Q. J. R. Meteorol. Soc. (1989), **115**, pp. 201–219 551.509.313

The slow equations

By PETER LYNCH
Irish Meteorological Service, Dublin, Ireland

(Received 27 January 1988; revised 5 August 1988)

SUMMARY

A filtered system of equations is derived, using the ideas of normal mode initialization. This *slow equation* system models the low frequency rotational atmospheric motions; there are no solutions corresponding to gravity waves. The prognostic element of the system is an equation expressing the conservation of potential vorticity. The slow equations differ from the general balance system at second order in the Rossby number, and are free from the spurious solutions found in that system. Integration of a barotropic model with the slow equations shows them to be highly accurate when compared with the primitive equations. Since the slow equation model has no high frequency solutions and is free from data shock, it may be useful for continuous assimilation of observational data.

1. INTRODUCTION

The object of numerical weather prediction is to forecast the evolution of the slow, rotational motions of the atmosphere. The high frequency gravity waves which are also solutions of the primitive equations have been causing headaches for modellers ever since Richardson's (1922) pioneering forecast.

The first successful numerical forecasts by Charney *et al.* (1950) circumvented the problem by filtering the equations so that only the slow motions remain as solutions. However, the resulting quasi-geostrophic equations were not always sufficiently accurate, and have systematic errors which make them unsuitable for operational forecasting. The balance equations are based on assumptions less drastic than those leading to the quasi-geostrophic system, and should therefore be more accurate. Daley (1982) has constructed a non-iterative procedure for integrating the balance system, and his results confirm its high accuracy. A balance system involving minimal filtering assumptions has also been proposed recently by Thompson (1980). However, the general balance system admits spurious non-physical solutions (with phase speeds far in excess of the gravity wave speed) in addition to the rotational modes (Moura 1976).

Since the primitive equations support gravity waves, forecasts made with them may be very noisy unless the initial data reflect the balance between the mass and wind fields found in the atmosphere. Modification of the data to ensure this balance is called initialization.

A more direct attack on the noise problem would be to develop a filtered system which simulates the rotational flow accurately but is free from gravity waves (and spurious high frequency solutions). Using the ideas of normal mode initialization, Daley (1980) devised a system in which the low frequency components of the flow are forecast, while appropriate gravity wave components are diagnosed at each moment. This system proved to be highly accurate when compared with a primitive equation model. Combining Daley's approach with the implicit normal mode method of Temperton (1985), it is possible to express this system in terms of the physical variables (obviating the need for transformations to and from normal mode space). We shall call the resulting system the *slow equations*. (A recent paper of Browning and Kreiss (1987) also discusses filtered systems derived using constraints based on initialization theory.)

The slow equations may also be deduced by means of scaling arguments similar to those used to derive the balance system, and based on smallness of the Rossby number, *Ro*. Both systems are accurate up to $O(Ro^2)$, and may be expected to yield results of comparable accuracy. There is a difference at $O(Ro^2)$ between the systems, with the consequence that the slow equations are free from the spurious unphysical solutions.

The slow system has one prognostic component, the equation of conservation of potential vorticity. This equation is in a form ideally suited to the application of the semi-Lagrangian method of integration. The diagnostic elements of the system may be written as standard Helmholtz equations, whose solution is straightforward. Parallel runs show the slow system to be very accurate when compared with a primitive equation model. When used with a zero timestep (i.e. omitting the semi-Lagrangian step) the slow system can be used to initialize data for a primitive equation run. This is precisely the implicit normal mode method of initialization (Temperton 1985; Juvanon du Vachat 1986). The method is equivalent to the filtering condition B derived heuristically by Bourke and McGregor (1983).

The slow equations may provide a suitable means for the assimilation of observational data during a forecast, as they can absorb inserted data without suffering high frequency shocks. The general balance equation, relating the mass and wind fields, is an inherent component of the slow system, and it may conveniently be used to define a local balance when data are inserted, in such a way that perturbations of given magnitude (i.e. in agreement with observations) are assimilated by the system. A further diagnostic component of the system, the imbalance equation, ensures the definition of an appropriate divergence field. The response of the primitive and slow equation models to inserted data is discussed in Lynch (1987).

2. Derivation of the slow equations

(*a*) *Background*

Daley (1980) has used the ideas of normal mode initialization to develop a method of integrating the primitive equations efficiently. The original equations can be written

$$\dot{\mathbf{X}} + \mathbf{L}\mathbf{X} + \mathbf{N}(\mathbf{X}) = \mathbf{0} \tag{1}$$

where $\mathbf{X}$ is the state vector of unknowns, $\mathbf{L}$ is a constant linear operator (matrix) and $\mathbf{N}$ is a nonlinear vector function. When transformed to Hough mode space, the system splits into two subsystems:

$$\dot{\mathbf{Y}} + \mathbf{\Lambda}_\mathbf{Y}\mathbf{Y} + \mathbf{N}_\mathbf{Y}(\mathbf{Y}, \mathbf{Z}) = \mathbf{0} \tag{2}$$

$$\dot{\mathbf{Z}} + \mathbf{\Lambda}_\mathbf{Z}\mathbf{Z} + \mathbf{N}_\mathbf{Z}(\mathbf{Y}, \mathbf{Z}) = \mathbf{0} \tag{3}$$

where $\mathbf{Y}$ and $\mathbf{Z}$ are respectively the coefficients of the slow and fast components of the flow and $\mathbf{\Lambda}_\mathbf{Y}$, $\mathbf{\Lambda}_\mathbf{Z}$ are diagonal matrices of eigenfrequencies. Machenhauer (1977) proposed that (2)–(3) could be initialized by setting the tendencies of the fast modes to zero. Assuming Machenhauer's criterion ($\dot{\mathbf{Z}} = 0$) to hold throughout the integration, Daley replaced (2)–(3) by the system

