A multi-layer spectral model and the semi-implicit method

By B. J. HOSKINS and A. J. SIMMONS

U.K. Universities Atmospheric Modelling Group, Department of Geophysics, University of Reading

(Received 30 August 1974; revised 27 November 1974)

SUMMARY

The formulaton of a multi-layer primitive equation model on the sphere is described. The horizontal representation is by means of spherical harmonics, truncated either in the triangular or rhomboidal manner. The time integration is performed using the semi-implicit method in which the linearized gravity wave terms are time averaged and thus the fast moving waves of this type are slowed. For a 5-layer hemispheric model with triangular truncation at wavenumber 21, storage of 38K words is needed and with the time scheme allowing a time-step of 90 minutes, one day's simulation requires 11 seconds of CDC 7600 time. The growth of a baroclinic wave on a simple basic state of differential solid body rotation is exhibited. The errors involved in this case in utilizing the large time-step allowed by the semi-implicit scheme are thoroughly examined by comparing wave amplitudes and phases, conservation properties and gravity wave treatment for different time-steps. These errors are found to be negligible. The conservation properties of the model are in fact extremely good. The vertical finite differencing scheme of Arakawa is studied in the same baroclinic instability simulation. The growth is similar though the conservation of angular momentum is greatly improved. The transform method used in all these integrations allows some aliasing, but this is shown to be negligible.

1. INTRODUCTION

In the last few years there has been increased interest in the use of truncated series of spherical harmonics to represent the horizontal variation of functions on the sphere in numerical models. This mathematically more attractive approach has been a competitor with the usual finite difference methods only since the introduction of the transform method (Orszag 1970; Eliasen *et al.* 1970). In this approach, products are performed by transforming to a suitable grid of points, performing the product and transforming back to wave space. Since this time, shallow water equation models have been developed, notably by Bourke (1972) and Machenauer and Rasmussen (1972).

The value of this spectral approach in suggesting a mathematical analysis of the dynamics of Rossby wave instability was shown in Hoskins (1973). This problem also provided a convenient test bed for the comparison of the spectral method with the usual second order finite difference method (Doron *et al.* 1974).

The development to a full primitive equation model on the sphere has taken two directions. Machenauer and Daley (1972) formulated a model with spectral representation in the vertical. In common with Bourke (1974) the present authors have taken the, perhaps, more cautious approach of using standard finite differences in the vertical. The model described in this paper is a natural extension to a multi-layer model of the one-layer model of Bourke (1972). Though developed independently, it is scarcely surprising that it is quite similar to his later model. To anticipate later sections, we note that the main difference is the use of an energy conserving scheme in the vertical.

A simultaneous development in the atmospheric modelling scene has been that of the semi-implicit integration method (Robert *et al.* 1972). One of the major disadvantages of the use of the primitive equations for numerical models has been the restriction of the time-step by the extremely high frequencies associated with the short gravity waves. The time-step limitation has certainly been overcome by the introduction of this method, but it has not been clear what price has been paid in accuracy of simulation and conservation.

The incorporation of the semi-implicit method in a spectral model involves very little extra work and is done in the model described here. To incorporate it in a finite difference model, however, necessitates a large increase in computation per time-step or the use of a splitting method (Burridge 1974) which has the disadvantage that it is only first order accurate in time. In this paper we attempt to evaluate the errors associated with a large semi-implicit time-step by a thorough analysis of integrations with different time-steps for the growth of a simple baroclinic wave.

The basic model is described in section 2, gravity waves and the semi-implicit scheme in section 3 and a basic model run in section 4. This is a baroclinic instability simulation starting from a deliberately very simple initial state of differential solid-body rotation. The study of the errors introduced by the semi-implicit method is made in section 5. Section 6 gives a brief account of the effect of using a different finite difference scheme in the vertical and the angular momentum conservation obtained. Finally, section 7 mentions some experiments which show that a possible source of error in the model is not important here. The reader who is not concerned with the details of the numerical model may wish to omit sections 2 and 3, and move directly to the simulation described in section 4.

Apart from the introduction of the model and analysis of the semi-implicit method, one of the points of this paper is to introduce some of the new ways of looking at the information provided by numerical models when integrated using the spectral method. In the authors' opinion this may be one of the most important uses for the method in its provocation of new dynamical insights into the content of models from the simplest right up to the general circulation models. It is intended that these displays will form the basis of subsequent papers using the spectral technique.

2. THE BASIC MODEL

The non-dimensional equations of motion for an inviscid, adiabatic, hydrostatic, perfect gas surrounding a rotating, spherical planet may be written

$$\frac{\partial \zeta}{\partial t} = \frac{1}{1 - \mu^2} \frac{\partial}{\partial \lambda} \mathscr{F}_V - \frac{\partial}{\partial \mu} \mathscr{F}_U, \qquad (1)$$

$$\frac{\partial D}{\partial t} = \frac{1}{1-\mu^2} \frac{\partial}{\partial \lambda} \mathscr{F}_U + \frac{\partial}{\partial \mu} \mathscr{F}_V - \nabla^2 \left(\frac{U^2 + V^2}{2(1-\mu^2)} + \phi + T \ln p_* \right), \quad (2)$$

$$\frac{\partial T'}{\partial t} = -\frac{1}{1-\mu^2} \frac{\partial}{\partial \lambda} (UT') - \frac{\partial}{\partial \mu} (VT') + DT' - \dot{\sigma} \frac{\partial T}{\partial \sigma} + \kappa \frac{T\omega}{p}, \quad . \tag{3}$$

$$\frac{\partial \ln p_*}{\partial t} = -\mathbf{V} \cdot \nabla \ln p_* - D - \frac{\partial \dot{\sigma}}{\partial \sigma}, \qquad (4)$$

Here $\mathscr{F}_{U} = V\zeta - \dot{\sigma}\frac{\partial U}{\partial\sigma} - T'\frac{\partial \ln p_{*}}{\partial\lambda}, \quad \mathscr{F}_{V} = -U\zeta - \dot{\sigma}\frac{\partial V}{\partial\sigma} - T'(1-\mu^{2})\frac{\partial \ln p_{*}}{\partial\mu}.$

