pure gravity-wave modes are the spherical harmonic functions $P_n^m(\sin \theta)e^{im\lambda}$, where P_n^m is an associated Legendre function. The spherical harmonic functions are eigenfunctions of the Laplacian operator with eigenvalues $-n(n+1)/a^2$. The pure gravity-wave frequencies are thus $\frac{1}{a} \sqrt{n(n+1)\Phi_l}$.

Examples of modes are presented here for the current operational 31-level ECMWF model and for a 50-level version of the model planned for operational implementation. The location of the model levels for these two vertical resolutions is shown in Fig. 9. Table 1 shows phase speeds and equivalent depths of the faster modes for a reference temperature, T_r , of 300K and a reference surface pressure, p_{sr} , of 1000hPa. Corresponding vertical structures of the divergence fields (weighted by \sqrt{p}) are shown in Fig. 10.

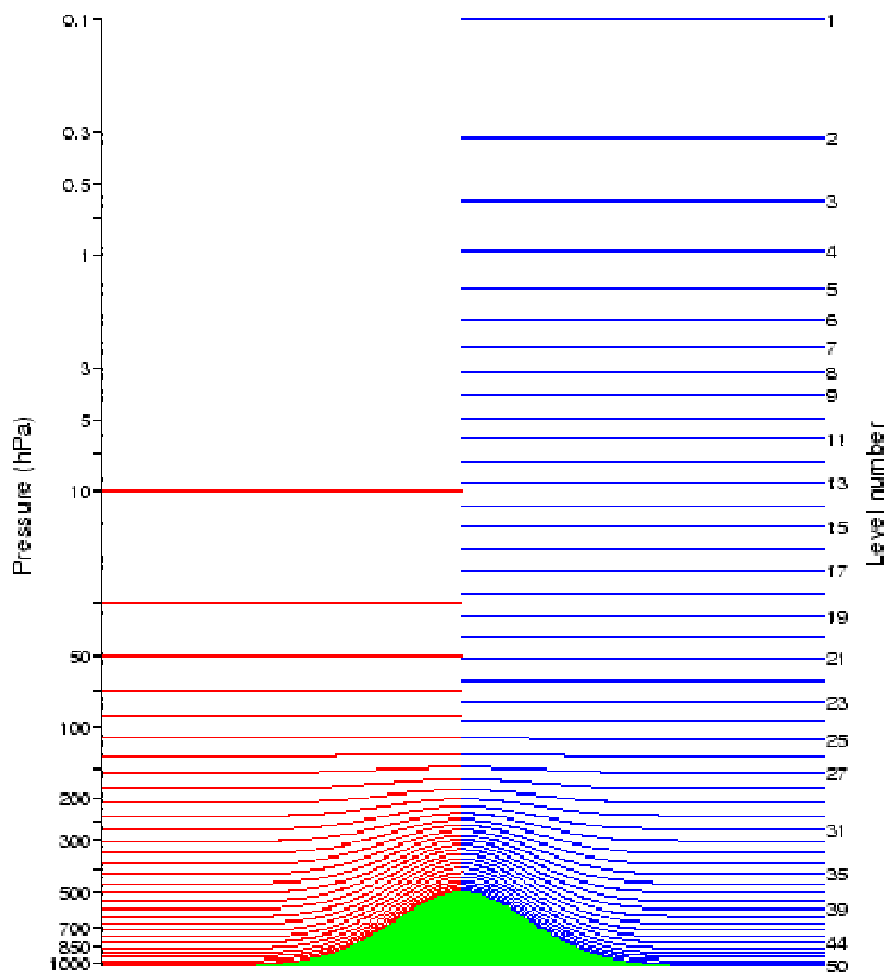


Figure 9. The distribution of full model levels for 31-level (left) and 50-level (right) vertical resolutions, plotted for a distribution of surface pressure which varies from 1013.25 to 500 hPa.

TABLE 1. GRAVITY-WAVE PHASE SPEEDS AND EQUIVALENT DEPTHS FOR THE VERTICAL MODES OF THE 31-LEVEL AND 50-LEVEL VERSIONS OF THE ECMWF MODEL FOR A REFERENCE TEMPERATURE OF 300 K AND A REFERENCE SURFACE PRESSURE OF 1000 hPa. ONLY MODES WITH GRAVITY-WAVE PHASE SPEEDS FASTER THAN 50 ms⁻¹ ARE INCLUDED.

Mode number l	Phase speed $\sqrt{\Phi_l}$ (ms ⁻¹) for 31-level model	Phase speed $\sqrt{\Phi_l}$ (ms ⁻¹) for 50-level model	Equivalent depth Φ_l/g (km) for 31-level model	Equivalent depth Φ_l/g (km) for 50-level model
1	343	347	12.0	12.3
2	203	262	4.2	7.0
3	119	194	1.4	3.8
4	78	145	0.6	2.1
5	56	112	0.3	1.3

TABLE 1. GRAVITY-WAVE PHASE SPEEDS AND EQUIVALENT DEPTHS FOR THE VERTICAL MODES OF THE 31-LEVEL AND 50-LEVEL VERSIONS OF THE ECMWF MODEL FOR A REFERENCE TEMPERATURE OF 300 K AND A REFERENCE SURFACE PRESSURE OF 1000 hPa. ONLY MODES WITH GRAVITY-WAVE PHASE SPEEDS FASTER THAN 50 ms⁻¹ ARE INCLUDED.

Mode number l	Phase speed $\sqrt{\Phi_l}$ (ms ⁻¹) for 31-level model	Phase speed $\sqrt{\Phi_l}$ (ms ⁻¹) for 50-level model	Equivalent depth Φ_l/g (km) for 31-level model	Equivalent depth Φ_l/g (km) for 50-level model
6		90		0.8
7		75		0.6
8		63		0.4
9		54		0.3

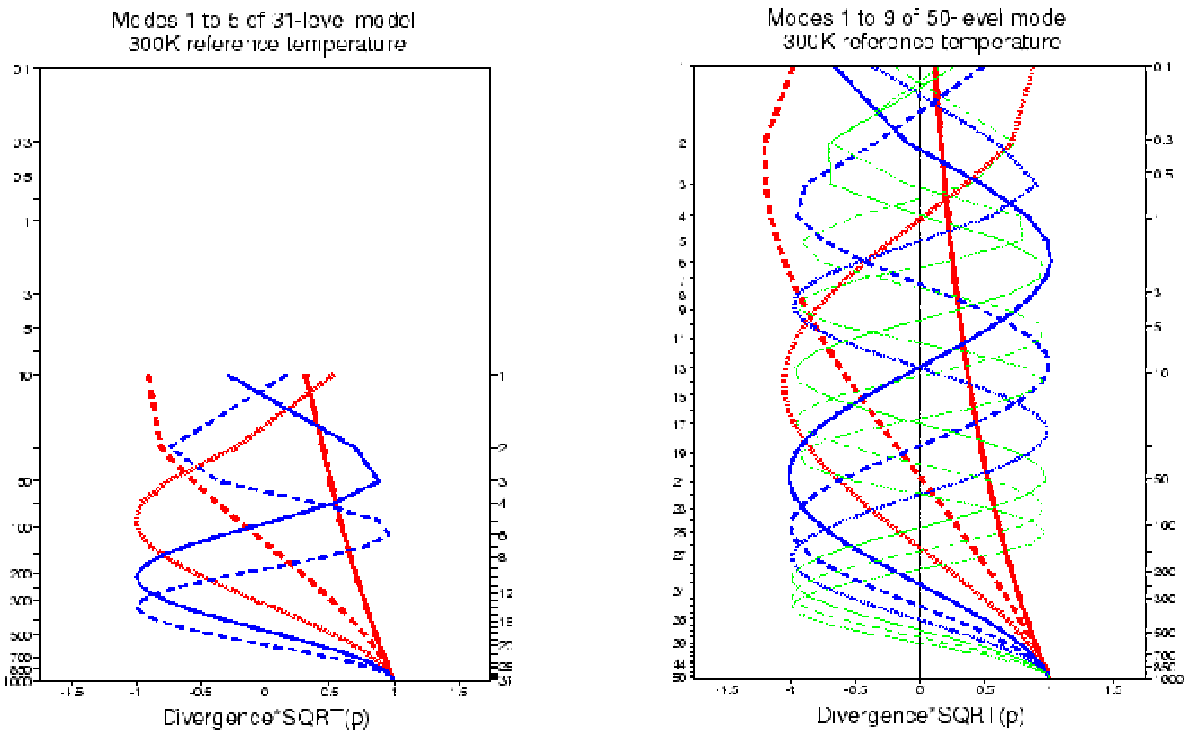


Figure 10. Structures of the vertical modes of the 31-level and 50-level models with gravity-wave phase speeds faster than 50 ms⁻¹, for a reference temperature of 300K and a reference surface pressure of 1000 hPa.

The gravest mode moves the fastest, and differs in structure from the other modes in that it is the only mode whose \sqrt{p} -weighted amplitude (or energy density) decreases with height. It is generally referred to as the external gravity wave in the context of numerical weather prediction, but is known in geophysical fluid dynamics as the Lamb wave. It is shown in [Appendix B](#) that there is a corresponding analytical solution of the continuous equations in which the divergence and temperature vary in the vertical as $p^{-\kappa}$, and the phase speed is given by $\sqrt{(RT_r)/(1-\kappa)}$. This gives a phase speed of 347ms⁻¹ for a reference temperature of 300K, with $R=287\text{m}^2\text{K}^{-1}\text{s}^{-2}$ and $\kappa=0.286$. This value is reproduced to the nearest ms⁻¹ by the 50-level resolution.

The remaining (“internal”) modes provide a mathematically-acceptable basis for representing the vertical structure of model variables, but do not correspond to modes of the continuous equations. They have an oscillatory structure

with amplitude approximately proportional to $1/(\sqrt{p})$, and represent standing waves that can exist because of the reflective nature of the upper boundary condition $\omega = 0$ applied in the model (Lindzen *et al.*, 1968).

It can be seen from Table 1 that only a limited number of modes are associated with gravity-wave phase speeds of the order of 50ms^{-1} or more. It is found necessary to initialize only these modes. Indeed, applying the non-linear normal-mode technique to the slower, higher-order modes is counter-productive, as the iterative solution fails to converge for these modes. The example presented in Fig. 8 illustrates the extent to which high frequency oscillations are suppressed by initializing up to the five gravest modes. Initializing four or five modes appears adequate in this case. These results were for a nine-level vertical resolution for which the gravity-wave phase speeds of the fourth and fifth modes were 70ms^{-1} and 39ms^{-1} respectively, for a 300K reference temperature.

2.3 Horizontal decomposition

The horizontal decomposition is illustrated most conveniently for the case of a global spectral model such as used at ECMWF. Firstly, the shallow-water momentum equations are rewritten in terms of the divergence, \hat{D}_l (given by (36)), and relative vorticity, $\hat{\xi}_l$:

$$\hat{\xi}_l = \frac{1}{a \cos \theta} \left(\frac{\partial \hat{v}_l}{\partial \lambda} - \frac{\partial}{\partial \theta} (\hat{u}_l \cos \theta) \right) \quad (38)$$

The non-linear shallow-water equations are:

$$\frac{\partial}{\partial t} \hat{\xi}_l = -\frac{2\Omega \cos \theta}{a} \hat{v}_l - 2\Omega \sin \theta \hat{D}_l + \hat{N}_{\xi l} \quad (39)$$

$$\frac{\partial}{\partial t} \hat{D}_l = 2\Omega \sin \theta \hat{\xi}_l - \frac{2\Omega \cos \theta}{a} \hat{u}_l - \nabla^2 \hat{P}_l + \hat{N}_{Dl} \quad (40)$$

and

$$\frac{\partial}{\partial t} \hat{P}_l = -\Phi_l \hat{D}_l + \hat{N}_{Pl} \quad (41)$$

These equations are recast in terms of the spectral coefficients ξ_n^m , D_n^m and P_n^m in spherical-harmonic expansions:

$$\hat{\xi}_l = \sum_{m=-M}^M \sum_{n=|m|}^M \xi_n^m(t) P_n^m(\sin \theta) e^{im\lambda} \quad (42)$$

$$\hat{D}_l = \sum_{m=-M}^M \sum_{n=|m|}^M D_n^m(t) P_n^m(\sin \theta) e^{im\lambda} \quad (43)$$

and

$$\hat{P}_l = \sum_{m=-M}^M \sum_{n=|m|}^M \phi_n^m(t) P_n^m(\sin \theta) e^{im\lambda} \quad (44)$$

As the prognostic variables are real valued, ξ_n^{-m} , D_n^{-m} and ϕ_n^{-m} are the complex conjugates of ξ_n^m , D_n^m and ϕ_n^m , and we need consider only the equations for $m \geq 0$. These equations can be separated into $M + 1$ sets, one for

each zonal wavenumber m . We introduce scaled variables:

$$\tilde{\xi}_n^m = \frac{\alpha}{\sqrt{n(n+1)}} \xi_n^m \quad (45)$$

$$\tilde{D}_n^m = i \frac{\alpha}{\sqrt{n(n+1)}} D_n^m \quad (46)$$

and

$$\tilde{\phi}_n^m = \frac{1}{\sqrt{\Phi_l}} \phi_n^m \quad (47)$$

For each m , we use $\tilde{\xi}$, \tilde{D} and $\tilde{\phi}$ to denote column vectors of dimension $M - m + 1$ with elements $\tilde{\xi}_n^m$, \tilde{D}_n^m and $\tilde{\phi}_n^m$, for $n = m, m + 1, \dots, M$.