$$\dot{\mathbf{Y}} + \mathbf{\Lambda}_\mathbf{Y}\mathbf{Y} + \mathbf{N}_\mathbf{Y}(\mathbf{Y}, \mathbf{Z}) = \mathbf{0} \tag{4}$$

$$\mathbf{\Lambda}_\mathbf{Z}\mathbf{Z} + \mathbf{N}_\mathbf{Z}(\mathbf{Y}, \mathbf{Z}) = \mathbf{0} \tag{5}$$

giving a prognostic equation for the slow modes and a diagnostic equation for the fast modes. I will call (4)–(5) the 'slow equations' (in normal mode form). Using the slow

equations, Daley developed an integration scheme which was stable, efficient and accurate: he compared a run with the slow equations and $\Delta t = 40$ min to a control run with the primitive equations and $\Delta t = 10$ min; the r.m.s. differences in surface pressure and in 500 hPa winds at 48 hours were only 0·6 hPa and 1 m s^{-1}. The great majority of the additional computational effort (per timestep) in Daley's scheme is due to the transformations between spectral (spherical harmonic) space and normal (Hough-) mode space; for a gridpoint model the transformations would be even more expensive.

Temperton (1985) has devised an initialization scheme which is completely equivalent to the normal mode method, but which operates in physical space. A very similar method has been presented by Juvanon du Vachat (1986). The central idea of Temperton's approach is to choose a linearization and to reformulate the physical equations so that the tendencies can be separated by inspection into slow and fast components:

$$\dot{\mathbf{X}} = \dot{\mathbf{X}}_R + \dot{\mathbf{X}}_G. \tag{6}$$

(Subscripts R and G denote the Rossby and gravity wave projections.) Using Machenhauer's criterion, $\dot{\mathbf{X}}_G = 0$ at $t = 0$, Temperton derived an initialization scheme which he called the implicit normal mode method; the same approach will be used here to derive a system of slow equations in physical space.

In the following section we will use the ideas of Daley and Temperton to develop a system of equations equivalent to that of Daley, but formulated in terms of the physical variables. The integration of this system does not require any transformations between physical and spectral space; nor does it require knowledge of the linear normal modes of the model being integrated.

(*b*) *Derivation of the slow equations*

A general baroclinic system of equations may be separated, by transformation to the vertical eigenmodes, into a number of systems equivalent to the shallow water equations. Therefore, we consider the shallow water system

$$\dot{\zeta} + f\delta = -N_\zeta \tag{7}$$

$$\dot{\delta} - f\zeta + \nabla^2\Phi = -N_\delta \tag{8}$$

$$\dot{\Phi} + \bar{\Phi}\delta = -N_\Phi \tag{9}$$

where N_ζ, N_δ and N_Φ represent the nonlinear terms, dots denote local time derivatives and otherwise the notation is conventional. The Coriolis parameter f is variable, but the β terms are included on the right-hand side. We assume that $f = f(x, y)$ depends upon both horizontal coordinates (as it does for a polar stereographic projection or for rotated latitude/longitude coordinates) and denote its derivatives by β_x and β_y. Then the nonlinear terms are

$$N_\zeta \equiv \mathbf{V}\cdot\nabla\zeta + \zeta\delta + (\beta_y v + \beta_x u)$$

$$N_\delta \equiv \mathbf{V}\cdot\nabla\delta + \delta^2 + (\beta_y u - \beta_x v) - 2J(u, v)$$

$$N_\Phi \equiv \mathbf{V}\cdot\nabla\Phi + (\Phi - \bar{\Phi})\delta.$$

The vorticity and continuity equations may be written

$$d(\zeta + f)/dt + (\zeta + f)\delta = 0 \qquad \text{and} \qquad d\Phi/dt + \Phi\delta = 0$$

and from these we easily derive an equation expressing the conservation of potential vorticity

$$\frac{d}{dt}\left(\frac{\zeta + f}{\Phi}\right) = 0. \tag{10}$$

Expanding the time derivative and keeping only linear terms on the left, this may also be written in the form

$$\frac{\partial(\bar{\Phi}\zeta - f\Phi)}{\partial t} = -(\bar{\Phi}N_\zeta - fN_\Phi). \tag{11}$$

This equation may alternatively be derived directly from (7) and (9).

Although (7)–(9) are not separable, we can easily deduce some crucial properties of their linear eigenmodes (Temperton 1985, 1988): (A) The slow modes are stationary, geostrophic and nondivergent. (B) The fast modes have zero linearized potential vorticity ($\Pi \equiv \bar{\Phi}\zeta - f\Phi$). Property (A) is easily seen by considering stationary solutions of the system (7)–(9). Property (B) follows from (11), linearized so that the r.h.s. vanishes, and the fact that the fast modes must have non-zero frequencies. These properties imply that the tendencies of divergence and (geostrophic) *imbalance* ($\varepsilon \equiv \nabla^2\Phi - f\zeta$) project entirely onto the fast modes, and the tendency of potential vorticity entirely onto the slow modes:

$$\dot{\delta} = 0 + \dot{\delta}_G \tag{12}$$

$$\dot{\varepsilon} \equiv \nabla^2\dot{\Phi} - f\dot{\zeta} = 0 + (\nabla^2\dot{\Phi} - f\dot{\zeta})_G \tag{13}$$

$$\dot{\Pi} \equiv \bar{\Phi}\dot{\zeta} - f\dot{\Phi} = (\bar{\Phi}\dot{\zeta} - f\dot{\Phi})_R + 0. \tag{14}$$

These are the explicit form of (6) above. An equation for the tendency of the geostrophic imbalance is easily derived from (7) and (9):

$$\dot{\varepsilon} + (\nabla^2 - f^2/\bar{\Phi})\bar{\Phi}\delta = -(\nabla^2 N_\Phi - fN_\zeta). \tag{15}$$

This *imbalance equation*, together with (8) and (10) (or (11)), gives a system of three equations completely equivalent to the original system (7)–(9).