 ζ is the absolute vorticity, D the divergence, $T = \overline{T}(\sigma) + T'$ the temperature, p_* the surface pressure, and ϕ the geopotential, all non-dimensionalized using as length scale the radius of the planet, a; as time scale the reciprocal of its angular velocity, Ω^{-1} ; temperature scale $a^2\Omega^2/R$ (R the gas constant); and pressure scale $p_0 = 1000$ mb. σ is the vertical co-ordinate pressure/ p_* ; λ is the longitude; and $\mu = \sin \theta$, where θ is the latitude. U and V are the

longitudinal and latitudinal velocities multiplied by $\cos \theta$. The horizontal advection operator is

$$\mathbf{V} \cdot \nabla = \frac{U}{1-\mu^2} \frac{\partial}{\partial \lambda} + V \frac{\partial}{\partial \mu}$$

A velocity potential α and streamfunction ψ may be introduced giving

$$U = -(1 - \mu^2) \frac{\partial \psi}{\partial \mu} + \frac{\partial \alpha}{\partial \lambda}, \quad V = \frac{\partial \psi}{\partial \lambda} + (1 - \mu^2) \frac{\partial \alpha}{\partial \mu},$$

$$\zeta = 2\mu + \frac{1}{1 - \mu^2} \frac{\partial V}{\partial \lambda} - \frac{\partial U}{\partial \mu} = 2\mu + \nabla^2 \psi,$$

$$D = \frac{1}{1 - \mu^2} \frac{\partial U}{\partial \lambda} + \frac{\partial V}{\partial \mu} = \nabla^2 \alpha.$$

The use of the horizontal equations of motion in their vorticity and divergence form follows Bourke (1972). It is very convenient for the application of the spectral and semi-implicit time scheme techniques. Interpretation of results is also aided by the direct output of these dynamically meaningful quantities.

In this model we use a spectral representation in the horizontal and a finite difference representation in the vertical. The variables ζ , D, T', ϕ at each layer and the variable $\ln p_*$ are represented by truncated series of spherical harmonics:

$$X = \Sigma X_n^m P_n^m(\mu) e^{im\lambda}.$$

The series for ζ and D directly imply representations of U and V. The usual truncation since the work of Ellsaesser (1966) has been rhomboidal: $|m| \leq M$, $n - |m| \leq J$. In some senses, this can be thought of as giving the same resolution for each zonal wavenumber. An alternative is the triangular truncation $n \leq M$. This allows a representation independent of position (or direction) on the sphere. In this paper, only integrations with the triangular truncation will be described.* The programme is written so that only minor changes are necessary to convert to the alternative truncation. Many integrations have been made with both methods and a detailed comparison has been made in Simmons and Hoskins (1975).



Figure 1. The arrangement of model variables in the vertical for the case of N equally spaced layers.

The programme is also flexible in the number, N, and spacing of the layers in the vertical, though the integrations described here will be for five equally spaced layers. The

[•] On the hemisphere, for programming purposes it was found to be convenient to use the 'jagged triangular truncation': for basic variables even about the equator: $n \le M - 1$,

for basic variables odd about the equator: $n \leq M$.

arrangement of variables in the vertical is exhibited in Fig. 1. The finite difference analogue of $\dot{\sigma}\partial X/\partial\sigma$ at $\sigma = \sigma_r$ is taken as $\overline{\dot{\sigma}\delta_{\sigma}X^{\sigma}}$, using the usual notation. This is formally second order accurate if the $\frac{1}{2}$ -levels are defined as the values at equally spaced intervals of a differentiable function (see Smagorinsky *et al.* 1965). The finite difference analogues of the integral of Eq. (4) from $\sigma = 0$ to 1 and $\sigma = 0$ to σ_{r+1} give

$$\frac{\partial}{\partial t} \ln p_{*} = -\sum_{r=1}^{N} A_{r} \Delta \sigma_{r},$$
$$\dot{\sigma}_{r+\frac{1}{2}} = (\sum_{s=1}^{r} \Delta \sigma_{s}) \sum_{1}^{N} A_{s} \Delta \sigma_{s} - \sum_{s=1}^{r} A_{s} \Delta \sigma_{s},$$

where $A = D + V \cdot \nabla \ln p_*$. The hydrostatic equation becomes

$$\phi_r - \phi_{r+1} = \alpha_r (T_r + T_{r+1}),$$
 . . . (6)

where $\alpha_r = \frac{1}{2} \ln (\sigma_{r+1}/\sigma_r)$. Assuming an isothermal lowest layer,

$$\phi_N - \phi_* = \alpha_N T_N, \qquad . \qquad . \qquad . \qquad (7)$$

where $\alpha_N = \ln (1/\sigma_N)$, and ϕ_* is the geopotential at the surface $\sigma = 1$.

Following Corby et al. (1972), the conversion term in the thermodynamic equation

$$\frac{\kappa T \omega}{p} = \kappa T \left[\mathbf{V} \cdot \nabla \ln p_* - \frac{1}{\sigma} \int_0^{\sigma} A d\sigma \right]$$

is represented by

$$\kappa T_r \left[\mathbf{V}_r \cdot \nabla \ln p_* - \frac{1}{\Delta \sigma_r} \left(\alpha_r \sum_{j=1}^r A_j \Delta \sigma_j + \alpha_{r-1} \sum_{j=1}^{r-1} A_j \Delta \sigma_j \right) \right].$$

The finite difference equivalents of

$$\mathbf{V} \cdot \dot{\sigma} \frac{\partial}{\partial \sigma} \mathbf{V} + \frac{1}{2} \mathbf{V}^2 \frac{\partial \dot{\sigma}}{\partial \sigma} = \frac{\partial}{\partial \sigma} (\frac{1}{2} \mathbf{V}^2 \dot{\sigma}),$$
$$\frac{\partial \dot{\sigma}}{\partial \sigma} T + \dot{\sigma} \frac{\partial T}{\partial \sigma} = \frac{\partial}{\partial \sigma} (\dot{\sigma} T),$$
$$\int_0^1 (\phi - \phi_*) A d\sigma = \int_0^1 \frac{T}{\sigma} \left(\int_0^{\sigma} A d\sigma' \right) d\sigma,$$

are all maintained by the above vertical scheme, hereinafter referred to as the T scheme. As indicated by Corby *et al.*, this implies that the total energy

$$\iint dS \, \frac{p_*}{g} \sum_{r=1}^N \left(\phi_* + \frac{1}{\kappa} T_r + \frac{1}{2} \mathbf{V}_r^2 \right) \Delta \sigma_r$$

is formally conserved by the T scheme. Mass is also formally conserved by this scheme. Thus changes in these two quantities will be caused only by time differencing and by the horizontal space scheme. We note that since the T scheme equivalent of

$$\int_0^1 T d\sigma = \int_0^1 (\phi - \phi_*) d\sigma$$

is not maintained, angular momentum is not formally conserved.