The shallow water equations (39), (40) and (41) may then be written in the form:

$$\frac{d}{dt} \begin{bmatrix} \tilde{\xi} \\ \tilde{D} \\ \tilde{\phi} \end{bmatrix} = i \begin{bmatrix} \beta & \mathbf{F} & \mathbf{0} \\ \mathbf{F} & \beta & \mathbf{C} \\ \mathbf{0} & \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \tilde{\xi} \\ \tilde{D} \\ \tilde{\phi} \end{bmatrix} + \begin{bmatrix} \tilde{N}_\xi \\ \tilde{N}_D \\ \tilde{N}_\phi \end{bmatrix} = i \mathbf{L} \begin{bmatrix} \tilde{\xi} \\ \tilde{D} \\ \tilde{\phi} \end{bmatrix} + \begin{bmatrix} \tilde{N}_\xi \\ \tilde{N}_D \\ \tilde{N}_\phi \end{bmatrix} \quad (48)$$

where β and \mathbf{C} are diagonal matrices, with diagonal elements β_n^m and c_n , where

$$\beta_n^m = \frac{2\Omega m}{n(n+1)} \quad (49)$$

and

$$c_n = \frac{1}{\alpha} \sqrt{n(n+1)\Phi_l} \quad (50)$$

\mathbf{F} is a symmetric tridiagonal matrix with diagonal elements zero, and off-diagonal elements f_n^m for $n > m$, with

$$f_n^m = \frac{2\Omega}{n} \sqrt{(n^2 - 1)(n^2 - m^2)/(4n^2 - 1)} \quad (51)$$

In the absence of rotation $\beta = \mathbf{F} = \mathbf{0}$, and the non-trivial eigenvectors of \mathbf{L} are gravity waves with frequencies $\pm c_n$, as noted earlier.

With rotation, the basic Coriolis effect (as occurs in an f-plane geometry) is represented by the matrix \mathbf{F} , while the matrix β represents the variation of the Coriolis effect with latitude, the so-called ‘‘beta’’ effect. Pure barotropic Rossby waves are obtained by imposing $\tilde{D} = 0$ and have frequencies β_n^m .

Equation (48) provides an explicit form for \mathbf{L} :

$$\mathbf{L} = \begin{bmatrix} \beta & \mathbf{F} & \mathbf{0} \\ \mathbf{F} & \beta & \mathbf{C} \\ \mathbf{0} & \mathbf{C} & \mathbf{0} \end{bmatrix} \quad (52)$$

Eigenvalues and eigenfunctions of \mathbf{L} can readily be obtained using standard software for matrix analysis. Separation into Rossby and gravity modes can in practice be achieved simply on the basis of wave frequencies; the eastward moving waves and faster westward moving waves are the modes to be initialized.

2.4 Implicit initialization

Although formally straightforward for global spectral models, explicit normal-mode initialization has a demanding requirement for computer storage of the normal modes in the case of high horizontal resolution. Similar considerations apply to global models based on alternative horizontal discretizations. Two additional problems may arise for a limited-area model. Firstly, the map projection may cause horizontal separability to be lost. Secondly, there may be a problem in defining lateral boundary conditions for the normal modes. This led to the development of implicit normal mode initialization. This is presented here for the case of a global spectral model; Temperton(1988) introduced the method for a regional model using a polar stereographic projection and finite elements with a non-uniform grid.

We start from the formula for initialization increments given by equation (16):

$$\mathbf{L}\Delta\mathbf{x}^{(n+1)} = i\mathbf{E}_G\mathbf{E}_G^T\left(\frac{d\mathbf{x}}{dt}\right)^{(n)} = i\left(\frac{d\mathbf{x}}{dt}\right)_G^{(n)}$$

The approach relies on being able to determine the “fast” gravity-wave component $(d\mathbf{x}/dt)_G^{(n)}$ from the total tendency $(d\mathbf{x}/dt)^{(n)}$ and on being able to solve the linear system for the increments $\Delta\mathbf{x}^{(n+1)}$. This can be achieved by approximating the operator \mathbf{L} . The approximation is to neglect the “beta-effect” in the vorticity equation. Equation (52) is replaced by:

$$\mathbf{L} = \begin{bmatrix} \mathbf{0} & \mathbf{F} & \mathbf{0} \\ \mathbf{F} & \beta & \mathbf{C} \\ \mathbf{0} & \mathbf{C} & \mathbf{0} \end{bmatrix} \quad (53)$$

Equation (49) indicates that the approximation will be a good one for small horizontal scales (large n) for which the non-zero elements β_n^m of the diagonal matrix β are small.

With this approximation, the linearized form of (48) gives:

$$\frac{d\tilde{\xi}}{dt} = i\mathbf{F}\tilde{D} \quad (54)$$

$$\frac{d\tilde{D}}{dt} = i\beta\tilde{D} + i(\mathbf{F}\tilde{\xi} + \mathbf{C}\tilde{\phi}) \quad (55)$$

and

$$\frac{d\tilde{\phi}}{dt} = i\mathbf{C}\tilde{D} \quad (56)$$

The Rossby modes are stationary and non-divergent:

$$\frac{d\tilde{\xi}}{dt} = \tilde{D} = 0 \quad (57)$$

and, from (55), satisfy the balance equation:

$$\mathbf{C}\tilde{\phi} = -\mathbf{F}\tilde{\xi} \quad (58)$$

From (54) and (56), a gravity mode (which has $\tilde{D} \neq 0$) must have:

$$\tilde{\xi} = \mathbf{F}\mathbf{C}^{-1}\tilde{\phi} \quad (59)$$

The expanded form of (16) is

$$\begin{bmatrix} \mathbf{0} & \mathbf{F} & \mathbf{0} \\ \mathbf{F} & \beta & \mathbf{C} \\ \mathbf{0} & \mathbf{C} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \Delta\tilde{\xi}^{(n+1)} \\ \Delta\tilde{D}^{(n+1)} \\ \Delta\tilde{\phi}^{(n+1)} \end{bmatrix} = i \begin{bmatrix} \left(\frac{d\tilde{\xi}}{dt}\right)_G^{(n)} \\ \left(\frac{dD}{dt}\right)_G^{(n)} \\ \left(\frac{d\tilde{\phi}}{dt}\right)_G^{(n)} \end{bmatrix} \quad (60)$$

Temperton(1989) showed that (57), (58) and (59) can be used to derive formulae for initialization increments that do not require explicit computation of gravity modes:

$$(\mathbf{F}^2 + \mathbf{C}^2)\Delta\tilde{D}^{(n+1)} = i\mathbf{F}\left(\frac{d\tilde{\xi}}{dt}\right)^{(n)} + i\mathbf{C}\left(\frac{d\tilde{\phi}}{dt}\right)^{(n)} \quad (61)$$

and

$$(\mathbf{F}^2 + \mathbf{C}^2)(\mathbf{C}^{-1}\Delta\tilde{\phi}^{(n+1)}) = i\left(\frac{d\tilde{D}}{dt}\right)^{(n)} - \beta\Delta\tilde{D}^{(n)} \quad (62)$$

with

$$\Delta\tilde{\xi}^{(n+1)} = \mathbf{F}(\mathbf{C}^{-1}\Delta\tilde{\phi}^{(n)}) \quad (63)$$

The matrix $(\mathbf{F}^2 + \mathbf{C}^2)$ is pentadiagonal, but the equations to be solved ((61) and (62)) can in fact each be split into two separate equations for odd and even spectral components with diagonally-dominant tridiagonal matrices on the left-hand sides. Their solution is thus straightforward.

In practice this implicit method has been found to give results that are negligibly different from explicit normal-mode initialization for most of the wavenumber range of high resolution models. A mixed scheme has been adopted for the ECMWF spectral model. The implicit method is used by default to initialize total wavenumbers in the range $n > 21$, with the explicit method used for $n \leq 21$, as the neglect of the β term becomes increasingly inaccurate

as the total wavenumber decreases.

2.5 Diabatic initialization

The formalism of non-linear normal-mode initialization presented in 2.1 is general, but in practice it was found necessary to suppress the diabatic (parametrized) components of the term $\mathbf{N}(\mathbf{x})$ in early applications of the method (Williamson and Temperton, 1981). In particular, the large and rapidly varying temperature tendencies arising from the parametrization of convection were found to inhibit convergence of the method. Moreover, the basic initialization condition expressed by (9):

$$\frac{d}{dt}\alpha_j = 0$$

is inappropriate for the westward-moving thermal tidal waves forced by the daily westward progression of the solar heating of the atmosphere.

The form of initialization in which only the adiabatic tendencies from the model are computed at each iteration is referred to as *adiabatic* non-linear normal-mode initialization. Although successful in preventing the growth of high-frequency oscillations in subsequent forecasts, adiabatic initialization also suppresses the slowly-varying large-scale circulations that are a feature of the tropical atmosphere. Two of the approaches that have been adopted to circumvent the problem are outlined below.

2.5 (a) Full-field initialization. A diagnosis and discussion of the problems of *adiabatic* initialization has been given by Wergen (1987), who also described the *diabatic* non-linear normal-mode initialization scheme introduced operationally at ECMWF in 1982. This scheme was used in conjunction with the Centre's Optimum Interpolation analysis until the latter was replaced by a variational data assimilation scheme early in 1996. The diabatic initialization scheme was modified in 1986 to include the treatment of tides specified below.

The approach involves estimating a steady large-scale diabatic forcing. This is applied as a tendency term that does not change during the iterations of the non-linear normal-mode initialization, thereby avoiding the problem of non-convergence. The tidal problem is dealt with by subtracting an estimate of the tidal tendency from the adiabatic tendency. The steps involved are:

- Performing a short forecast (~2h) from the uninitialized analysis, using all parametrizations but suppressing the diurnal cycle in the radiation scheme;
- Time-averaging the diabatic tendencies computed at each timestep of this forecast;
- Projecting the time-averaged tendencies onto gravity modes, keeping only low-frequency components (with periods > 11h);
- Subtracting tidal tendencies computed from a (10-day) time series of the most recent analyses;
- Adding the resulting fixed filtered tendencies to the adiabatic tendencies computed for each iteration of the initialization procedure;
- Computing the initialization increment in the usual way, using the modified tendencies from the preceding step.

This approach, though rather cumbersome, worked well in practice.

2.5 (b) Incremental initialization. The simpler alternative approach of incremental initialization (Puri *et al.*, 1982; Ballish *et al.*, 1992) was adopted when ECMWF moved to the variational data assimilation scheme. This approach does not necessarily rely on using the normal-mode technique for initialization. In particular, it can be used in conjunction with the digital filtering technique described later.

We illustrate the approach here in its basic form. Let \mathbf{x}_b denote a background state, and as before \mathbf{x}_U the uninitialized state and \mathbf{x}_I the initialized state. Let $Init(\mathbf{x})$ denote the result of an adiabatic initialization of \mathbf{x} . The incremental initialization scheme is then:

$$\mathbf{x}_I = \mathbf{x}_b + Init(\mathbf{x}_U) - Init(\mathbf{x}_b) \quad (64)$$

Provided the differences between \mathbf{x}_U and \mathbf{x}_b are small, the scheme preserves the large-scale diabatic balance present in the background field. However, if a component is handled poorly by the initialization procedure and is improved in \mathbf{x}_U compared with \mathbf{x}_b , the thermal tide for example, the improvement in \mathbf{x}_U may not carry through to the initialized analysis \mathbf{x}_I .