We now assume that the gravity wave projections of tendency vanish. Thus, the tendency terms in (8) and (15) are dropped; the potential vorticity equation (10) is left unchanged. The resulting system is

$$\frac{d}{dt}\left(\frac{\zeta + f}{\Phi}\right) = 0 \tag{16}$$

$$\nabla^2\Phi - f\zeta = -N_\delta \tag{17}$$

$$(\nabla^2 - f^2/\bar{\Phi})\bar{\Phi}\delta = -(\nabla^2 N_\Phi - fN_\zeta). \tag{18}$$

We shall call this system the 'slow equations'. They comprise the equation of conservation of potential vorticity (16), the nonlinear balance equation (17) and the imbalance or omega equation (18). They are analogous to Daley's equations (4)–(5), but refer to the physical variables, obviating the need for transformations to and from Hough space.

Although Daley used Eqs. (4) and (5), his criterion to determine the prognostic or diagnostic treatment of each coefficient was based on its associated frequency. His system is thus more general than (16)–(18). A system using the frequency cut-off criterion and formulated in terms of the physical variables can be devised using the Laplace transform technique (Lynch 1985).

The above analysis may be repeated with the β terms included in the linear analysis (Temperton 1988). The resulting equations are a little more cumbersome, and it is unclear whether any advantage is obtained (cf. Lynch 1985, section 4). Similarly, the total derivatives of δ and ε may be omitted, rather than their local tendencies; Thompson (1980) argues that this is a physically more meaningful approximation. Again, the result is a somewhat more complicated form of the nonlinear terms. It is not clear whether the alternative approximation yields better results. Finally, Tribbia (1984) has developed a higher order method of initialization. Rather than assuming that the nonlinear terms in (3) are constant, he represents them by low order polynomials in time. The first-order version of his technique is the same as Machenhauer's method. The second-order version consists of setting to zero the second (time) derivative of the gravity modes. One could use Tribbia's criterion to derive slow equations of higher order; any practical advantages would have to be weighed against the added complexities of the resulting system.

3. The slow equations and the balance system

(*a*) *Introduction*

The slow equations are similar, but not identical, to the general balance system. There is a single prognostic variable, the potential vorticity. Various small terms are retained in the slow equations which are normally dropped in the balance system. However, the most important difference is the omission of the tendency of (geostrophic) imbalance ($\dot{\varepsilon}$) in the slow equations. This may be justified by scaling arguments similar to those which lead to the balance system.

The essential approximation in deriving the balance system is the replacement of the divergence equation by a diagnostic relationship between the streamfunction and the geopotential. The resulting system of equations is highly implicit, and its numerical solution presents some difficulties (Charney 1962; Daley 1982). The general balance system admits spurious high frequency solutions in addition to the Rossby wave solutions (Moura 1976); it will be shown below that the slow equations are free from this problem.

(*b*) *Perturbation expansion*

We consider the shallow water equations, linearized about a state of rest:

$$\partial u/\partial t - fv + \partial \Phi'/\partial x = 0 \tag{19}$$

$$\partial v/\partial t + fu + \partial \Phi'/\partial y = 0 \tag{20}$$

$$\partial \Phi'/\partial t + \bar{\Phi}(\partial u/\partial x + \partial v/\partial y) = 0. \tag{21}$$

The corresponding vorticity and divergence equations are

$$\partial \zeta/\partial t + f\delta + \beta v = 0 \tag{22}$$

$$\partial \delta/\partial t + (\nabla^2 \Phi' - f\zeta + \beta u) = 0. \tag{23}$$

We introduce length and velocity scales L and V, assume an advective time scale L/V and scale Φ' geostrophically as $2\Omega LV$. The non-dimensionalized forms of (19)–(21) then become

$$Ro\, \partial u/\partial t - \mu v + \partial \Phi/\partial x = 0 \tag{24}$$

$$Ro\, \partial v/\partial t + \mu u + \partial \Phi/\partial y = 0 \tag{25}$$

$$Ro^{-1} Fr\, \partial \Phi/\partial t + (\partial u/\partial x + \partial v/\partial y) = 0 \tag{26}$$

where $\mu = \sin\phi$, $Ro = V/2\Omega L$ is the Rossby number and $Fr = V^2/\bar{\Phi}$ is the Froude number. The vorticity and divergence equations become

$$Ro\,\partial\zeta/\partial t + \mu\delta + (L/a)\,\sigma v = 0 \tag{27}$$

$$Ro\,\partial\delta/\partial t + \{\nabla^2\Phi - \mu\zeta + (L/a)\sigma u\} = 0 \tag{28}$$

where $\sigma = \cos\phi$ and a is the earth's radius.

We assume the following values for the scale factors, etc.

$V = 20\ \mathrm{m\,s^{-1}}$, $L = 10^6\ \mathrm{m}$, $2\Omega = 1{\cdot}5 \times 10^{-4}\mathrm{s^{-1}}$, $\bar{\Phi} = 5 \times 10^4\ \mathrm{m^2 s^{-2}}$, $a = 6 \times 10^6\ \mathrm{m}$.