The equations are integrated using the transform method in a manner identical with that of Bourke (1974). Spectral expansions for ζ , D, T', and $\ln p_{\star}$ and the implied ones

for $U, V, \partial \ln p_*/\partial \lambda$ and $(1 - \mu^2) \partial \ln p_*/\partial \mu$ are used to give grid point values of these quantities. The nonlinear terms in Eqs. (1)-(4) are then evaluated and the spectral tendencies computed by the inverse transformation utilizing integration by parts. It may be shown that this procedure is exact for the retained modes provided that the grid used is one defined by enough equally spaced lines of longitude and enough Gaussian angles of latitude (Eliasen *et al.* 1970). For the quadratic terms the number of longitudes, M_g , and latitudes J_g , must satisfy

> rhomboidal: $M_g \ge 3M + 1$, $J_g \ge (2M + 3J + 1)/2$ triangular: $M_g \ge 3M + 1$, $J_g \ge (3M + 1)/2$.

There are terms for which this grid is insufficient for removing aliased interactions. These terms are the triple correlation involved in the energy conversion term and in the vertical advection terms. These require in theory, $M_g \ge 4M + 1$ and $J_g \ge (4M + 4J + 1)/2$ for rhomboidal and $J_g \ge (4M + 1)/2$ for triangular truncation. In most of this paper aliasing will be allowed in these terms. This point will be returned to briefly in section 7.

3. GRAVITY WAVES AND THE SEMI-IMPLICIT SCHEME IN A MULTI-LAYER MODEL

Quite generally, any vertical difference scheme for the model shown in Fig. 1 gives finite difference analogues of Eqs. (2)-(5) of the form

$$\partial \ln p_*/\partial t = \mathscr{P} - \pi \mathbf{D}, \qquad . \qquad . \qquad . \qquad . \qquad (10)$$

Here the variables **D**, **T'**, **T**, \mathbf{T} , $\phi - \phi_*$ with values at each layer are represented by column vectors, as are also \mathcal{T} and \mathcal{D} . **g** and τ are constant matrices, \mathcal{P} and $\ln p_*$ are scalars and π is a row vector. The separation in Eq. (9) is such that the matrix τ contains all the references to the basic temperatures \mathbf{T} which multiply the divergence to produce the temperature tendency. In Eq. (10) π is a constant vector and this term contains all the direct effect of divergence on the surface pressure tendency. In appendix I we give the matrices and vectors for the T scheme.

With continuous time representation, elimination of T, ϕ and $\ln p_*$ from Eqs. (8)-(11) gives an equation for D:

$$\left(\frac{\partial^2}{\partial t^2} - \mathbf{B}\nabla^2\right)\mathbf{D} = \frac{\partial\mathcal{D}}{\partial t} - \nabla^2(\mathbf{g}\mathcal{F} + \mathbf{T}\mathcal{P}). \qquad (12)$$

Here the matrix $\mathbf{B} = \mathbf{g\tau} + \mathbf{T}\pi$. With the right-hand side set to zero, this is the gravity wave equation for a multi-layer model. The eigenvalues of **B** are the squares of the eigenvelocities associated with the vertical structure which is the corresponding eigenmode of **B**. (The gravity wave eigenmodes for a standard case are exhibited in Fig. 4.) The right-hand side may be thought of as the generation of gravity waves.

We now consider a representation at discrete times and use the standard notation $\delta_t X = (X^{t+\Delta t} - X^{t-\Delta t})/2\Delta t$, $\overline{X}^t = \frac{1}{2}(X^{t+\Delta t} + X^{t-\Delta t})$. In a manner similar to that of Robert *et al.* (1972) we use an averaging in time for the gravity wave terms separated above. This semi-implicit scheme gives

$$\delta_t \mathbf{D} = \mathscr{D} - \nabla^2 (\bar{\boldsymbol{\phi}}^t + \mathbf{\tilde{T}} \overline{\ln p}_*^t), \quad . \qquad . \qquad (13)$$

$$\delta_t \mathbf{T} = \mathscr{F} - \tau \overline{\mathbf{D}}^t, \qquad (14)$$

$$\delta_t \ln p_* = \mathscr{P} - \pi \overline{\mathbf{D}}^t. \qquad . \qquad . \qquad . \qquad (15)$$

Using the relation

$$\overline{X}^t = X^{t-\Delta t} + \Delta t \delta_t X,$$

elimination of $\overline{\mathbf{T}}^t$, $\overline{\phi}^t$ and $\overline{\ln p}_*^t$ may be carried out as before. This gives the semi-implicit time scheme analogue of Eq. (12):

$$(\mathbf{I} - \mathbf{B}\Delta t^2 \nabla^2) \overline{\mathbf{D}}^t = \mathbf{D}^{t-\Delta t} + \Delta t [\mathcal{D} - \nabla^2 (\boldsymbol{\phi}^{t-\Delta t} + \mathbf{T} \ln p_{\star}^{t-\Delta t})] - -\Delta t^2 \nabla^2 (\mathbf{g} \mathcal{F} + \mathbf{T} \mathcal{P}), \qquad (16)$$

I being the unit matrix. With the spectral representation used here, this becomes the algebraic equation

$$\left(\frac{1}{c_n}\mathbf{I} + \mathbf{B}\Delta t^2\right)\overline{\mathbf{D}_n^{mt}} = \frac{1}{c_n}\mathbf{D}_n^{m(t-\Delta t)} + \Delta t \left[\frac{1}{c_n}\mathcal{D}_n^m + \boldsymbol{\phi}_n^{m(t-\Delta t)} + \mathbf{T}\ln p_{\ast n}^{m(t-\Delta t)}\right] + \Delta t^2(\mathbf{g}\mathcal{F}_n^m + \mathbf{T}\mathcal{P}_n^m), \qquad (17)$$

for each component \mathbf{D}_n^m . Here $c_n = n(n + 1)$.