As applied in a conventional data assimilation system, the state \mathbf{x}_b is the short-range forecast that provides the background for the analysis, \mathbf{x}_U is the result of the analysis and \mathbf{x}_I is the starting point for the next forecast. The approach has, however, wider applicability. (64) could also be used, for example, to provide initial conditions for a forecast from an analysis produced using a different data assimilation system, such as that of a different forecasting centre. In this case, \mathbf{x}_b would be the (uninitialized) analysis produced by a standard data assimilation using the forecast model for which initialization is required, and \mathbf{x}_U would be the analysis produced by the different data assimilation system.

3. CONTROL OF GRAVITY WAVES IN THE ECMWF VARIATIONAL DATA ASSIMILATION SYSTEM

The basic variational data assimilation problem is to determine the model state \mathbf{x} that minimizes a scalar cost function J . J comprises three elements:

$$J = J_b + J_o + J_c \quad (65)$$

where J_b is the background cost function defined in terms of the deviation of the analysis from the background state \mathbf{x}_b and J_o is the observation cost function defined in terms of the deviation of the analysis from the observations in the case of three-dimensional variational assimilation (3D-Var). In the case of four-dimensional assimilation (4D-Var), J_o is the observation cost function defined in terms of the deviation from the observations of a forecast from the state \mathbf{x} . The two elements J_b and J_o are discussed in the companion lecture notes on *Data assimilation concepts and methods*. The primary control of gravity waves in the data assimilation comes in general through a multivariate formulation of the J_b term, although in 4D-Var a contribution can come through the model integration that is implicit in the term J_o .

The term J_c represents additional constraints on the analysis. These could include physical constraints, for example that the relative humidity be between 0 and 100%, or a constraint on high frequencies based on digital filtering. In the current implementation at ECMWF, J_c is given by

$$J_c = \varepsilon \left\| \left(\frac{d\mathbf{x}}{dt} \right)_G - \left(\frac{d\mathbf{x}_b}{dt} \right)_G \right\|^2 \quad (66)$$

where $(d/dt)_G$ denotes the projection of the tendency (d/dt) onto gravity waves and $\| \|^2$ denotes a simple energy-based norm defined by a weighted sum of squares of spectral coefficients. The overall weighting factor ε was adjusted during the development of ECMWF's 3D-Var system to remove oscillations in surface pressure (typically a fraction of an hPa in magnitude) that were found to occur in the absence of J_c (Courtier et al., 1998).



Examples of the evolution of the cost function J and its three components over 70 iterations of a 3D-Var minimization are presented in Fig. 11. The upper panel shows how in absolute terms J_c makes only a tiny contribution to the overall cost function. The background state \mathbf{x}_b is used as the starting value for the minimization, and J_b and J_c are thus initially zero. J_o is subsequently decreased substantially¹, at the expense of some increase in J_b and a slight increase in J_c . The plots in the lower panel of Fig. 11 indicate an increase in small-scale gravity-wave activity when J_c is excluded from the minimization (but computed as a diagnostic). In this case J_c grows to about twice the value that develops when it provides part of the constraint on the analysis. It should be noted that the elements of the cost function are defined globally; despite its small overall value J_c may provide an important local constraint, close to steep orography in particular.

The variational data assimilation scheme is implemented in an incremental form in which the minimization is carried out at a lower resolution than that of the background forecast. Until March 1999, initialization was applied twice in the procedure. The low-resolution analysis requires an interpolated low-resolution background field, and this was initialized applying adiabatic non-linear normal-mode initialization for scales with $n \geq 20$. Then, on completion of the low-resolution analysis, incremental adiabatic non-linear normal-mode initialization was applied in forming the high-resolution analysis \mathbf{x}_a :

$$\mathbf{x}_a = \mathbf{x}_b + \text{Init}(\mathbf{x}_b + (\delta\mathbf{x}_a)_{LR \rightarrow HR}) - \text{Init}(\mathbf{x}_b) \quad (67)$$

Here $(\delta\mathbf{x}_a)_{LR \rightarrow HR}$ is the increment of the low-resolution analysis (the difference between the low-resolution analysis and the low-resolution background) interpolated to the high resolution. This incremental initialization was applied operationally only for scales with $n \geq 20$ from May 1997 onwards, for the primary purpose of adjusting the low-resolution analysis to the high-resolution orography. It was, however, applied to all scales in the original operational implementation of 3D-Var, as discussed further below.

1. The sharp fall near iteration number 30 is due to the initiation of variational quality control.

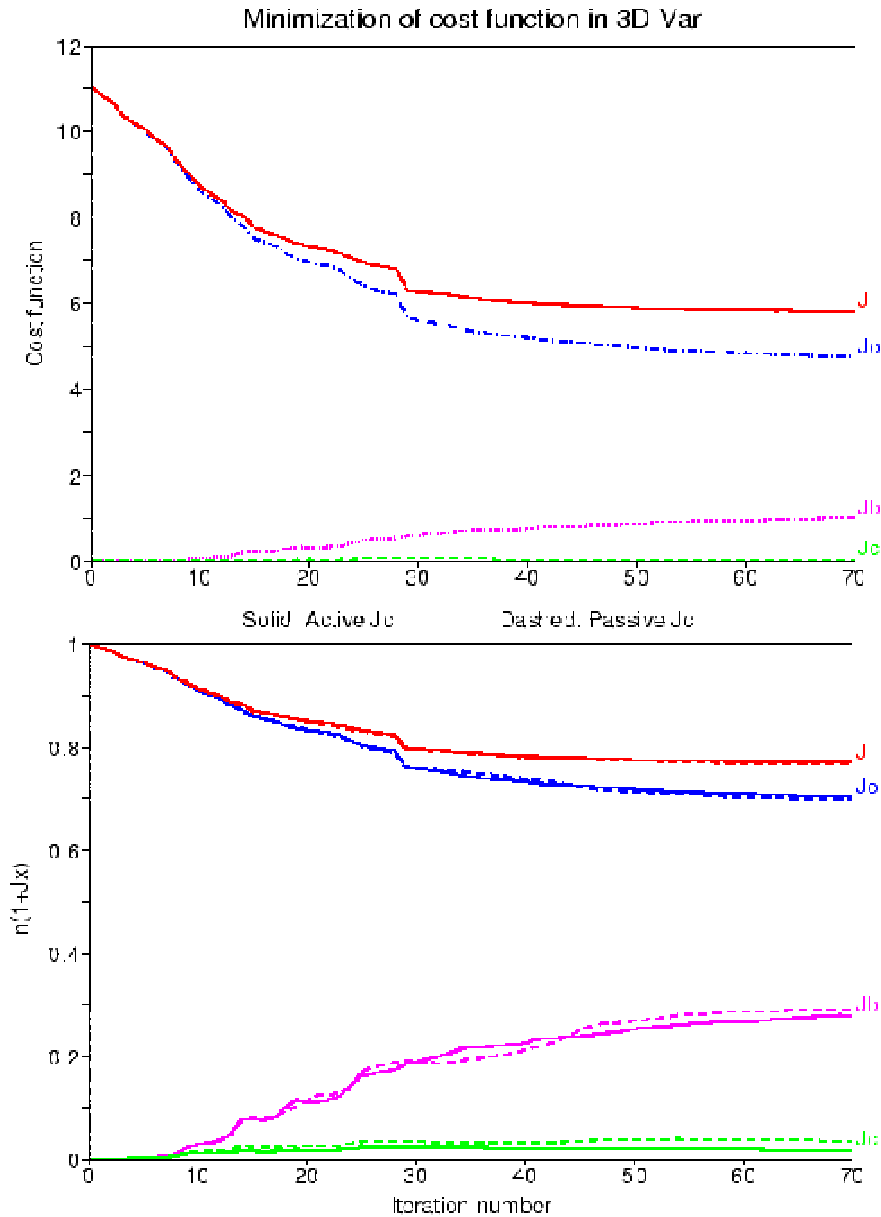


Figure 11. Evolution of the cost function J and its components J_o , J_b and J_c during the minimization in a standard cycle of 3D-Var (upper panel) and plotted on a logarithmic scale (lower panel) both for the standard cycle (solid) and for a modified cycle (dashed) in which J_c was calculated for diagnostic purposes but not activated in the minimization.

Some examples of the time evolution of the surface pressure are presented in Figs. 12 and 13. Plots are shown for the two points $40^\circ\text{N } 90^\circ\text{W}$ and $30^\circ\text{N } 90^\circ\text{E}$ used to illustrate results from earlier studies of initialization in Figs. 7 and 8. The forecasts were carried out after three cycles of 3D-Var using the initialization configuration in question. Fig. 12 shows the impact of excluding the constraint J_c . Both forecasts shown were from assimilations in which the incremental initialization was applied on all scales, and this is evidently sufficient to prevent high-frequency gravity-wave oscillations whether J_c is activated or not. Suppressing J_c gives little change at $40^\circ\text{N } 90^\circ\text{W}$ over the Great Plains, but has modified the starting value of surface pressure by about 0.5hPa at the Himalayan point. A pronounced semi-diurnal tidal oscillation can be seen at the latter point.

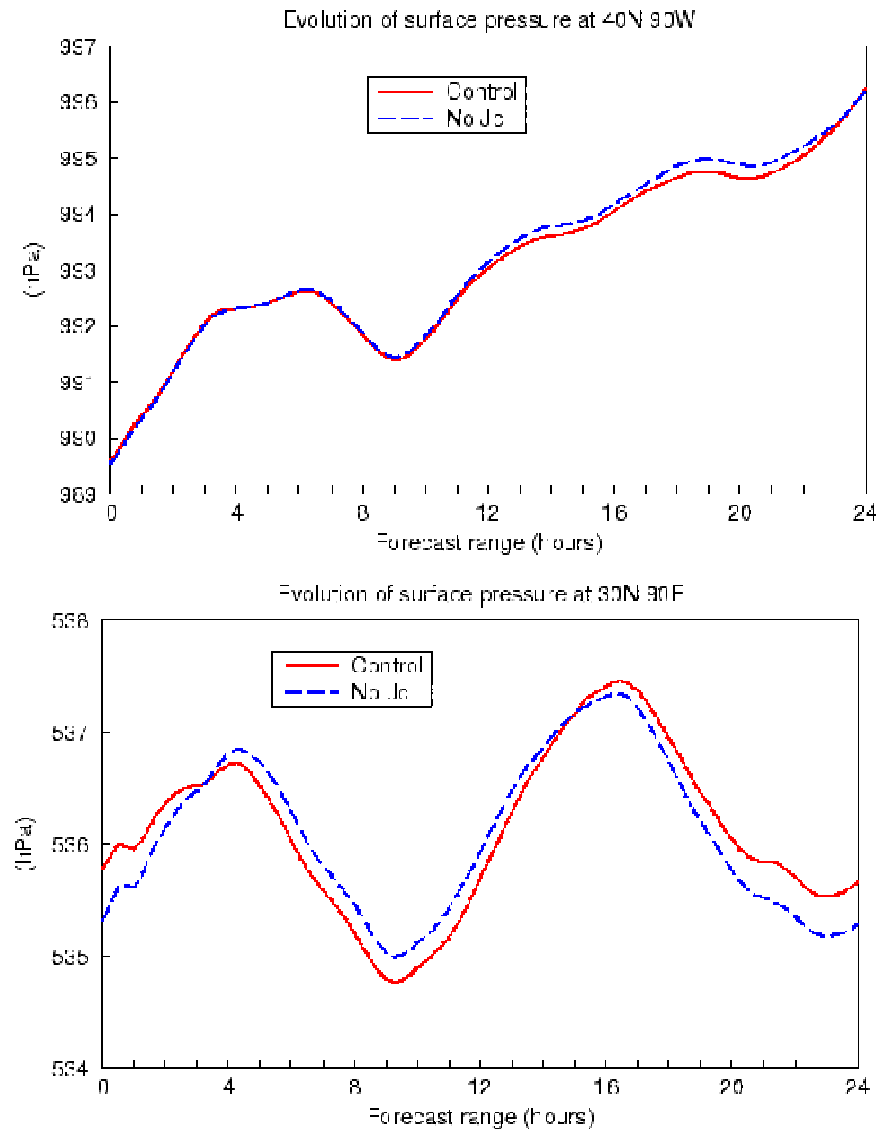


Figure 12. Surface pressure (hPa) as a function of time for a control forecast (solid) and a forecast following three cycles of 3D-Var in which J_c was not activated (dashed), at 40°N 90°W (upper) and 30°N 90°E (lower).