Then the non-dimensional numbers take values as follows: $Ro \sim 10^{-1}$, $Fr \sim 10^{-2}$, $Ro^{-1}Fr \sim Ro$, $L/a \sim Ro$. We have assumed that $L \ll a$ so that $L/a = \mathrm{O}(Ro)$. This is true for synoptic scales of motion; for planetary scales $L \sim a$, and the scaling must be reexamined (Burger 1958).

We will now derive perturbation forms of the equations, correct up to $\mathrm{O}(Ro^2)$. First, the dependent variables are expanded in the Rossby number

$$q = q_0 + q_1 Ro + q_2 Ro^2 + \cdots.$$

The expansions are substituted into the equations of motion and terms of like order in Ro are equated. We perform the expansion process in two different ways: the conventional method leads to the balance system; and a reformulated method, starting from equivalent but different forms of the original equations, yields the slow equations.

Method A. We expand Eqs. (26), (27) and (28). The zero-order terms give

$$\delta_0 = 0 \qquad \text{and} \qquad \nabla^2\Phi_0 = \mu\zeta_0$$

representing non-divergent, geostrophically balanced flow. Next, we consider the terms of order Ro, which give

$$\partial\Phi_0/\partial t + \delta_1 = 0$$

$$\partial\zeta_0/\partial t + \mu\delta_1 + \sigma v_0 = 0$$

$$\nabla^2\Phi_1 - \mu\zeta_1 + \sigma u_0 = 0.$$

In dimensional form, the equations correct up to order $\mathrm{O}(Ro^2)$ are

$$\Phi_t + \bar{\Phi}\delta = 0 \qquad \text{(i)}$$

$$\zeta_t + f\delta + \beta v = 0 \qquad \text{(ii)} \qquad \text{[BE]}$$

$$\nabla^2\Phi - f\zeta + \beta u = 0 \qquad \text{(iii)}.$$

This is the conventional (linearized) balance system (denoted [BE]). We note that there are two prognostic equations, and will show below that this leads to spurious solutions of the system in addition to the Rossby wave solutions.

We have assumed that $L \ll a$ and $L/a = \mathrm{O}(Ro)$, which is true for synoptic scales. Since δ_0 vanishes, the wind field may be partitioned into solenoidal and potential components as follows

$$\mathbf{V} = \mathbf{V}_\psi + Ro\,.\,\mathbf{V}_\chi.$$

Thus, in the β terms of [BE] the winds may be replaced by their rotational components. The equations are then equivalent to the linearized form of Eqs. (5.1–3) in Charney (1962). For planetary scales the full winds must be kept (Charney 1973). It is the inclusion of the term βu_χ in [BE(iii)] which leads to the spurious solutions (Moura 1976).

Method B. We now define the *imbalance* as $\varepsilon = \nabla^2\Phi - f\zeta + \beta u$ and observe from the foregoing analysis that $\varepsilon = O(Ro^2)$. With an advective timescale appropriate to low frequency motions we expect that the tendency of ε is also small. From (19), (21) and (22) we easily deduce

$$\partial\varepsilon/\partial t + (\bar{\Phi}\nabla^2 - f^2)\delta - f\beta v - \beta(fv - \Phi_x) = 0. \tag{29}$$

In non-dimensional form this becomes

$$Ro\,\partial\varepsilon/\partial t + \{(Ro^2Fr^{-1})\nabla^2 - \mu^2\}\delta - (L/a)\mu\sigma v - (L/a)\sigma(\mu v - \Phi_x) = 0 \tag{30}$$

and we recall that $Ro^2Fr^{-1} = O(1)$.

Consider the system of equations (22), (23) and (29) or, in non-dimensional form, (27), (28) and (30). The dependent variables are expanded as before and coefficients of like powers of Ro considered together. The zero-order terms yield

$$\delta_0 = 0 \qquad \varepsilon_0 \equiv \nabla^2\Phi_0 - \mu\zeta_0 = 0$$

which represents non-divergent geostrophically balanced flow, as before. Now the first-order terms give the following system:

$$\partial\zeta_0/\partial t + \mu\delta_1 + \sigma v_0 = 0$$

$$\varepsilon_1 \equiv \nabla^2\Phi_1 - \mu\zeta_1 + \sigma u_0 = 0$$

$$(\nabla^2 - \mu^2)\delta_1 - \mu\sigma v_0 = 0.$$

In dimensional form the equations correct up to $O(Ro^2)$ are

$$\zeta_t + f\delta + \beta v = 0 \qquad \text{(i)}$$

$$\nabla^2\Phi - f\zeta + \beta u = 0 \qquad \text{(ii)} \qquad \text{[SE]}$$

$$(\bar{\Phi}\nabla^2 - f^2)\delta - f\beta v = 0 \qquad \text{(iii)}.$$

This system is denoted [SE] as it is a form of the slow equations, of the same character as those derived in the previous section. The first two equations are identical to those which were obtained above for the [BE] system. However, we now have a diagnostic equation for δ in place of the continuity equation. There is only one prognostic component, and the only normal modes are the Rossby waves—the system is free from spurious modes. This fact, together with the extra simplicity of the system, makes it more attractive than the conventional balance system.

(c) *The transition between the two systems*

Since both the systems [BE] and [SE] have been derived from the same starting point, with the same scaling assumptions, and since both are correct at $O(Ro)$, any differences between them must be $O(Ro^2)$ at most. We will show now that this is the case.