If gravity waves are not amplified by the vertical scheme, then the matrix **B** has positive eigenvalues and so the matrix $\mathbf{A}_n = \mathbf{I}/c_n + \mathbf{B}\Delta t^2$ may be inverted. Since \mathbf{A}_n is a constant this inversion need be computed once only. The values of divergence, temperature and surface pressure at the new time step may then be obtained directly. The vorticity equation is approximated by the usual centred difference in time and direct solution is possible.

For a simple gravity wave mode composed of one spherical harmonic with no forcing, the continuous Eq. (12) gives the frequency relation $\sigma^2 = c_n v_r^2$, where v_r^2 is an eigenvalue of the matrix **B**. It is easily shown from Eqs. (13) to (15) that the semi-implicit scheme gives a frequency $\sigma_s = (\tan^{-1} \sigma \Delta t)/\Delta t$. For an explicit integration method, the time averages in Eqs. (13) to (15) are dropped and the equivalent frequency is $\sigma_e = (\sin^{-1} \sigma \Delta t)/\Delta t$. These approximations are shown in Fig. 2. For $\sigma \Delta t < 1$, the explicit method overestimates and



Figure 2. Approximations (σ_o) to the correct gravity wave frequency (σ) when using the explicit and semiimplicit time schemes with time-step Δt .

642

the semi-implicit method underestimates the frequency. For $\sigma \Delta t > 1$, the explicit method gives computational instability. The semi-implicit method always gives stability but as the correct frequency tends to infinity, that computed tends to $\pi/2\Delta t$.

The semi-implicit method described above was developed and tested independently but is almost identical with that described by Bourke (1974). However, here, the standard layer temperatures \mathbf{T} are considered constant in time so that the matrix \mathbf{B} and the matrix inversion are calculated once only. The time taken for one semi-implicit time step is of the order of 1% more than that for an explicit time step. The storage required is increased only by the amount needed for the inverse of \mathbf{A}_n .

4. THE GROWTH OF A BAROCLINIC WAVE

(a) The numerical model

For almost all the integrations described in this paper we use a primitive equation model with five equally spaced levels in the vertical and with a spectral representation in the horizontal. The truncation is triangular at total wavenumber 21 (hereinafter referred to as triangular 21). For a hemispheric integration, the number of degrees of freedom for each variable at each layer is 231. The total storage required for the programme and arrays is 38K words, and the time taken for one time step, including the calculation of energy, total mass, and angular momentum diagnostics, on a CDC 7600 is 0.67 seconds. If the model is run with the maximum possible time-step of 90 minutes, the time taken for a one day integration on the hemisphere is under 11 seconds. The integrations described here were performed using only zonal wavenumbers 0, 8 and 16. The number of degrees of freedom is then 31, the time-step time without diagnostics 0.084 seconds. In this section the time-step used is $\frac{1}{2}h$.

(b) Initial state (numerical values are given in appendix II)

The average temperature, \overline{T} , chosen for each layer is shown in Fig. 3. There is some crude representation of a stratosphere by the increased static stability between the top



Figure 3. The initial average temperature in each layer and solid body rotation velocity at the equator.

layers. In Fig. 4 are exhibited the corresponding eigenmodes of the matrix **B**. They show the distribution of divergence with height in the gravity wave modes. The 'external' mode has no change of sign with height. The 'first internal' mode has one zero and has smaller phase speed. Each succeeding mode has one more zero and moves less quickly.

The basic velocity is chosen to be a very simple differential solid body rotation, with equatorial velocities shown in Fig. 3. This was chosen for a first experiment for reasons of simplicity and possible comparison with concurrent analytic baroclinic instability studies (Hollingsworth 1975). It emerged that the unstable modes are realistic despite the unrealistic nature of the velocity distribution.



Figure 4. The gravity wave mode divergence structures and the associated velocities for the five-layer model described in the text. The modes range from the fast ' external ' one with the same sign in divergence at all heights to the very slow ' fourth internal mode ' with four zeros.

A very small initial perturbation of non-dimensional amplitude 0.01 was added to the $(8, 9)^*$ vorticity coefficient with the same amplitude and phase at all heights. This corresponds to a maximum wind perturbation of 1 m s^{-1} .

The temperature variation at each layer and the surface pressure distribution are obtained by demanding that the initial divergence and divergence tendency be zero, and applying a smoothness condition. The art of initialization is not a major emphasis in this paper, but details of the method are given in appendix II. The balanced temperature field has an equator to pole temperature difference of -5K, 37K, 50K, 55K at the respective layers. The surface pressure difference is 2 mb. The wave velocity perturbation produces an extremely small temperature perturbation and a pressure perturbation of less than 1 mb.

(c) Initial time steps

We perform *n* initial time steps: a forward one of $\Delta t/2^{n-1}$ and centred ones of $\Delta t/2^{n-2}$, ..., Δt each from time zero before starting the basic time scheme (see Fig. 5). This method is based on a suggestion of Miyakoda (1960) and has been found an effective way of reducing the initial shock, especially when using the large time steps permitted by use of the semi-implicit method. Here we use n = 3 initial time steps.



Figure 5. The scheme for n = 3 initial time-steps. Step 1 is forward from t = 0 to $t = \Delta t/4$. Steps 2, 3 and 4 are centred from t = 0 to $\Delta t/2$, Δt and $2\Delta t$ respectively. The integration continues with standard centred steps.

(d) Results

The initial vorticity, divergence and temperature at level 5 ($\sigma = 0.9$) and the surface pressure are shown in Fig. 6. Since the perturbation is not a normal mode, it takes a few days for the developing baroclinic wave to become organized. At days 5, 6 and 7 (also shown in Fig. 6) the surface pressure minimum is 994mb, 988mb, 980mb respectively. At day 8 it is 963mb. Evidence will be given below that this integration contains significant error due to the horizontal truncation after day 6. The usual phase relations and development of the temperature and pressure waves are clear, and the model attempts to form some cold and warm frontal structure. The low level divergence gives an indication of the low level vertical velocity. Strong upward motion occurs in the warm air running from the incipient

[•] This standard notation means that the vorticity coefficient of the mode m = 8, n = 9 is equal to 0.01. Since the vorticity is a real function, it is implied that the coefficient of m = -8, n = 9 is also 0.01.