Fig. 13 shows the effect of removing the incremental initialization of the larger scales ($n < 20$), and of removing initialization (and the J_c constraint) completely. Removing the larger-scale initialization allows some high-frequency oscillations to develop, although even at the Himalayan point the amplitude is barely over 0.1hPa. Completely removing initialization has a larger effect, especially in the first few steps at the Himalayan point. Gravity-wave oscillations are nevertheless much smaller than in the forecast from the uninitialized analysis shown in Fig. 7, presumably because the present forecasts come from a consistent and much more modern data assimilation system.

Further examination of these issues was carried out as part of the development of the 50-level version of the model which became operational in March 1999. It was found that J_c continues to play a small but useful role, but that the initialization steps could be eliminated without significant deterioration of analysis and forecast quality. As the amplitudes of the internal modes vary approximately as $1/(\sqrt{p})$ for small p , elimination of the initialization steps

avoided a problem of large initialization increments close to the top of the 50-level model.

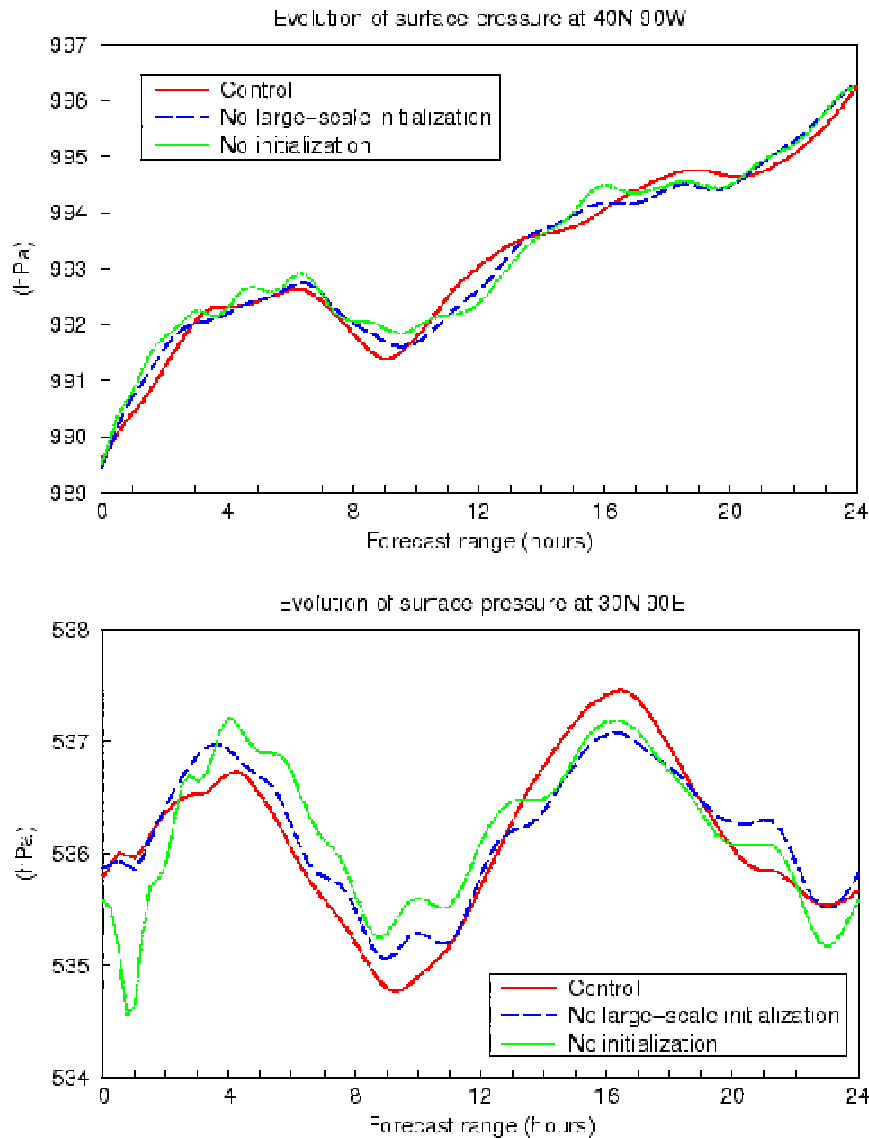


Figure 13. Surface pressure (hPa) as a function of time for a control forecast (solid) and for forecasts following three cycles of 3D-Var with no large-scale initialization (dashed) and no initialization at all (dotted), at 40°N 90°W (upper) and 30°N 90°E (lower).

The removal of the incremental initialization of the scales with $n < 20$ was implemented operationally at ECMWF in May 1997 at the same time as a change to the background term J_b , moving from the formulation described by [Courtier *et al.*\(1998\)](#) (referred to as “old” J_b) to that reported by [Bouttier *et al.*\(1997\)](#) (the “new” J_b). The new J_b was used for the forecasts shown in [Figs. 12](#) and [13](#). Some idealized tests carried out prior to the change provide examples of the working of the initialization and of the background constraint.

Fig. 14 shows increments in 850hPa height due to several idealized isolated observations of this field, specified such that the observed deviation from the background field was the same at each point. Results are shown for a single cycle of 3D-Var. The old J_b reduces to a univariate formulation in the tropics, and produces localized increments of similar magnitude at all locations (upper-left panel). However, the incremental initialization (applied to all scales) removes most of the increment in the tropics (lower-left panel). The new J_b imposes a semi-empirical (close to linear) balance. It produces more of a large-scale increment, and smaller local increments in the tropics (upper-right panel). More of each local increment survives initialization. Incremental initialization thus plays a smaller role in imposing balance on the analysis in the case of the new J_b . This was an important factor in the decision to remove initialization for scales $n < 20$ in the operational system.

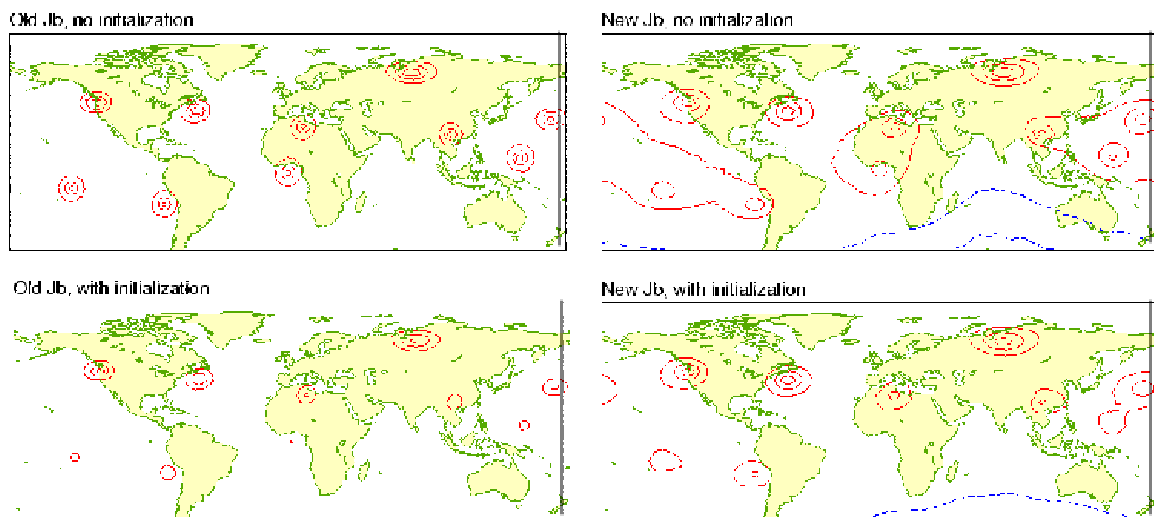


Figure 14. Analysis increments in 850hPa height for a set of idealized height observations at 850hPa, for the J_b operational prior to May 1997 (left) and that operational after May 1997 (right), with no initialization (upper) and after incremental non-linear normal-mode initialization (left).

Close-ups of the height increments at one location and the associated wind increments are presented in Fig. 15. The multivariate formulations of the old and new J_b both produce increments which are close to being in geostrophic balance, and initialization changes are much smaller than in the tropics. They are slightly smaller with the new than the old J_b . The height increments are reduced by initialization and the wind increments are increased.

The new J_b produces a divergent component to the wind increment at the ground, as can be seen in the plots for 1000hPa shown in Fig. 16. The increment in divergence is shallow, and survives initialization because the latter is applied only to the first five, relatively deep, vertical modes.

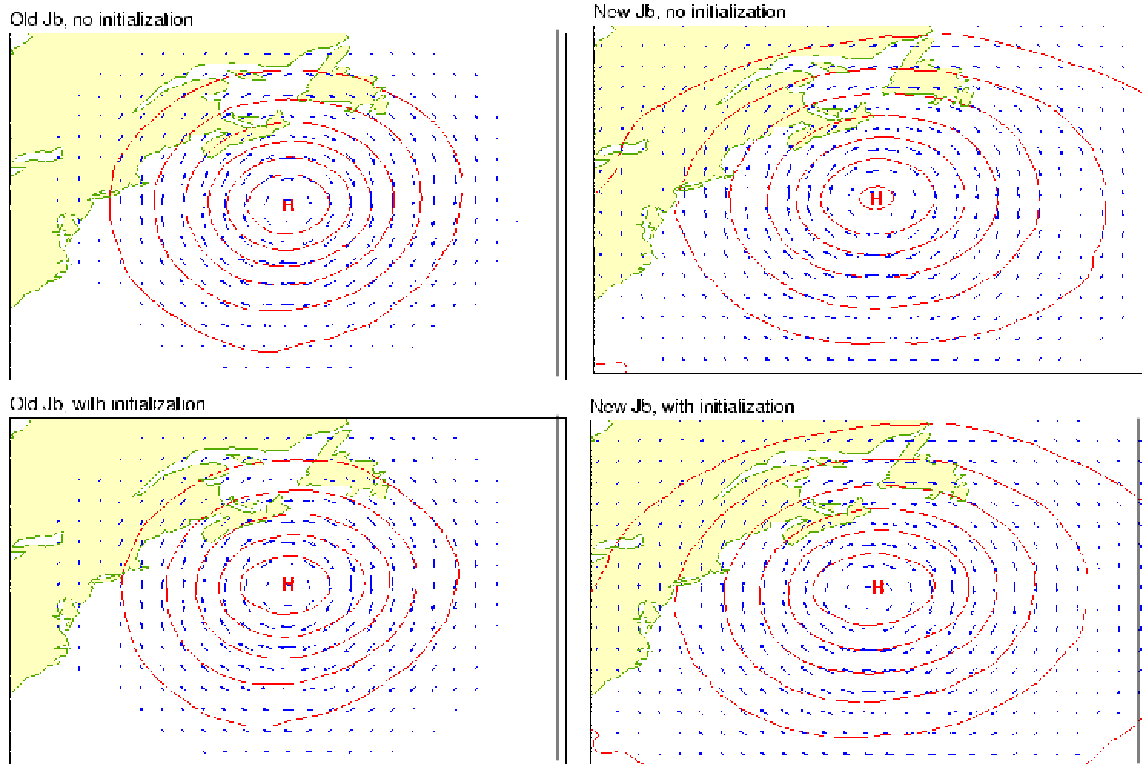


Figure 15. Analysis increments in 850hPa height and wind for a set of idealized height observations at 850hPa, for the J_b operational prior to May 1997 (left) and that operational after May 1997 (right), with no initialization (upper) and after incremental non-linear normal-mode initialization (left).

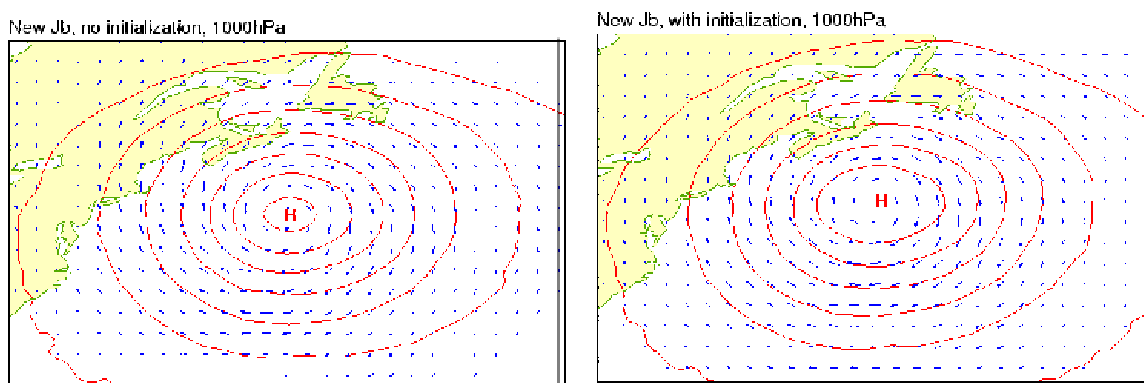


Figure 16. Analysis increments in 1000hPa height and wind for a set of idealized height observations at 850hPa, for the J_b operational after May 1997, with no initialization (left) and after incremental non-linear normal-mode initialization (right).