From [BE] it is straightforward to derive the equation

$$\partial(\nabla^2\Phi - f\zeta)/\partial t + (\bar{\Phi}\nabla^2 - f^2)\delta - f\beta v = 0. \tag{31}$$

This equation differs from [SE(iii)] only in having a tendency term. Now, by scaling this equation as before, it appears immediately that this tendency term vanishes at the zeroth and first orders in Ro. But the removal of the term renders the balance and slow systems identical. Therefore, they are equivalent at $O(Ro)$.

Despite the difference between [BE] and [SE] being only $O(Ro^2)$, there is no direct means of transition between them. Considering [BE] *per se*, the tendency term in (31) must be kept. To make the transition we must take a step backwards to the original primitive system. The presence of the tendency term in (31) is a fundamental difference between the balance and slow systems, and justifies different names being given to them.

(*d*) *Linear normal modes*

Since two time derivatives are omitted in the slow system [SE], the dispersion relation is linear in the frequency. There is only a single normal mode for given wavenumbers. The balance system [BE] has two time derivatives; a naive argument would suggest that the system should have a quadratic dispersion relation and therefore admit modes other than the Rossby wave solutions. Let us examine this in more detail.

The system [BE], written in terms of the streamfunction and velocity potential, is

$$\partial\Phi/\partial t + \bar{\Phi}\nabla^2\chi = 0$$
$$\partial\nabla^2\psi/\partial t + f\nabla^2\chi + \beta(\chi_y + \psi_x) = 0$$
$$\nabla^2\Phi - f\nabla^2\psi + \beta(\chi_x - \psi_y) = 0.$$

Seeking solutions of the form $A\exp\{\mathrm{i}(kx + ly - \nu t)\}$ we get a quadratic dispersion relation:

$$(\alpha\beta k)\nu^2 + \{\beta^2(\alpha k^2 + l^2)/K^2 - K^4(\bar{\Phi} + f^2/K^2) + 2f\,.\,\beta\mathrm{i}l\}\nu - \beta k K^2\bar{\Phi} = 0 \quad (32)$$

where $K^2 = (k^2 + l^2)$ and α (=1 or 0) is a tracer indicating inclusion or exclusion of the term βu_χ in the balance equation. For simplicity, let us assume that $l = 0$. If $\alpha = 1$ and we define $c_R \equiv -\beta/K^2$ and $c_G = \sqrt{(\bar{\Phi} + f^2/K^2)}$, the phase speed $c = \nu/k$ must satisfy

$$c_R c^2 + (c_G^2 - c_R^2)c - \bar{\Phi}c_R = 0. \quad (33)$$

Since $|c_R| \ll |c_G|$, the second term in the coefficient of c may be omitted. The two roots may easily be found approximately: for $|c|$ small we omit the quadratic term, arriving at

$$c \doteq c_R(1 + f^2/K^2\bar{\Phi})^{-1}$$

which is the familiar Rossby wave speed; for $|c|$ large the constant term is dropped, giving

$$c \doteq c_S \equiv c_G^2/(-c_R).$$

Since $|c_G| \gg |c_R|$, this solution is much faster than a gravity wave. It is in fact a spurious solution, with no physical significance. Such physically unrealistic high frequency, eastward-propagating solutions were discussed by Moura (1976).

Our naive reasoning has proved correct: the two time derivatives in the balance system give rise to two types of normal mode solution. If the term βu_χ is dropped from the balance equation, we put $\alpha = 0$ and the quadratic term in the dispersion relation (32) disappears. Then the only solution is that having the Rossby wave speed, and the system is free from spurious solutions. In general, the condition necessary to avoid spurious solutions is that the balance equation have no term involving a divergence element (χ, u_χ, v_χ or δ).

(e) *A general filtered system*

The foregoing analysis was applied to the linear shallow water equations (19)–(21). A similar approach may be used in a more general nonlinear context. We write the divergence equation in the form

$$d\delta/dt + \varepsilon = 0 \tag{34}$$

where ε, the *imbalance*, is the sum of all remaining terms. Now we derive an equation for the rate of change of ε, which may be written formally as

$$d\varepsilon/dt + \eta = 0. \tag{35}$$

The derivation of this equation may be algebraically involved. Thompson (1980) has derived an expression for η which includes the effects of nonlinear advection (although the β terms are omitted). Note that (35) may also be written

$$d^2\delta/dt^2 + \eta = 0. \tag{36}$$

To obtain a general filtered system, we omit the time derivatives in (34) and (35). (Either partial or total time derivatives may be omitted; Thompson (1980) argues that only the total derivatives have intrinsic physical meaning, being independent of the coordinate system.) Equation (34) now becomes a generalized balance equation, expressing the relationship between the mass and rotational flow fields. Equation (35) becomes a diagnostic equation for the divergence (analogous to the omega equation); we may call it the imbalance equation, since it is the approximate form of the equation for the tendency of imbalance. We complete the system by adding the equation for conservation of potential vorticity. This general slow system, with a single prognostic component, models the slow, rotational motions and is free from both gravity waves and spurious solutions.

Hinkelmann (1969) proposed a general method for defining initial data for the primitive equations. He suggested that the observed wind and mass fields should be adjusted so that

$$d(\nabla \cdot \mathbf{V})/dt = 0 \qquad \text{and} \qquad d^2(\nabla \cdot \mathbf{V})/dt^2 = 0. \tag{37}$$

That is, these two conditions should be used to derive diagnostic relationships, which could then be used to filter the initial data. As an alternative, he pointed out that these conditions could be used to replace two prognostic equations by diagnostic relationships, yielding a general filtered system; this is precisely what we have done above. Thus, the slow equations are an implementation of the filtering technique originally suggested by Hinkelmann.