Figure 6. The development of a baroclinic wave. The top three rows show the relative vorticity, divergence and temperature fields respectively at level 5 ($\sigma = 0.9$). The bottom row shows the surface pressure. The columns are for days 0, 5, 6 and 7 respectively. The maps are polar stereographic projections with the pole at the lower left-hand corner. A background grid of 20° latitude by 20° longitude is indicated. Contours are drawn every 0.1 Ω , 0.02 Ω , 5K, 4 mb in the respective rows.

warm front region along the incipient cold front, and is a maximum ahead of the surface low. The vorticity and divergence, especially, show the strong SW-NE tilt of the mode, particularly in equatorial regions. The consequent feed of angular momentum is such as to form a strong jet in mid-latitudes despite there being no jet structure in the basic state. This is in agreement with the detailed analysis of Hollingsworth (1975). More detailed description of the limitations of this model due to its truncation and of the baroclinic wave structure is given in Simmons and Hoskins (1975).

Fig. 7 gives an indication of the vorticity spectrum at level 5 at days 6 and 7. Since the smallest scales retained, particularly in wavenumber 16, have significant amplitude, the probability of truncation error is apparent. The (8, 11) mode in the vorticity is dominant at all layers from days 4 to 6 apart from days 4 and 6 at level 1. Thus this mode as exhibited in Fig. 8 gives a good indication of the growth, movement and structure of the baroclinic wave. The usual westward tilt with height becoming somewhat smaller with time is apparent, as is also the slight reduction in amplitude with a marked decrease in the 'stratosphere' layer. This decrease in amplitude is much more dramatic in the temperature field. The mode accelerates from a movement of 5° longitude between days $1\frac{1}{2}$ to $2\frac{1}{2}$, to 8° longitude between days $5\frac{1}{2}$ to $6\frac{1}{2}$.



Figure 7. The vorticity spectrum at level 5 at day 6 (dashed) and day 7 (continuous). The waves have zonal wavenumber, m, 0, 8 and 16 and total wavenumber, n, from m up to 21. Only modes with n odd are nonzero for an integration on the hemisphere and only these are indicated. The value plotted is the square root of the contribution to the mean square vorticity. If surface pressure variation is small, then for equal energy in all modes $|\xi_n| = c_n^{\dagger} = [n(n+1)]^{\dagger}$. These points nearly lie on straight lines through $n = -\frac{1}{2}$.





Figure 8. The structure, development and movement of the (8, 11) vorticity mode. The horizontal bar is proportional to the amplitude of the mode at that layer and time. The centres of these bars indicate the positions of the maxima on the sphere (longitude indicated below the diagram).

5. TIME SCHEME EXPERIMENTS

In this section our aim is to investigate thoroughly the errors introduced by using the long time steps made possible by the semi-implicit method. To put the question in context, we note that the longest possible time step for the model described here when an explicit time scheme is used is 15 minutes. For the semi-implicit scheme, it is 90 minutes. As stated above, there is a negligible increase in computation and storage required. A comparison of a semi-implicit and explicit time integrations, both using 5 minute time steps, has shown that the differences are negligible compared with those referred to below. Therefore we compare only semi-implicit integrations for the case described in section 4 using 5 minute, 30 minute and 90 minute time steps. These runs will be referred to as \$5, \$30, \$90 respectively. They are started using 3, 3 and 5 initial time steps (see section 4(c)). The investigation concentrates on three aspects: vorticity wave amplitudes and phases; conservation properties; and treatment of gravity waves.

(a) Vorticity amplitudes and phases

In Fig. 9 are exhibited the amplitude and phase of the (8, 11) component of vorticity at the lowest layer. This component is the dominant one from day 2 to day 6. Only the S5 and S90 curves are drawn, the S30 waves nearly everywhere lying between the two. For comparison we include the curves obtained from an integration using a triangular 42



Figure 9. Comparisons of the (8, 11) vorticity mode amplitudes and phases in the S5 (continuous) and S90 (dotted) runs and in a run with double the resolution (dashed):
(a) amplitude: (b) phase plus 58° per day. Angles should be divided by 8 to obtain changes in longitude on the sphere.

truncation and a 30 minute time step. The amplitudes in the triangular 21 experiments are much closer to each other than any of them are to that of the triangular 42 experiment. The amplitudes are significantly different after day 6. In order to exhibit the very small phase differences a phase speed of 58° per day has been added to all the phases. Again the triangular 21 experiments are more similar to one another than to the triangular 42 experiments and again this is particularly true after day 6.

This behaviour is not limited to the dominant component. As a further example we quote the comparisons for the (8, 15) vorticity coefficient. At day 4, the triangular 21 experiments all give the amplitude to within 2% and the phase to within 0.31°. The triangular 42 gives an amplitude 5% different and a phase 0.78° different from the S30 case. At day 6, the figures are 1%, 0.06° and 19% and 10.83°.

We conclude that errors in vorticity amplitudes and phases due to time truncation are very small and that they are negligible compared with the errors due to space truncation.

(b) Conservation properties

The total energy and mass of the atmosphere are formally conserved by the continuous equations and by the vertical finite difference scheme but not by the horizontal spectral scheme or by the time differencing. Fig. 10 shows the spurious change in total energy normalized by the *change* in kinetic energy from day 0 to day 6. At day 6 the total energy



Figure 10. The spurious change in total energy normalized by the change in kinetic energy at day 6. The value plotted at a time is the maximum up to that time. From the top, the curves are for S90, S30, S5. The ordinate is logarithmic.

has changed by 10^{-3} , 10^{-4} , 2.5×10^{-5} of the kinetic energy change for S90, S30, S5 respectively. The increase in error after this time is rather more rapid. This is probably related to the increase in energy in the high wavenumbers.

In Fig. 11 is exhibited the change in total mass of the atmosphere normalized by the mass itself. At day 6 the proportional change of mass is approximately 5×10^{-8} , 10^{-8} , 2×10^{-10} for S90, S30, S5 respectively. The rate of loss of mass in the S90 experiment, if considered as representative for a general circulation experiment, corresponds to 1 mb in 300 years.

The conservation properties improve by approximately an order of magnitude between S90 and S30, and between S30 and S5. However, even for S90 they are extremely good.