The thermal tide provides a final example. The variational analysis is able to “draw” to the tidal signal present in the surface pressure observations, but the signal is not fully retained in the ensuing forecast. The analysis thus produces increments which improve the description of the tides. A fraction of the improvement is lost, however, if incremental initialization is applied to large scales. Fig. 17 illustrates how the analysis generally fits better the

surface-pressure observations from a frequently reporting tropical island station when the initialization is restricted to scales $n \geq 20$.

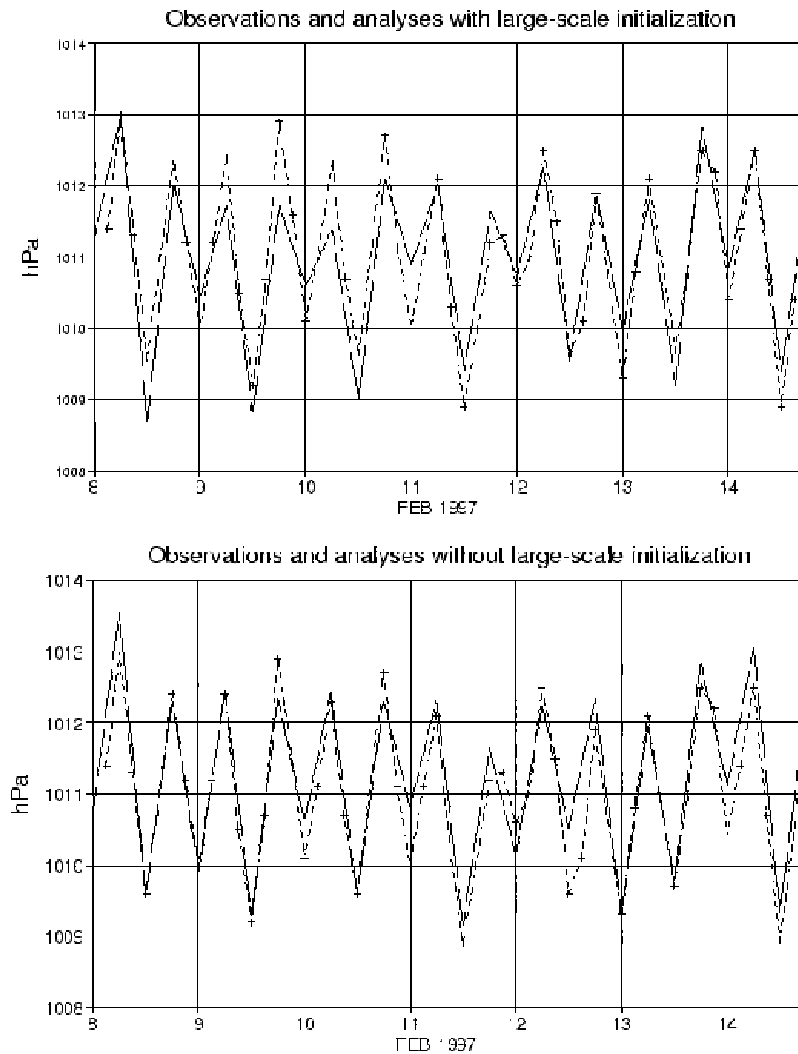


Figure 17. Surface pressure (hPa) from 00UTC 8 February to 18UTC 14 February 1997, as observed at Seychelles International Airport (5°S, 56°E; dashed line) and as analysed at this location (solid) with (upper) and without (lower) large-scale initialization.

4. DIGITAL FILTERING

The method of digital filtering provides an approach to initialization that is conceptually simple and easier to implement than non-linear normal-mode initialization. It involves generating a sequence of model fields and then applying a filter to the resulting time series for each model grid-point and variable or each model spectral coefficient. The filter is chosen to reduce the amplitudes of high-frequency components of the time series to acceptable levels. The forecast is then run from an appropriate point within the filtered time series.

4.1 Adiabatic, non-recursive filtering

We consider first the simplest case of adiabatic, non-recursive filtering. We denote the uninitialized analysis by $\mathbf{x}^{(0)}$. A forward adiabatic integration is carried out for N timesteps to generate the set of values:

$$\{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(N)}\}$$

where $\mathbf{x}^{(n)}$ denotes the model state after n timesteps. Also, a backward adiabatic integration of the same length is carried out to generate values:

$$\{\mathbf{x}^{(-1)}, \mathbf{x}^{(-2)}, \dots, \mathbf{x}^{(-N)}\}$$

The filtered initial state, or initialized analysis, is then given generally by:

$$\mathbf{y}^{(0)} = \text{Init}(\mathbf{x}^{(0)}) = \frac{1}{h} \sum_{n=-N}^N h_n \mathbf{x}^{(n)} \quad (68)$$

where

$$h = \sum_{n=-N}^N h_n \quad (69)$$

The original application of digital filtering for initialization by Lynch and Huang (1992) used a modification of a basic filter defined by:

$$h_n = \frac{\sin(n\nu_c \Delta t)}{n\pi} \quad (70)$$

This represents the discrete equivalent of the filter of a continuous function that leaves low-frequency components unchanged but removes high-frequency components completely, multiplying a fourier component $\exp(i\nu t)$ by $H(\nu)$, where

$$H(\nu) = \begin{cases} 1 & |\nu| \leq |\nu_c| \\ 0 & |\nu| > |\nu_c| \end{cases} \quad (71)$$

The modified ‘‘Lanczos’’ filter was defined by:

$$h_n = \left\{ \frac{\sin[n\pi/(N+1)]}{n\pi/(N+1)} \right\} \left\{ \frac{\sin(n\nu_c \Delta t)}{n\pi} \right\} \quad (72)$$

The term

$$\frac{\sin[n\pi/(N+1)]}{n\pi/(N+1)}$$

in (72) provides what is known as a Lanczos window.

The filters whose coefficients are given by (70) and (72) have the property of leaving the phase of a sinusoidal wave unchanged (apart from a possible 180° shift) while reducing the amplitude of the wave. The reduction in amplitude is shown as a function of wave period in Fig. 18. The calculation is for a cut-off period ($2\pi/v_c$) of six hours, a span ($2N\Delta t$) also of six hours and a timestep (Δt) of 15 minutes. The solid line denotes the amplitude response for the continuous filter $H(v)$. The basic filter (dashed line) exhibits the familiar Gibbs oscillations, which are greatly reduced by application of the Lanczos filter (dotted line). The Lanczos filter gives less attenuation of longer-period waves, but weaker filtering of waves with periods just shorter than six hours.

The impact on waves of unit input amplitude and periods of 3, 6, 12 and 24 hours is shown in Fig. 19. The amplitudes of the 6-, 12- and 24-hour period waves are reduced less when the Lanczos window is included. The phase of the 3-hour wave is reversed by the basic filter. It has smaller amplitude and no phase reversal with the Lanczos window.

An example of the noise reduction found when

and Huang (1992) applied the method to a version of the HIRLAM model is presented in Fig. 20. The Lanczos filter was used with six-hour cutoff and span, and the model was run with a six-minute timestep. Fig. 20 shows that digital filtering initialization reduced noise more effectively than the normal-mode initialization scheme developed for HIRLAM. Other diagnostics confirmed the success of the digital filtering approach.

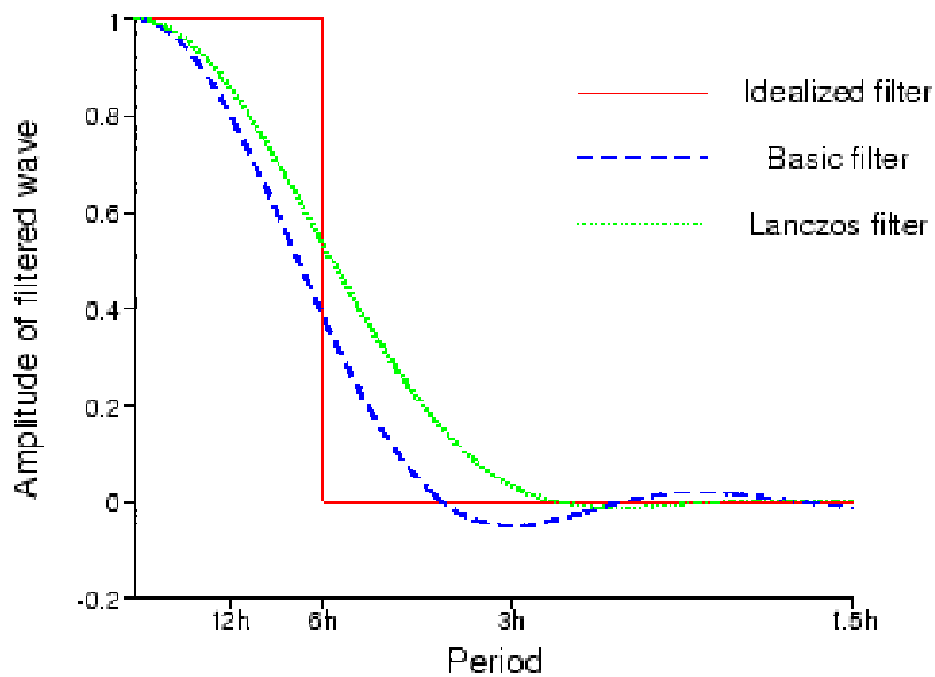


Figure 18. Amplitude of the filtered wave as a function of wave period for an input sinusoidal wave of unit amplitude. The idealized continuous filter has a 6-hour cutoff, and the corresponding basic discrete filter and basic filter modified by a Lanczos window are shown for a 15-minute timestep and 6-hour span, following Lynch and Huang (1992).

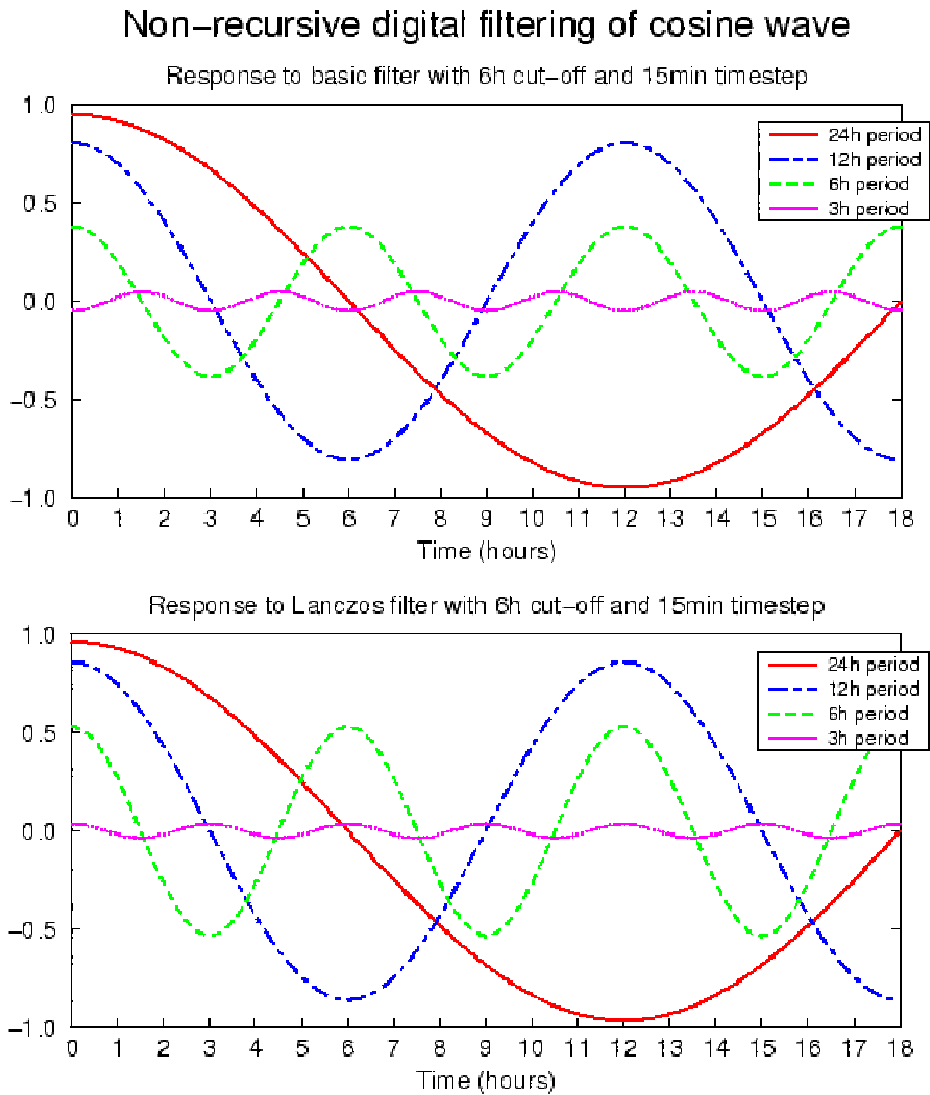


Figure 19. The response of waves of periods 3, 6, 12 and 24 hours to the basic filter (upper) and the filter modified by the Lanczos window (lower).