It was further argued by Hinkelmann that an improved balance would be achieved by requiring

$$d^n(\nabla \cdot \mathbf{V})/dt^n = 0 \qquad \text{and} \qquad d^{n+1}(\nabla \cdot \mathbf{V})/dt^{n+1} = 0 \tag{38}$$

for $n > 1$. We can re-express these criteria as

$$d^n\delta/dt^n = 0, \qquad d^n\varepsilon/dt^n = 0. \tag{39}$$

Clearly, for $n = 2$ these are essentially the same as Tribbia's (1984) second-order initialization technique. Likewise, there is a correspondence between the procedures for larger n. It is evident that higher order systems of slow equations could be derived by using (38) rather than (37), to derive diagnostic relationships, together with the potential vorticity conservation equation.

We will see below that the difference between a forecast with the slow equations and a reference run with a primitive equation model is of about the same size as the errors obtained by Daley (1982) using the balance system. As the two systems yield results of comparable accuracy, the slow equations may be considered to have a clear advantage over the balance system by virtue of their greater simplicity and the absence of unphysical solutions.

4. Numerical integration of the slow equations

The numerical integration of the slow system is straightforward. We rewrite the system (16)–(18) here for convenience:

$$\frac{d}{dt}\left(\frac{\zeta + f}{\Phi}\right) = 0 \tag{40}$$

$$\nabla^2 \Phi - f\zeta = -N_\delta \tag{41}$$

$$(\nabla^2 - f^2/\bar{\Phi})\bar{\Phi}\delta = -(\nabla^2 N_\Phi - fN_\zeta). \tag{42}$$

The potential vorticity equation (40) is in a form ideally suited to the semi-Lagrangian approach. This method ensures unconditional stability for the time integration, which is a great advantage. After each timestep, the diagnostic relationships (41) and (42) are used to calculate the remaining variables. Certain quantities not known at the new time must be lagged, or evaluated using values from the previous time. The error thus incurred could have been reduced by an iterative procedure, but this was found to be unnecessary. The method used bears some similarity to the non-iterative method used by Daley (1982) to solve the balance equations.

(*a*) *Discretization*

The only prognostic variable is the potential vorticity $\Pi = (\zeta + f)/\Phi$. Since it is convenient to calculate ζ and Φ at the same points, a D grid is used, with v calculated a half gridstep east and u half a gridstep north of each Φ point. The area covered by the grid can be seen in Fig. 1. There are 40×26 points with a spacing of $2° \times 2°$ on a rotated latitude/longitude grid with a pole at 150°E 30°N.

The initial data are the (uninitialized) operational analyses of geopotential height and winds at 500 hPa valid at 0000 UTC, 19 November 1986. The mean height calculated from this analysis is 5439 m. The wind fields were interpolated from a C to a D grid. Calculation of the vorticity ζ, and thence Π, at Φ points gives us the starting values for the integration.

(*b*) *The Lagrangian timestep*

The prognostic element of the slow system is Eq. (40), expressing conservation of potential vorticity. To integrate this equation, we use a two-time-level semi-Lagrangian technique. Thus, (40) is approximated by

$$\Pi_{I,J}^{n+1} = \Pi_*^n \tag{43}$$

where the potential vorticity Π at the gridpoint (I,J) at time $(n+1)\Delta t$ is equal to its value at a departure point, denoted by an asterisk, at time $n\Delta t$. The departure point is estimated using a technique described by McDonald and Bates (1987). This technique uses centring in both time and space, and bilinear interpolation. With the departure point thus estimated, the value of Π at this point at time $n\Delta t$ is calculated by a bicubic