(c) Gravity wave treatment

The most sensitive indicators of the presence and structure of gravity wave modes are the divergence coefficients. As discussed in section 3, a gravity wave has $D_n^m \propto e^{i\sigma t}$ corresponding to $D \propto \cos(m\lambda + \sigma t) P_n^m(\mu)$. The frequency, σ , and the vertical structure are determined by the matrix **B**. For a particular total wavenumber n = 10, the effect of using a semi-implicit scheme with various time-steps is exhibited in the table.



Figure 11. The spurious change of total mass normalized by the mass itself - otherwise as Fig. 10.

	External mode Velocity 302m s ⁻¹ Period 3.51h			First internal mode Velocity 101m s ⁻¹ Period 10.5h			Second internal mode Velocity 32.5m s ⁻¹ Period 32.6h		
Δt (min)	$\sigma \Delta t$	σ_{e}/σ	Period(h)	$\sigma \Delta t$	σ_c/σ	Period(h)	$\sigma \Delta t$	σ_c/σ	Period(h)
5	0.149	0.994	3.53	0.020	0.999	10.5	0.016	1.000	32.6
30	0.892	0·816	4.30	0.300	0.972	10.8	0.096	0.997	32.7
90	2.68	0 ∙452	7.76	0.900	0.814	12.9	0.289	0.973	33.5

TABLE 1. TREATMENT OF GRAVITY WAVES WITH TOTAL WAVENUMBER n = 10

The value of $\sigma \Delta t$ fixes the position on the graph in Fig. 2. The reduction in the computed σ and the lengthened period, $2\pi/\sigma$, are also given.

A divergence coefficient may be considered as composed of a sum of components due to slow moving baroclinic waves and fast moving gravity waves. In this paper, the baroclinic wave is eastward propagating and amplifying. To obtain an idea of the 'signature' of the gravity waves, we first consider the divergence coefficient associated with a slow, eastward propagating wave of constant amplitude, and a single gravity wave mode composed of



Figure 12. A schematic of the complex coefficient D_n^m shown in an Argand diagram for the case $D_n^m = ae^{-i\omega t} + b_{\overline{w}}e^{i\sigma t} + b_{\overline{w}}e^{i\sigma t}$

where a = 0.015 is the amplitude of a slow eastward propagating mode with $2\pi/\omega = 6$ days, σ corresponds to the frequency of the first internal gravity wave mode for n = 10 ($2\pi/\sigma \sim 10.5$ h), b_E is the amplitude of the eastward moving component and b_W the amplitude of the westward moving component. In the left column, the gravity mode is pure leastward ($b_W = 0$), in the centre it is stationary ($b_E = b_W$), and in the right column it is pure westward ($b_E = 0$). The top row is for ($b_E + b_W$) $\sigma = 0.8a\omega$ and the bottom for ($b_E + b_W$) $\sigma = 2.4a\omega$. The dots represent values every $1\frac{1}{4}$ hr and the lines to the origin indicate the values at discrete days. 3 days' values are plotted, starting from the positive real axis. The tick marks on the axes indicates values every 0.01.

eastward and westward propagating components. The 'harmonic dial' for this coefficient D_n^m in various cases is exhibited in Fig. 12. When the gravity wave is in the same direction as the meteorological mode, there is a tendency to form cusps or loops on the inside of the harmonic dial. When it moves in the opposite direction they are formed on the outside.

The actual harmonic dials for D_{10}^8 in the integrations up to day 7 are shown in Fig. 13. There is clear evidence of small gravity wave fluctuations on the growing baroclinic wave.



Figure 13. The harmonic dials for D_{10}^{a} in the S5 (above) and S90 (below) integrations for layers 5 to 1 reading from left to right. 7 days are shown and the notation is as in Fig. 12.

Noting the periods and vertical structures and comparing with the table and with Fig. 4, some definite identification of modes can be made. However, for the purposes of this paper, we comment only that the individual gravity waves are well represented by even the S90 run. The external mode does not appear to have been generated and, consistent with the table, the internal modes are not significantly slowed. The reduction in speed would be more marked for the higher wavenumber external and first internal modes, but study of the results suggests that this effect is not important due to the lack of energy in these modes.

6. THE ARAKAWA VERTICAL DIFFERENCING SCHEME AND ANGULAR MOMENTUM CONSERVATION

In an attempt to evaluate the effect of using an alternative vertical difference scheme, the usefulness of having angular momentum conservation, and the treatment of this conservation by the spectral method, we repeated the above experiment using the vertical scheme of Arakawa (1972). This scheme is based on the thermodynamic and hydrostatic equations (Eqs. (3) and (5)) in their forms:

$$\frac{\partial T'}{\partial t} = -\mathbf{V} \cdot \nabla T' - \sigma^{\kappa} \dot{\sigma} \frac{\partial}{\partial \sigma} (T\sigma^{-\kappa}) + \kappa T \left(\frac{\partial \ln p_{*}}{\partial t} + \mathbf{V} \cdot \nabla \ln p_{*} \right),$$
$$\frac{\partial \phi}{\partial \sigma^{\kappa}} = -\frac{1}{\kappa} T \sigma^{-\kappa}.$$

Inserting finite differences, the vertical advection in the thermodynamic equation is represented by $\sigma^{\kappa} \overline{\delta \delta_{\sigma}} (T \sigma^{-\kappa})^{\sigma}$. The hydrostatic equation becomes

$$\phi_r - \phi_{r+1} = \alpha_r T_r + \beta_{r+1} T_{r+1},$$

where $\alpha_r = \frac{1}{2\kappa} \left[\left(\frac{\sigma_{r+1}}{\sigma_r} \right)^{\kappa} - 1 \right], \ \beta_{r+1} = \frac{1}{2\kappa} \left[1 - \left(\frac{\sigma_r}{\sigma_{r+1}} \right)^{\kappa} \right].$

The height of the lowest layer is determined by insisting that the finite difference equivalent of

$$\int_0^1 \left[T - (\phi - \phi_*)\right] d\sigma = 0,$$

remains true in the form

$$\sum_{1}^{N} [T_r - (\phi_r - \phi_*)] \Delta \sigma_r = 0.$$

This implies that

$$\phi_{N} - \phi_{*} = \sum_{r=1}^{N-1} \Delta \sigma_{r} \left[T_{r} - \sum_{s=r}^{N-1} (\alpha_{s} T_{s} + \beta_{s+1} T_{s+1}) \right] + T_{N} \Delta \sigma_{N}.$$

Thus the height of the lowest layer depends on the temperatures at all heights. In particular for a model with equally spaced layers, it can be shown that the height of the lowest layer is always badly overestimated, this overestimation depending strongly on the temperature at the top layer. Thus the integral condition is retained at the expense of putting all the errors into the height of the lowest layer. It can be shown that the above scheme conserves total energy and angular momentum.