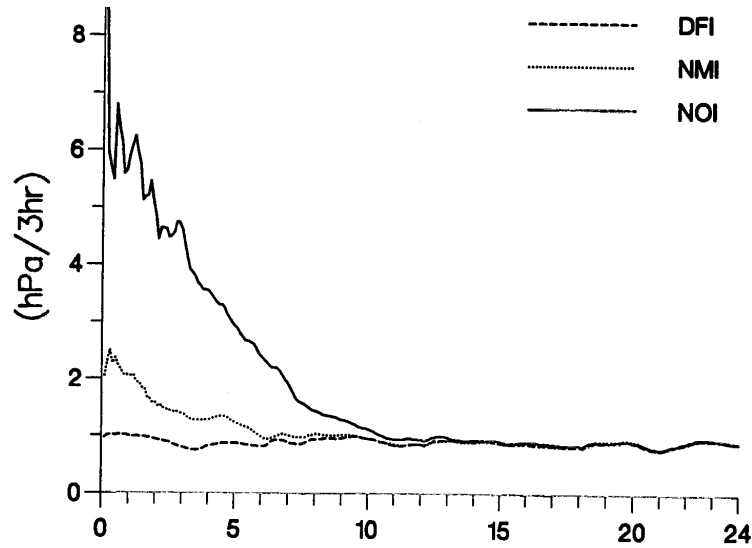


Figure 20. Evolution of the mean absolute surface pressure tendency (hPa /3h) in 24-hour forecasts starting from an uninitialized analysis (solid line) and from analyses initialized using non-linear normal-mode initialization (dotted) and non-recursive digital filtering (dashed), from Lynch and Huang(1992).

4.2 Diabatic, recursive filtering

The backward integration used to generate the values $\{\mathbf{x}^{(-1)}, \mathbf{x}^{(-2)}, \dots, \mathbf{x}^{(-N)}\}$ in the approach described in the preceding subsection cannot be carried out using parametrizations of irreversible physical processes. Use of a *recursive* filter offers one way to use the digital filtering method for diabatic initialization (Lynch and Huang, 1994).

Consider the sequence of values $\{\mathbf{x}^{(0)}, \mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)}, \dots\}$ of model variables from consecutive timesteps of a forecast starting from the uninitialized analysis $\mathbf{x}^{(0)}$, possibly including diabatic and frictional processes. A recursive filter of order N is defined in general by the $N + 1$ values $\{a^{(0)}, a^{(1)}, \dots, a^{(N)}\}$ and the N values $\{b^{(1)}, b^{(2)}, \dots, b^{(N)}\}$ in the following expression for the filtered value $\mathbf{y}^{(n)}$ at step n :

$$\mathbf{y}^{(n)} = \sum_{j=0}^N a^{(j)} \mathbf{x}^{(n-j)} + \sum_{j=1}^N b^{(j)} \mathbf{y}^{(n-j)} \quad (73)$$

The process is started by applying lower-order filters to compute values of $\mathbf{y}^{(n)}$ for $n < N$. A non-recursive implementation of this filter is set out in Appendix C.

We illustrate the case of the Second-order Quick-Start filter presented by Lynch and Huang (1994). This has $N = 2$ and:

$$y^{(n+1)} = (a^{(0)} x^{(n+1)} + a^{(1)} x^{(n)} + a^{(2)} x^{(n-1)}) + (b^{(1)} y^{(n)} + b^{(2)} y^{(n-1)}) \quad (74)$$

with

$$y^{(0)} = x^{(0)} \quad (75)$$

and

$$y^{(1)} = ax^{(1)} + (1-a)x^{(0)} \quad (76)$$

If we define:

$$\mu_c = \tan\left(\frac{\nu_c \Delta t}{2}\right) \quad (77)$$

and

$$s = \mu_c \sqrt{1 + \sqrt{2}} \quad (78)$$

the filter coefficients are given by:

$$a = \frac{1}{1 + \mu_c} \quad (79)$$

$$a^{(0)} = \frac{1}{2}a^{(1)} = a^{(2)} = \left(\frac{s}{1-s}\right)^2 \quad (80)$$

$$b^{(1)} = 2\left(\frac{1-s}{1+s}\right) \quad (81)$$

and

$$b^{(2)} = -\left(\frac{1-s}{1+s}\right)^2 \quad (82)$$

The upper panel of [Fig. 21](#) shows the input and filtered waves for a 12-hour input period, a cutoff frequency ($2\pi/\nu_c$) of three hours and a timestep of 15 minutes. The filter causes little reduction in amplitude of this wave, but introduces a delay or phase-lag of a little more than half an hour. A very similar delay can be seen in the lower panel of [Fig. 21](#) for waves with longer periods.

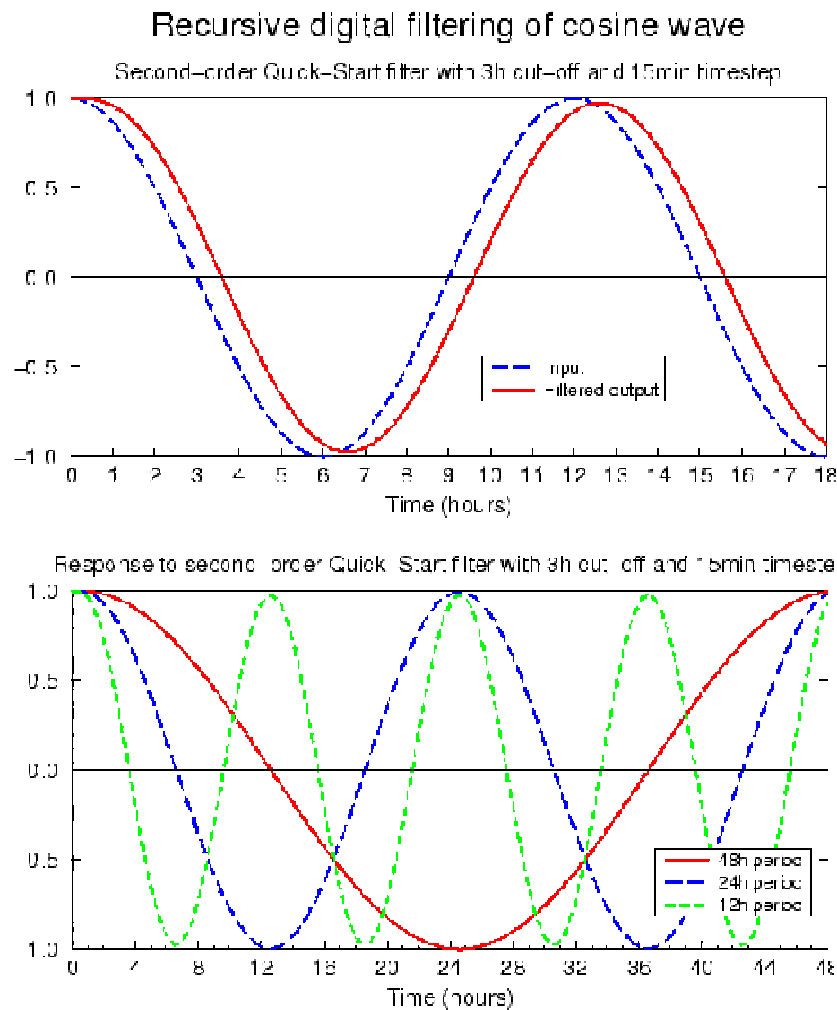


Figure 21. The input wave and filtered output wave for 12-hour input period and a second-order recursive filter (upper) and the filtered output for 12-, 24- and 48-hour input periods (lower).

The filtered response to input waves of periods 12, 3, 1 and 0.5 hours are shown in Fig. 22, for 15- and 1-minute timesteps. The amplitude of the wave with 3-hour period is reduced by about 30%, and it too is delayed by a little over half an hour. Waves with shorter periods are damped considerably over the first hour or so. That with half-hour period is soon damped completely in the case of the 15-minute timestep.

The rapid initial damping of short-period waves and the uniformity of the phase-lags for the longer-period waves means that an initialized forecast may be successfully launched by applying the second-order filter for a span of an hour or two, and then setting the model's clock back by half an hour or so to account for the delay, before extending the forecast from the end-point of the filtered sequence. The effect of applying this procedure on the level of noise in a forecast using the HIRLAM system can be seen by comparing the dashed and solid curves in Fig. 23. The procedure is evidently successful, and is sufficient for most forecasting purposes. It does not, however, provide initialized conditions at the analysis time or for the following hour or so, such as may be useful for diagnostic purposes or physical initialization. The alternative of simply assuming that the filtered value at the end of the span applies at the time of the uninitialized analysis is shown by the dotted curve in Fig. 23. This introduces a phase shift of about an hour, the effect of which can be reduced by adopting the incremental approach presented in 2.5.

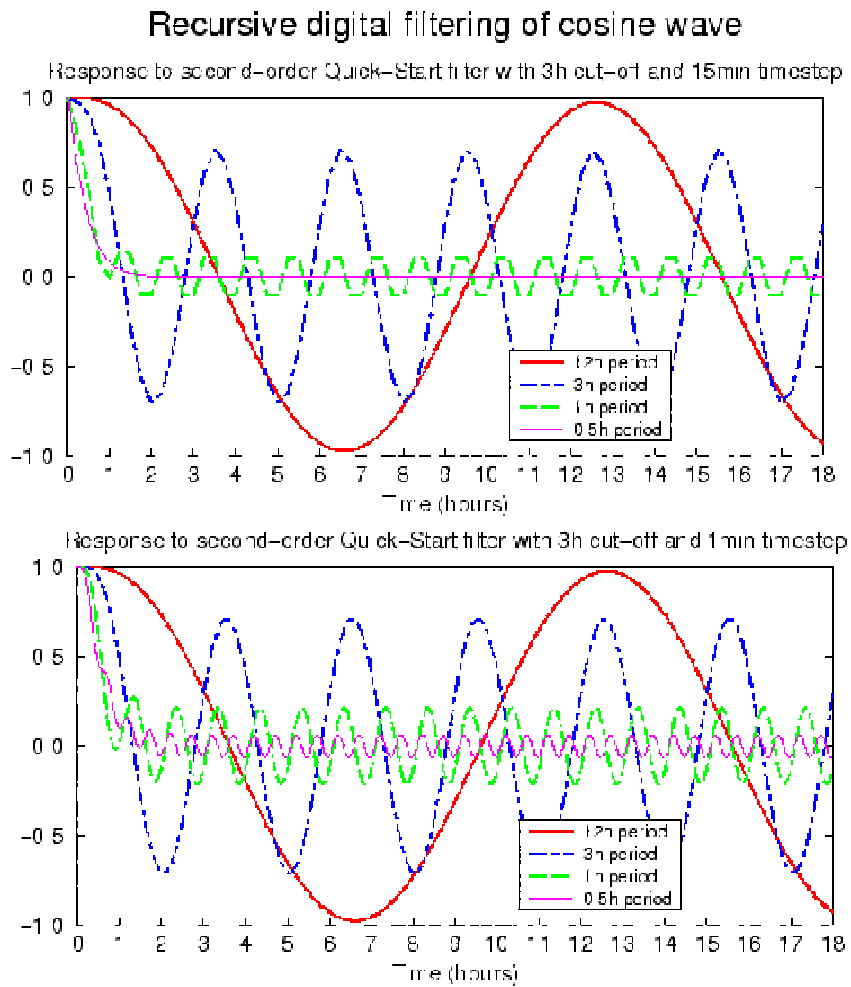


Figure 22. The filtered output wave for 12-, 3-, 1- and 0.5-hour input periods and a second-order recursive filter with 15-minute timestep (upper) and 1-minute timestep (lower).