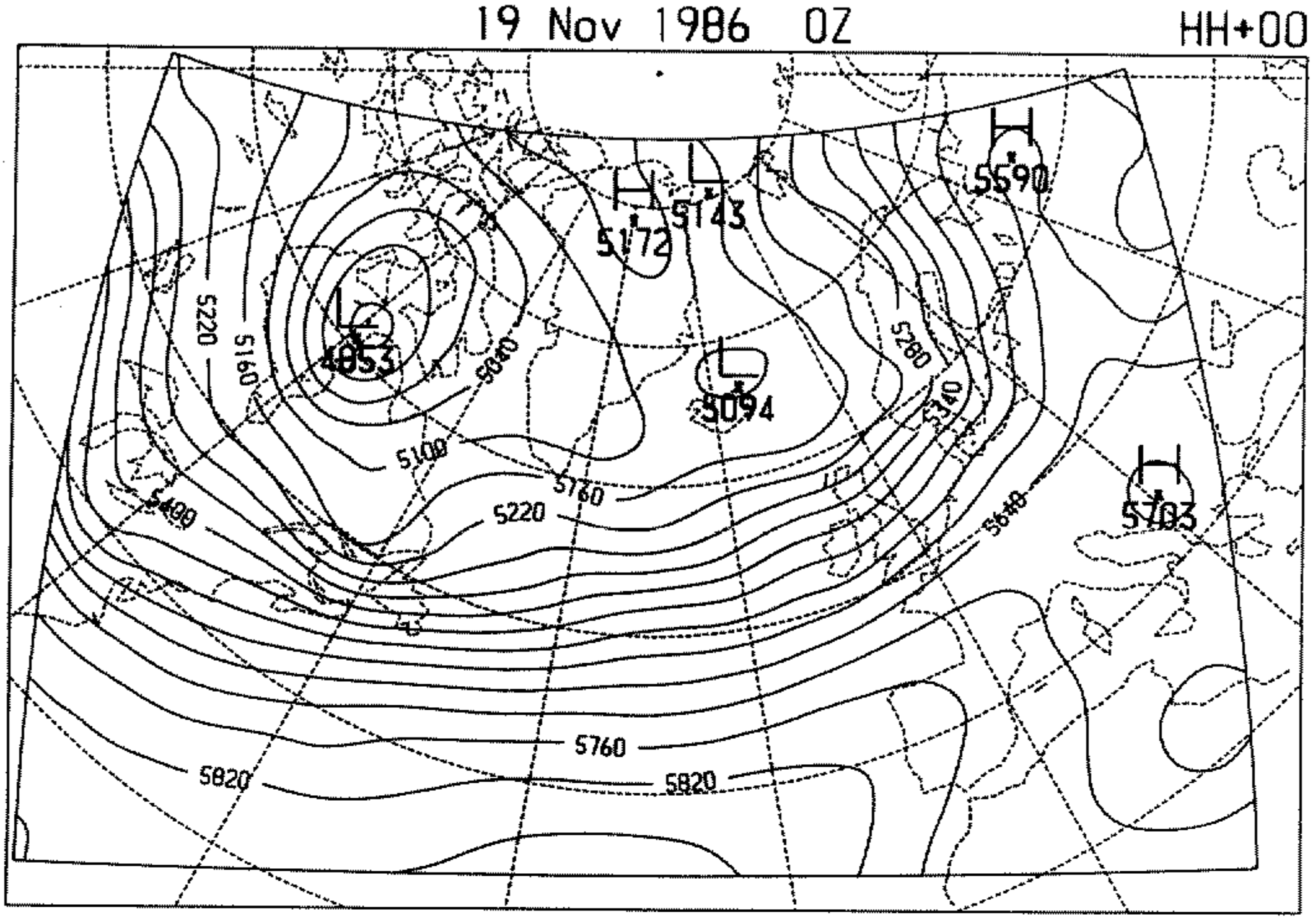

Figure 1. Analysed 500 hPa geopotential height (metres) valid at 00 GMT, 19 November 1986.

interpolation, and by (43) this immediately gives us the value at point (I, J) at the new time. The semi-Lagrangian method is discussed more fully in Bates and McDonald (1982).

(c) *Diagnostic steps*

From (41) and the definition of Π, we get an equation for the geopotential:

$$(\nabla^2 - f^2/\bar{\Phi})\Phi = -N_\delta - f^2 + f\{\Pi - (f/\bar{\Phi})\}\Phi. \tag{44}$$

The nonlinear right-hand terms are calculated using values at $n\Delta t$, except for Π at $(n+1)\Delta t$. Note that the coefficient of Φ on the r.h.s. is small, so that the resulting time truncation error should also be small. Then Φ^{n+1} is obtained by solution of the Helmholtz equation (44). We use SOR for simplicity, although faster methods are available (e.g. see appendix of Temperton 1988). From Π^{n+1} and Φ^{n+1} we immediately get ζ^{n+1}.

The next step is to compute N_ζ and N_Φ and then the right-hand side of (42) using the most up-to-date values available. This Helmholtz equation is then solved for the divergence δ^{n+1}.

The winds at time $(n+1)\Delta t$ must be retrieved from the vorticity and divergence fields. This process requires the specification of a single wind component on the boundary. In the method of Sangster (1960) the streamfunction and velocity potential are calculated by solving two Poisson equations, and the wind is then deduced from them. A more direct method (Lynch 1988), requiring the solution of only one Poisson equation, yields equally accurate results and was used in this study.

Finally, the vorticity is recalculated in order to agree with the wind values specified on the boundaries, and Π is recomputed. Point values and global averages of various

quantities, energy components and other diagnostics are calculated and stored at each timestep. At the end of the integration the winds are transformed back to a C grid to allow direct comparison with a primitive equation model.

5. Comparison with a primitive equation model

To assess the accuracy of the slow equations, a series of parallel runs using slow and primitive equation models was carried out. The slow equation model [SE] is as described in section 4 above. The primitive equation model (denoted [PE]) is a one-level version of the model described by McDonald (1986). It uses a semi-Lagrangian advection scheme and a semi-implicit adjustment step. The grids are identical for the two models, except that [PE] uses a C grid, whereas [SE] uses a D grid, so that u and v points are interchanged.

As a reference run, we take the [PE] 24-hour forecast with a timestep Δt of 10 minutes. The initial analysis of height is shown in Fig. 1, the 24-hour forecast in Fig. 2 and the difference in Fig. 3. To limit gravity wave noise in the primitive equation model, a light divergence damping is applied, with a coefficient value $2\times10^7\ \mathrm{m^2\ s^{-1}}$. This damps out the noise in about 12 hours.

Preliminary runs with the [SE] model differed significantly from the reference, suggesting certain changes. The average divergence was too high, and there were large differences between the runs near the western boundary. To counteract the first effect, the divergence was set to zero on the boundary prior to the solution of (42). This resulted in a reduction of the r.m.s. divergence from $1{\cdot}8\times10^{-6}\ \mathrm{s^{-1}}$ to a more reasonable value $5{\cdot}7\times10^{-8}\ \mathrm{s^{-1}}$. The large differences near the western boundary were due to a discrepancy in the treatment of the boundary conditions. To remove this, ζ and Π were recalculated after retrieval of the winds at each timestep. This led to more compatibility between the two models.