The Arakawa scheme was used in an experiment otherwise identical with S30 described above. Up to day 8, all the vorticity coefficients were almost always within 3% of those in S30 and the energy conservation was extremely similar. Fig. 14 shows the conservation of



Figure 14. The spurious change in total angular momentum normalized by that part of the angular momentum due to the average change in solid body rotation at day 6. This change is approximately equivalent to 0.1m s^{-1} at the equator. The upper curve is for the S30 run and the lower one for the run using the Arakawa vertical scheme.

angular momentum normalized by that part due to the average change in solid-body rotation at day 6. This change at the equator is only 0.1 m s^{-1} . The upper curve shows the S30 run. S5 and S90 were everywhere within 2%, suggesting that the error is due to the vertical scheme. This is confirmed by the two orders of magnitude decrease when using the Arakawa scheme. Again the error associated with the horizontal spectral and semi-implicit methods is very small.

We conclude that the Arakawa vertical scheme gives a much improved angular momentum conservation, but that the change in the baroclinic instability is negligible. This indicates that the vertical resolution in the runs described here is sufficient to well resolve the mode.

7. Aliasing in the triple correlation terms

To test the effect of allowing aliasing in the triple correlation terms, three runs were made on a baroclinic instability problem using a 2-layer model. The runs were

- (i) with the model as above
- (ii) truncating V. $\nabla \ln p_*$ spectrally before computing $\dot{\sigma}$ and ω/p
- (iii) using a grid large enough for the triple terms.

The latter two runs did not alias the triple products. The conservation properties (whose

computation is themselves subject to aliasing and truncation) were almost identical in the runs, and individual wave coefficients identical to the fourth significant figure. Clearly the aliasing of the triple products is not a problem here.

8. CONCLUSION

We have provided some details of the formulation of a very efficient multi-layer primitive equation model on the sphere, using a spectral horizontal representation and a semi-implicit time scheme. For the resolution employed, a computer store of 38K words is required for a hemispheric integration, and the time taken for a one-day simulation on the CDC 7600 is 11 seconds. A detailed analysis of the value of different types and levels of truncation is given in Simmons and Hoskins (1975), where comparisons are made with second order finite difference integrations in a baroclinic instability problem. It seems that the resolution used here may well be sufficient for some general circulation studies.

The semi-implicit time-stepping technique has been analysed both theoretically and by comparison of integrations with different time steps. It is concluded that the errors involved in using the largest possible time-step are negligible compared with those due to space truncation.

The conservation of energy, mass and, when conserved by the vertical scheme, angular momentum are exceedingly good.

It is hoped to proceed with this model in two directions. One is the study of particular systems such as the baroclinic wave simulated in this paper. This will be aided by the extra dimension offered by the availability of wave amplitudes and phases at each instant as well as the usual physical space fields. The other approach must be to insert representations of physical processes occurring in the atmosphere so that a spectral general circulation model can be compared with the usual finite difference models. The inherent errors in each are different and often, in our experience, opposite in sign so that a clear idea may be obtained as to which properties are numerical in character and which a result of the physics included.

ACKNOWLEDGMENT

The constant interest and support provided by Professor R. P. Pearce and Dr. A. Hollingsworth has been much appreciated.

References

Arakawa, A.	1972	'Design of the UCLA general circulation model,' U.C.L.A. Met. Dep. Tech. Report No. 7.
Bourke, W.	1972	'An efficient one-level primitive equation spectral model,' Mon. Weath. Rev., 100, pp. 683-689.
	1974	'A multi-level spectral model. I. Formulation and hemis- pheric integrations,' Mon. Weath. Rev., 102, pp. 687- 701.
Burridge, D.	1974	'Recent developments in the 10-layer model,' Talk at the R. Met. Soc. summer conference, Exeter, 1974.
Corby, G. A., Gilchrist, A. and Newson, R. L.	1972	[•] A general circulation model of the atmosphere, suitable for long period integrations,' <i>Quart. J. R. Met. Soc.</i> , 98 , pp. 809–832.
Doron, E., Hollingsworth, A., Hoskins, B. J. and Simmons, A. J.	1974	'A comparison of grid-point and spectral methods in a meteorological problem,' <i>Ibid.</i> , 100, pp. 371-383.
Eliasen, E., Machenauer, B. and Rasmussen, E.	1970	On a numerical method for integration of the hydrodynami- cal equations with a spectral representation of the horizontal fields.' Inst. of Theor. Met., Univ. of

Copenhagen, Report No. 2.

Elsaesser, H. W.	1966	[•] Evaluation of spectral versus grid methods of hemispheric numerical weather prediction, J. Appl. Met., 5, pp. 246-262.
Hollingsworth, A.	1975	'Baroclinic instability on the sphere,' Quart. J. R. Met. Soc., 101, pp. 495-528.
Hoskins, B. J.	1973	'Stability of the Rossby-Haurwitz wave,' <i>lbid.</i> , 99, pp. 723-745.
Machenauer, B. and Daley, R.	1972	'A baroclinic primitive equation model with a spectral representation in three dimensions,' Inst. of Theor. Met., Univ. of Copenhagen, Report No. 4.
Machenhauer, B. and Rasmussen, E.	1972	'On the integration of the spectral hydrodynamical equa- tions by a transform method,' <i>Ibid.</i> , Report No. 3.
Miyakado, K.	1960	'The methods of numerical time integration of one dimen- sional linear equations and their inherited errors,' J. Met. Soc. Japan, 38, pp. 259-287.
Orszag, S. A.	1970	' Transform method for calculation of vector coupled sums: application to the spectral form of the vorticity equa- tion,' J. Atmos. Sci., 27, pp. 890-895.
Robert, A. J. Henderson, J. and Turnbull, C.	1972	An implicit time integration scheme for baroclinic modes of the atmosphere, Mon. Weath. Rev., 100, pp. 329-335.
Simmons A. J. and Hoskins, B. J.	1975	[•] A comparison of spectral and finite difference simulations of a growing baroclinic wave. [•] Quart. J. R. Met. Soc., 101, pp. 551-565.
Smagorinsky, J., Manabe, S. and Holloway, J. L.	1965	[•] Numerical results from a nine-level general circulation model of the atmosphere, [•] Mon. Weath. Rev., 93, pp. 727-768.