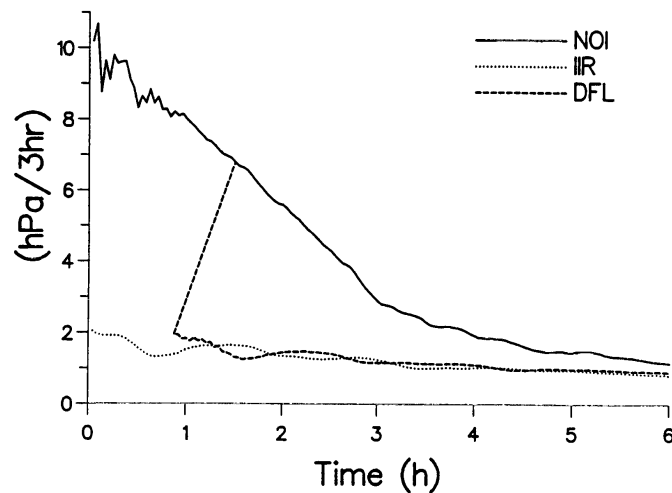


Figure 23. Evolution of the mean absolute surface pressure tendency (hPa /3h) over six-hour forecasts starting from an uninitialized analysis (solid line) and from using two types of recursive digital filtering schemes, from Lynch and Huang(1994).

4.3 Diabatic, non-recursive filtering

Another approach to diabatic initialization is to carry out an initial backward adiabatic integration over a time interval T followed by a forward diabatic integration over an interval $2T$, and then to apply non-recursive filtering to the sequence of values from the diabatic integration. Lynch *et al.*(1997) describe a particularly efficient variation of this approach that has been used to initialize both limited-area and (in incremental form) global models at Météo-France.

The principal cost of digital filtering initialization is that of the model integration. Improved efficiency arises partly from using a Dolph-Chebyshev filter (Lynch, 1997) that requires a shorter span to achieve the same degree of noise reduction as the Lanczos filter. A further gain comes from filtering the results of the backward adiabatic integration. This yields a filtered value at a time $T/2$ prior to the analysis time, and this provides initial conditions for a diabatic forecast over an interval T . The initialized analysis is derived from a non-recursive filtering of this diabatic forecast.

Fig. 24 shows that similar levels of noise reduction are obtained by using the Lanczos and (shorter-span) Dolph-Chebyshev filters without filtering the backward adiabatic forecast, and by using the Dolph-Chebyshev filter on both the backward adiabatic and (half-length) forward diabatic integrations.

A disadvantage of this general approach to diabatic initialization is that the initialized fields are subject to error due to changes brought about by diabatic processes over the first half of the forward integration that is filtered. This error is reduced by using filters that need a shorter span to be effective, and by the filtering of the backward integration, which enables the length of the diabatic forecast to be halved.

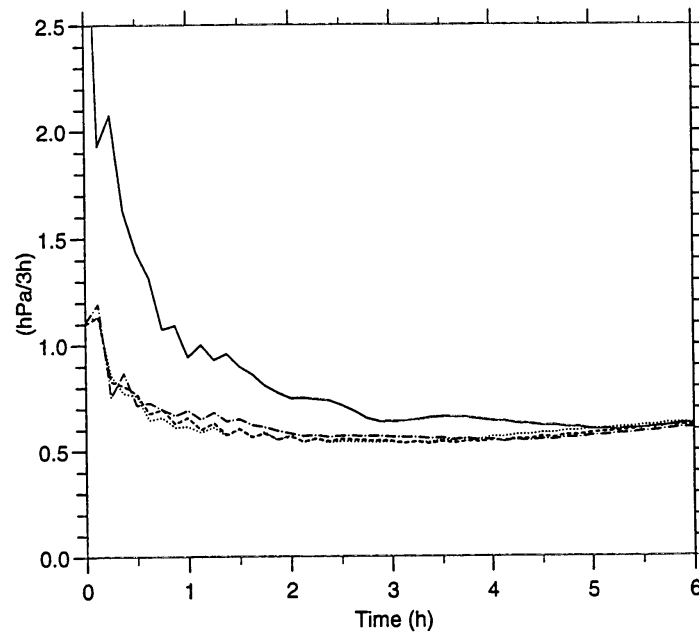


Figure 24. Evolution of the mean absolute surface pressure tendency (hPa /3h) for the first six hours of a forecast starting from an uninitialized analysis (solid line) and from analyses initialized using three non-recursive diabatic digital filtering schemes, from Lynch *et al.*(1997).

APPENDIX A . DEFINITION OF OPERATORS FOR THE ECMWF VERTICAL FINITE-DIFFERENCE SCHEME

Specification of the forms for the matrices γ and τ and the vectors δ and μ has been given by [Simmons](#) and [Burrige\(1981\)](#), although with a notation different to that used here. We present the general case of a reference temperature that varies with pressure, with values T_{rl} at the “full” levels of the model, for $l = 1, 2, \dots, L$. We assume a hybrid vertical coordinate in which the pressure at the “half” levels, $p_{l+1/2}$, is defined as a function of the surface pressure, p_s , with $p_{1/2} = 0$ and $p_{L+1/2} = p_s$.

The forms are:

$$\gamma_{kl} = \begin{cases} 0 & l < k \\ R\alpha_l & l = k \\ R \ln\left(\frac{p_{l+1/2}}{p_{l-1/2}}\right) & l > k \end{cases} \quad (\text{A.1})$$

$$\tau_{kl} = \frac{\Delta p_l}{\Delta p_k} \kappa T_{rk} \gamma_{lk} + \frac{1}{2\Delta p_k} \{S_{kl}(T_{r(k+1)} - T_{rk}) + S_{(k-1)l}(T_{rk} - T_{r(k-1)})\} \quad (\text{A.2})$$

$$\delta_l = \frac{\Delta p_l}{p_{sr}} \quad (\text{A.3})$$

and

$$\mu_l = \frac{RT_{rl}}{p_{l+1/2}} \left(p_s \frac{\partial p}{\partial p_s} \right)_{l+1/2} + \sum_{k=l+1}^L RT_{rk} \left(\frac{1}{p_{k+1/2}} \left(p_s \frac{\partial p}{\partial p_s} \right)_{k+1/2} - \frac{1}{p_{k-1/2}} \left(p_s \frac{\partial p}{\partial p_s} \right)_{k-1/2} \right) \quad (\text{A.4})$$

Here

$$\alpha_l = \begin{cases} 1 - \frac{p_{l-1/2}}{\Delta p_l} \ln\left(\frac{p_{l+1/2}}{p_{l-1/2}}\right) & l > 1 \\ \ln 2 & l = 1 \end{cases} \quad (\text{A.5})$$

and

$$S_{kl} = (\Delta p_l) \left(\frac{\partial p}{\partial p_s} \right)_{k+1/2} - \begin{cases} 0 & l > k \\ \Delta p_l & l \leq k \end{cases} \quad (\text{A.6})$$

All expressions involving a pressure are evaluated for the reference surface pressure p_{sr} .

If the reference temperature profile is isothermal, with $T_{rl} = T_r$, (A.2) reduces to

$$\tau_{kl} = \frac{\Delta p_l}{\Delta p_k} \kappa T_{rk} \gamma_{lk} \quad (\text{A.7})$$

and noting that

$$\frac{1}{p_{L+1/2}} \left(p_s \frac{\partial p}{\partial p_s} \right)_{L+1/2} = 1$$

(A.4) becomes simply

$$\mu_l = RT_r \tag{A.8}$$

We illustrate results by comparing modes computed for the idealized temperature profile shown in Fig. 25 with modes computed for isothermal reference profiles with temperatures of 245K and 300K. The idealized profile is constructed to be linear in $\ln p$ in regions representative of the troposphere and stratosphere. This distribution of temperature has a mean value of approximately 245K

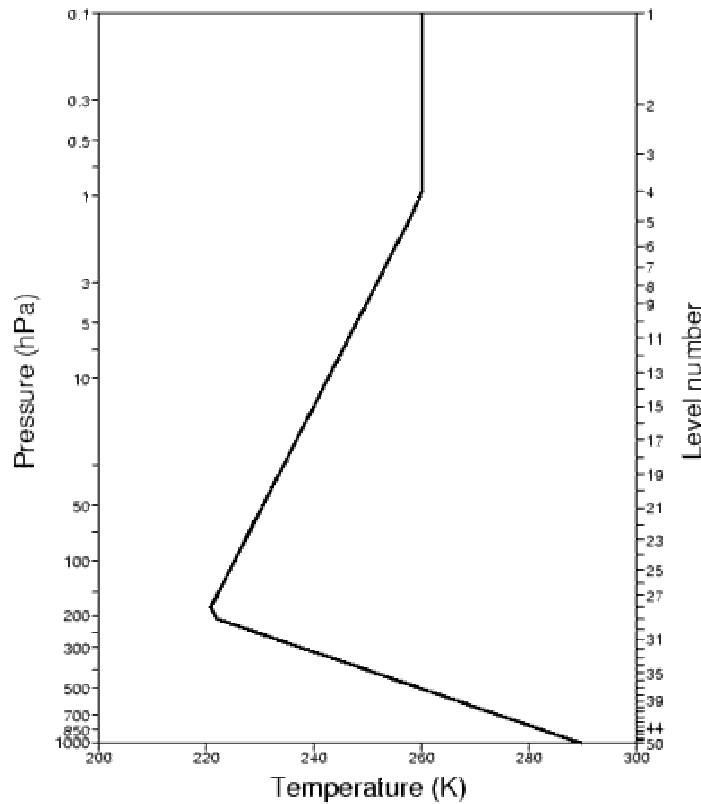


Figure 25. An idealized vertical profile of temperature.

Gravity-wave phase speeds for the first nine modes are presented in Table 2 and mode structures for the profile

TABLE 2. GRAVITY-WAVE PHASE SPEEDS FOR THE FIRST NINE MODES OF A 50-LEVEL VERSION OF THE ECMWF MODEL FOR THE REFERENCE TEMPERATURE PROFILE SHOWN IN Fig. 25 AND FOR TWO UNIFORM REFERENCE TEMPERATURES. THE REFERENCE SURFACE PRESSURE IS 1000HPA.

Mode number l	Phase speed (ms^{-1}) for reference temperature profile shown in Fig. 25	Phase speed (ms^{-1}) for 245K reference temperature	Phase speed (ms^{-1}) for 300K reference temperature
1	316	313	347

TABLE 2. GRAVITY-WAVE PHASE SPEEDS FOR THE FIRST NINE MODES OF A 50-LEVEL VERSION OF THE ECMWF MODEL FOR THE REFERENCE TEMPERATURE PROFILE SHOWN IN Fig. 25 AND FOR TWO UNIFORM REFERENCE TEMPERATURES. THE REFERENCE SURFACE PRESSURE IS 1000hPA.

Mode number l	Phase speed (ms^{-1}) for reference temperature profile shown in Fig. 25	Phase speed (ms^{-1}) for 245K reference temperature	Phase speed (ms^{-1}) for 300K reference temperature
2	249	237	262
3	180	175	194
4	131	131	145
5	100	102	112
6	78	82	90
7	64	67	75
8	54	57	63
9	47	49	54

shown in Fig. 25 and for the 245K reference temperature are shown in Fig. 26. The corresponding mode structures for the 300K reference temperature can be seen in the right-hand panel of Fig. 10. The gravity-wave phase speeds computed for the mean temperature of 245K are quite similar (within about 5%) to those computed for the reference state with varying tropospheric and stratospheric temperatures. Mode structures are qualitatively similar, but evidently reflect the differences in static stability in the troposphere and stratosphere. Phase speeds for the 245K reference temperature are smaller than those for the 300K reference by a factor equal to the square root of the ratio of the temperatures, and mode structures are identical for the two isothermal reference states. This is because the uniform reference temperature T_r appears only as a simple factor multiplying each element of the matrix \mathbf{G} defined by equation (28)

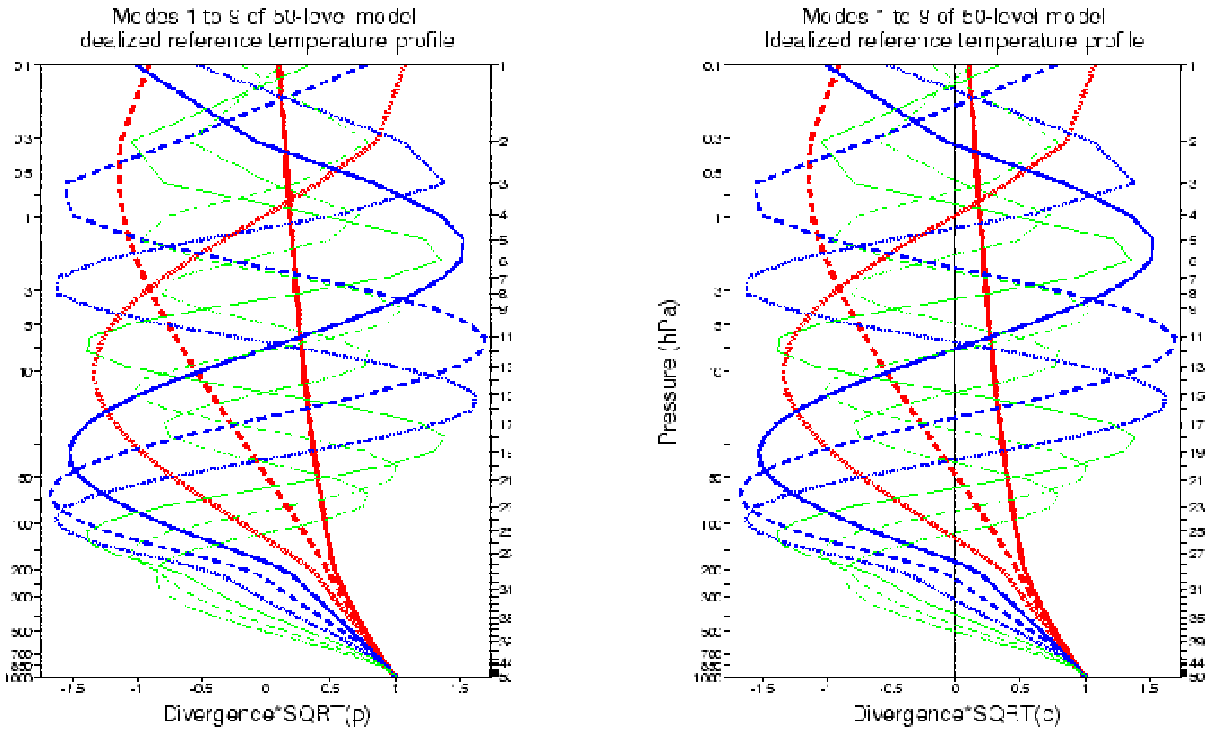


Figure 26. Structures of the first nine vertical modes of the 50-level model for the reference temperature profile shown in Fig. 25 and for a uniform reference temperature of 245K. The reference surface pressure is 1000hPa.

APPENDIX B . THE LAMB WAVE

The vertical structure and equivalent depth of the Lamb wave can be specified analytically. A temperature profile of the form:

$$T^* = -\alpha T_r \left(\frac{p}{p_{sr}} \right)^\alpha (\ln p_s)' \quad (\text{B.9})$$

enables the term $\phi' + RT_r(\ln p_s)'$ in (17) and (18) to be written:

$$\phi' + RT_r(\ln p_s)' = RT_r \left(\frac{p}{p_{sr}} \right)^\alpha (\ln p_s)' \quad (\text{B.10})$$

With D' also varying in the vertical as $(p/p_{sr})^\alpha$, the temperature and surface-pressure equations, (19) and (20), can both be satisfied only if $\alpha = -\kappa$. Thus, if we write:

$$RT_r(\ln p_s)' = \hat{P} \quad (\text{B.11})$$

$$u' = \left(\frac{p}{p_{sr}} \right)^{-\kappa} \hat{u} \quad (\text{B.12})$$

$$v' = \left(\frac{p}{p_{sr}}\right)^{-\kappa} \hat{v} \quad (\text{B.13})$$

and

$$D' = \left(\frac{p}{p_{sr}}\right)^{-\kappa} \hat{D} \quad (\text{B.14})$$

equations (17) to (20) reduce to the shallow water equations:

$$\frac{\partial \hat{u}}{\partial t} = 2\Omega \sin \theta \hat{v} - \frac{1}{a \cos \theta} \frac{\partial}{\partial \lambda} \hat{P} \quad (\text{B.15})$$

$$\frac{\partial \hat{v}}{\partial t} = -2\Omega \sin \theta \hat{u} - \frac{1}{a} \frac{\partial}{\partial \theta} \hat{P} \quad (\text{B.16})$$

$$\frac{\partial \hat{P}}{\partial t} = -\frac{RT_r}{1-\kappa} \hat{D} \quad (\text{B.17})$$

The phase speed of the plane wave in the absence of rotation is thus $\sqrt{(RT_r)/(1-\kappa)}$.

Since $\kappa = R/c_p$ and $c_p = c_v + R$, where c_v is the specific heat at constant volume, the phase speed may be written $\sqrt{(c_p/c_v)RT_r}$. This may be recognized as the speed of sound in a gas of temperature T_r .

With height, z , as the vertical coordinate, the vertical structure $(p/p_{sr})^{-\kappa}$ becomes $\exp((g\kappa z)/(RT_r))$. Wave energy density varies as $\exp(-gz(1-2\kappa)/(RT_r))$. Horizontal winds and temperature thus increase exponentially with increasing height, but wave energy density decreases away from the ground. The vertical velocity vanishes identically. This may be seen by writing, to first order in wave amplitude:

$$z = z_0(p) + z' \quad (\text{B.18})$$

The vertical velocity, w , is then given to this order by:

$$w = \frac{Dz}{Dt} = \frac{\partial z'}{\partial t} + \omega' \frac{dz_0}{dp} \quad (\text{B.19})$$

z_0 and z' are given by:

$$z_0 = \frac{RT_r}{g} \ln\left(\frac{p_{sr}}{p}\right) \quad (\text{B.20})$$

and

$$z' = \frac{1}{g}(\phi' + RT_r(\ln p_s)') \quad (\text{B.21})$$

Using (B.9) and (B.10), (B.21) becomes:

$$z' = \frac{RT'}{g\kappa} \quad (\text{B.22})$$

and substituting (B.20) and (B.22) into (B.19):

$$w = \frac{R}{g\kappa} \frac{\partial T'}{\partial t} - \frac{\omega' RT_r}{gp} \quad (\text{B.23})$$

The right-hand side of (B.23) is equal to zero by virtue of equation (19).

APPENDIX C . THE NON-RECURSIVE IMPLEMENTATION OF THE RECURSIVE FILTER

Consider the recursive filter defined in general by equation (73):

$$\mathbf{y}^{(n)} = \sum_{j=0}^N \mathbf{a}^{(j)} \mathbf{x}^{(n-j)} + \sum_{j=1}^N \mathbf{b}^{(j)} \mathbf{y}^{(n-j)}$$

Define a vector $\mathbf{a}^{(n)}$ of dimension $(n+1)$ by

$$\mathbf{a}^{(n)} = (0, 0, \dots, 0, \mathbf{a}^{(N)}, \mathbf{a}^{(N-1)}, \dots, \mathbf{a}^{(1)}) \quad (\text{C.24})$$

and define $\mathbf{b}^{(n)}$ similarly.

Let $\mathbf{X}^{(n)}$ and $\mathbf{Y}^{(n)}$ denote matrices whose rows are given by $(\mathbf{x}^{(j)})^T$ and $(\mathbf{y}^{(j)})^T$ for $j = 0, 1, 2, \dots, n$.

The non-recursive form of the recursive filter is then:

$$\mathbf{Y}^{(n)} = \mathbf{F}^{(n)} \mathbf{X}^{(n)} \quad (\text{C.25})$$

and $\mathbf{F}^{(n)}$ can be constructed inductively:

$$\mathbf{F}^{(n)} = \begin{bmatrix} \mathbf{F}^{(n-1)} & \mathbf{0} \\ \mathbf{a}^{(n-1)} + \mathbf{b}^{(n-1)} \mathbf{F}^{(n-1)} & \mathbf{a}^{(0)} \end{bmatrix} \quad (\text{C.26})$$

REFERENCES

- Ballish, B., Cao, X., Kalnay, E., and Kanamitsu, M. (1992) Incremental nonlinear normal mode initialization. *Mon. Wea. Rev.*, **120**, 1723-1734.
- Bouttier, F., Derber, J., and Fisher, M. (1997) The 1997 revision of the J_b term in 3D/4D-Var. *ECMWF Tech. Memo.*, **238**, 54pp.
- Chapman, S. and Lindzen, R.S. (1970) Atmospheric Tides. *D. Reidel, Holland*, 200pp.
- Courtier, P., Andersson, E., Heckley, W., Pailleux, J., Vasiljevic, D., Hamrud, M., Hollingsworth, A., Rabier, F., and Fisher, M. (1998) The ECMWF implementation of three dimensional variational assimilation (3D-Var). Part I: Formulation. *Q. J. R. Meteorol. Soc.*, **124**, in press.



- Kasahara, A. (1976) Normal modes of ultra-long waves in the atmosphere. *Mon. Wea. Rev.*, **104**, 669-690.
- Lilly, D.K. and Kennedy, P.J. (1973) Observations of a stationary mountain wave and its associated momentum flux and energy dissipation. *J. Atmos. Sci.*, **30**, 1135-1152.
- Lindzen, R.S., Batten, E.S. and Kim, J.W. (1968) Oscillations in atmospheres with tops. *Mon. Wea. Rev.*, **96**, 133-140.
- Longuet-Higgins, M.S. (1968) The eigenfunctions of Laplace's tidal equations over a sphere. *Phil. Trans. Roy. Soc. London*, **A262**, 511-607.
- Lynch, P. and Huang, X.-Y. (1992) Initialization of the HIRLAM model using a digital filter. *Mon. Wea. Rev.*, **120**, 1019-1034.
- Lynch, P. and Huang, X.-Y. (1994) Diabatic initialization using recursive filters. *Tellus*, **46A**, 583-597.
- Lynch, P. (1997) The Dolph-Chebyshev window: A simple optimal filter. *Mon. Wea. Rev.*, **125**, 655-660.
- Lynch, P., Giard, D., and Ivanovici, V. (1997) Improving the efficiency of a digital filtering scheme for diabatic initialization. *Mon. Wea. Rev.*, **125**, 1976-1982.
- Machenhauer, B. (1977) On the dynamics of gravity oscillations in a shallow water model, with application to normal-mode initialization. *Contrib. Atmos. Phys.*, **50**, 253-271.
- Ninomiya, K. (1983) Internal-gravity-wave-like variations of temperature, humidity and wind observed in the troposphere downstream of heavy rainfall. *J. Meteorol. Soc. Japan*, **61**, 163-169.
- Puri, K., Bourke, W., and Seaman, R. (1982) Incremental linear normal mode initialization in four-dimensional data assimilation. *Mon. Wea. Rev.*, **110**, 1773-1785.
- Simmons, A.J., and Burridge, D.M. (1981) An energy and angular-momentum conserving vertical finite-difference scheme and hybrid vertical coordinates. *Mon. Wea. Rev.*, **109**, 758-766.
- Taylor, G.I. (1929) Waves and tides in the atmosphere. *Proc. Roy. Soc.*, **A126**, 169-183.
- Temperton, C. (1988) Implicit normal-mode initialization. *Mon. Wea. Rev.*, **116**, 1014-1031.
- Temperton, C. (1989) Implicit normal-mode initialization for spectral models. *Mon. Wea. Rev.*, **117**, 436-451.
- Temperton, C., and Williamson, D.L. (1981) Normal mode initialization for a multi-level gridpoint model. Part I: Linear aspects. *Mon. Wea. Rev.*, **109**, 729-743.
- Webster, P.J. (1983) Large-scale structure of the tropical atmosphere. *Large-scale Dynamical Processes in the Atmosphere*, Eds. B.J. Hoskins and R.P. Pearce, Academic Press, 235-275.
- Wergen, W. (1987) Diabatic nonlinear normal mode initialization for a spectral model with a hybrid vertical coordinate. *ECMWF Technical Report No. 59*, 83pp.
- Williamson, D.L., and Temperton, C. (1981) Normal mode initialization for a multi-level gridpoint model. Part II: Nonlinear aspects. *Mon. Wea. Rev.*, **109**, 744-757.