Appendix I

DETAILS OF THE VERTICAL AND TIME SCHEMES

For the T scheme, the matrices and vectors defined in section 3 are

$$\begin{split} \mathcal{D}_{r} &= \frac{1}{1-\mu^{2} \partial \lambda} \mathscr{F}_{vr} + \frac{\partial}{\partial \mu} \mathscr{F}_{vr} - \nabla^{2} \left(\frac{U_{r}^{2} + V_{r}^{2}}{2(1-\mu^{2})} \right), \\ \mathcal{F}_{r} &= -\nabla \cdot (\mathbf{V}_{r}T_{r}^{\prime}) + T_{r}^{\prime}D_{r}^{\prime} - \frac{1}{2\Delta\sigma_{r}} \left\{ \dot{\sigma}_{r+\frac{1}{2}}(T_{r+1}^{\prime} - T_{r}^{\prime}) + \dot{\sigma}_{r-\frac{1}{2}}(T_{r}^{\prime} - T_{r-1}^{\prime}) \right. \\ &+ \left(\overline{T}_{r+1} - \overline{T}_{r}\right) \left[\left(\sum_{s=1}^{r} \Delta\sigma_{s} \right) \sum_{1}^{N} \Delta\sigma_{s} \mathbf{V}_{s} \cdot \nabla \ln p_{*} - \sum_{s=1}^{r} \Delta\sigma_{s} \mathbf{V}_{s} \cdot \nabla \ln p_{*} \right] \\ &+ \left(\overline{T}_{r} - \overline{T}_{r-1}\right) \left[\left(\sum_{s=1}^{r-1} \Delta\sigma_{s} \right) \sum_{1}^{N} \Delta\sigma_{s} \mathbf{V}_{s} \cdot \nabla \ln p_{*} - \sum_{s=1}^{r-1} \Delta\sigma_{s} \mathbf{V}_{s} \cdot \nabla \ln p_{*} \right] \\ &+ 2\kappa \overline{T}_{r} \left[\alpha_{r} \sum_{s=1}^{r} \Delta\sigma_{s} \mathbf{V}_{s} \cdot \nabla \ln p_{*} + \alpha_{r-1} \sum_{s=1}^{r-1} \Delta\sigma_{s} \mathbf{V}_{s} \cdot \nabla \ln p_{*} \right] \\ &+ 2\kappa T_{r}^{\prime} \left[\alpha_{r} \sum_{s=1}^{r} \Delta\sigma_{s} (D_{s} + \mathbf{V}_{s} \cdot \nabla \ln p_{*}) + \alpha_{r-1} \sum_{s=1}^{r-1} \Delta\sigma_{s} (D_{s} + \mathbf{V}_{s} \cdot \nabla \ln p_{*}) \right] \right\} \\ &+ \kappa (\overline{T}_{r} + T_{r}^{\prime}) \mathbf{V}_{r} \cdot \nabla \ln p_{*}, \\ \tau_{rs} &= \frac{1}{2\Delta\sigma_{r}} \left\{ (\overline{T}_{r+1} - \overline{T}_{r}) \left[\Delta\sigma_{s} \left(\sum_{i=1}^{r} \Delta\sigma_{i} \right) - \left| \begin{array}{c} 0 & (r < s) \\ \Delta\sigma_{s} & (r > s) \end{array} \right] \\ &+ (\overline{T}_{r} - \overline{T}_{r-1}) \left[\Delta\sigma_{s} \left(\sum_{i=1}^{r-1} \Delta\sigma_{i} \right) - \left| \begin{array}{c} 0 & (r - 1 < s) \\ \Delta\sigma_{s} & (r - 1 \ge s) \end{array} \right] \\ &+ 2\kappa \overline{T}_{r} \Delta\sigma_{s} \left[\begin{array}{c} \alpha_{r} & (s < r) \\ \alpha_{r} & (s < r) + \end{array} \right] \right\}, \end{split}$$



Appendix II

INITIALIZATION

The inverse balance equation method of obtaining pressure and height fields to balance the imposed velocities such as to give zero initial divergence tendency gives, from Eq. (17), for each spherical harmonic component five equations of the form $\phi_n^m + \overline{T} \ln p_{*n}^m = y_n^m$ for the six unknowns. Therefore

$$T_n^m = g^{-1}(\phi_n^m - \phi_{*n}^m) = -g^{-1}\phi_{*n}^m + g^{-1}y_n^m - \ln p_{*n}^m g^{-1}\overline{T}.$$

In our case there is no topography. It is found that $g^{-1}y$ and $g^{-1}T$ have a large two grid wave in the vertical. Two methods for choosing $\ln p_*$ and thus determining ϕ and T are suggested:

(a) extrapolate \overline{T} and y to the surface ($\phi = 0$) and let

$$\ln p_* = y_*/T_*$$

(b) choose $\ln p_*$ so as to minimize the two grid wave in **T**.

In a wide variety of cases, the two methods have been shown to give almost the same answer and to give a satisfactorily balanced initial condition. In a few cases, however, the temperature still shows oscillation in the vertical. In this paper we use method (b) and define $\ln p_*$ by demanding that a binomial filter (1, -4, 6, -4, 1) on T produce zero result. Thus $\ln p_*$ is chosen so that a cubic polynomial can be fitted through the five values of T.

Since T and $\ln p_*$ occur in \mathcal{D} , \mathcal{T} and \mathcal{P} , an iterative procedure is necessary to determine them in the manner described above. We use five iterations starting with zero values.

Model constants are:

$$R = 287 \text{m}^2 \text{s}^{-2} \text{K}^{-1}, a = 6,371 \text{km}, \Omega = 7.292 \times 10^{-5} \text{s}^{-1}$$

The dimensional layer average temperatures for the run described here are

 $\overline{T} = (220, 230, 250, 267, 280)$ K.

The superrotation is equivalent to equatorial velocities

$$(45, 35, 22, 12, 4)$$
m s⁻¹