After these adjustments, the differences between the reference run and the [SE] model with a 10-minute timestep were very small. The difference between the height forecasts had an r.m.s. value of only 5·9 m and a maximum difference of 20 m; the difference field is shown in Fig. 4. The r.m.s. difference between the wind forecasts was only $1{\cdot}5\ \mathrm{m\,s^{-1}}$. These differences are comparable in magnitude to typical observational errors.

Both models are capable of being integrated with long timesteps, since they both use stable numerical schemes (the operational 3-D version of [PE] uses a timestep of 90 minutes). To investigate the time truncation errors, both [PE] and [SE] were run with a one-hour timestep, and the results were compared with the corresponding runs with $\Delta t = 10$ minutes. For [PE] the r.m.s. differences in height and winds were 5·8 m and $0{\cdot}9\ \mathrm{m\,s^{-1}}$ respectively. For [SE] the relevant figures were 3·6 m and $0{\cdot}5\ \mathrm{m\,s^{-1}}$. These differences are small enough to suggest that the errors associated with a one-hour timestep are quite tolerable in both cases.

The effects of two different methods of retrieving the winds from the vorticity and divergence were examined. We made two parallel 24-hour forecasts, using the method of Sangster (1960) and the direct method of Lynch (1988) to retrieve the winds at each timestep. The same normal boundary winds at the same gridpoints were used in each case. The r.m.s. difference between the forecast wind fields after 24 hours was only $0{\cdot}13\ \mathrm{m\,s^{-1}}$; the height fields differed by less than one metre. For all practical purposes, the forecasts were identical. A full description of the two methods is given in Lynch (1988).

The results of the comparison between the primitive and slow equation models presented above indicate that the slow equations are capable of accurately modelling the

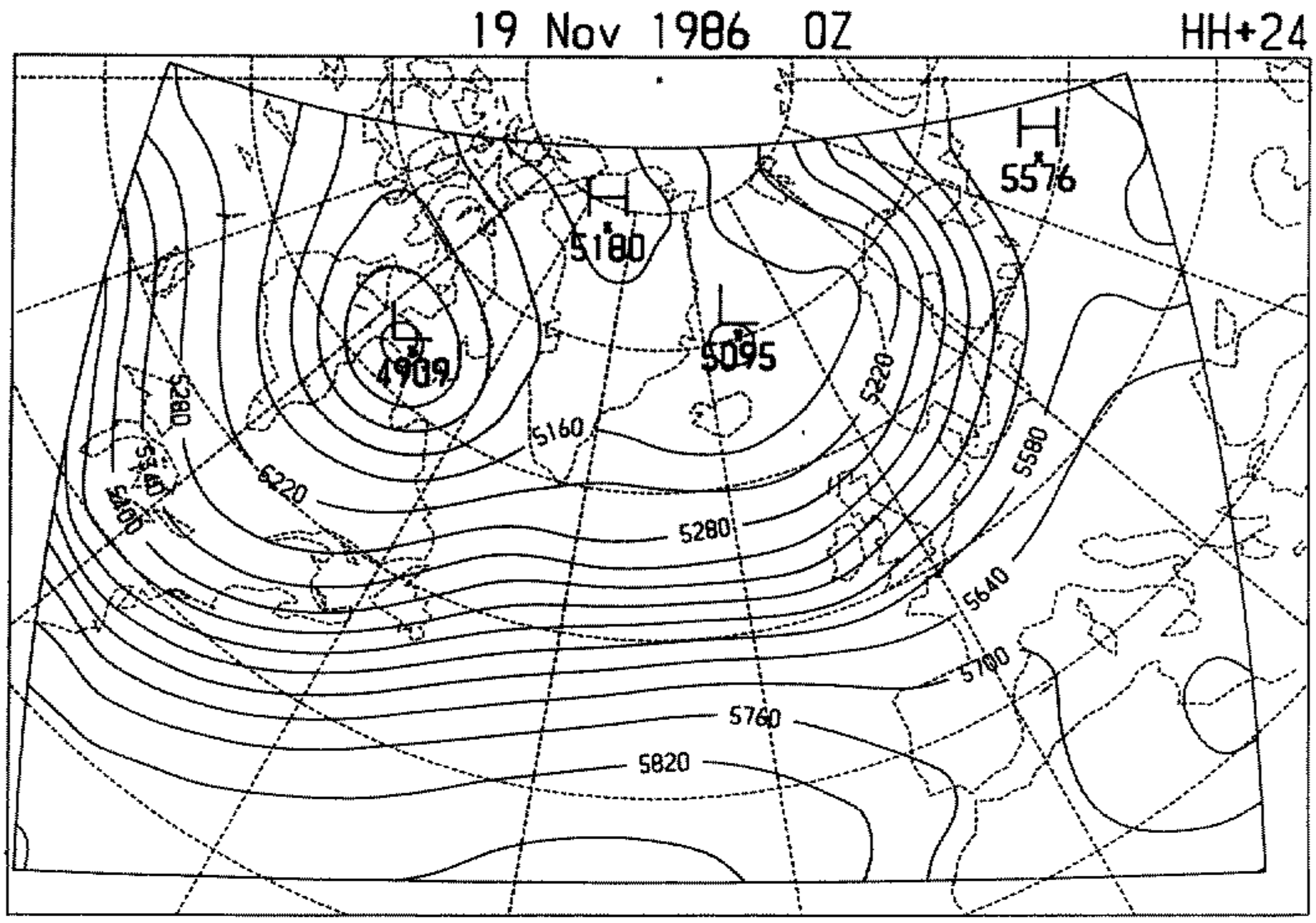

Figure 2. Forecast 500 hPa geopotential height valid at 00 GMT, 20 November 1986. This is the 24-hour forecast made with the primitive equation model [PE].

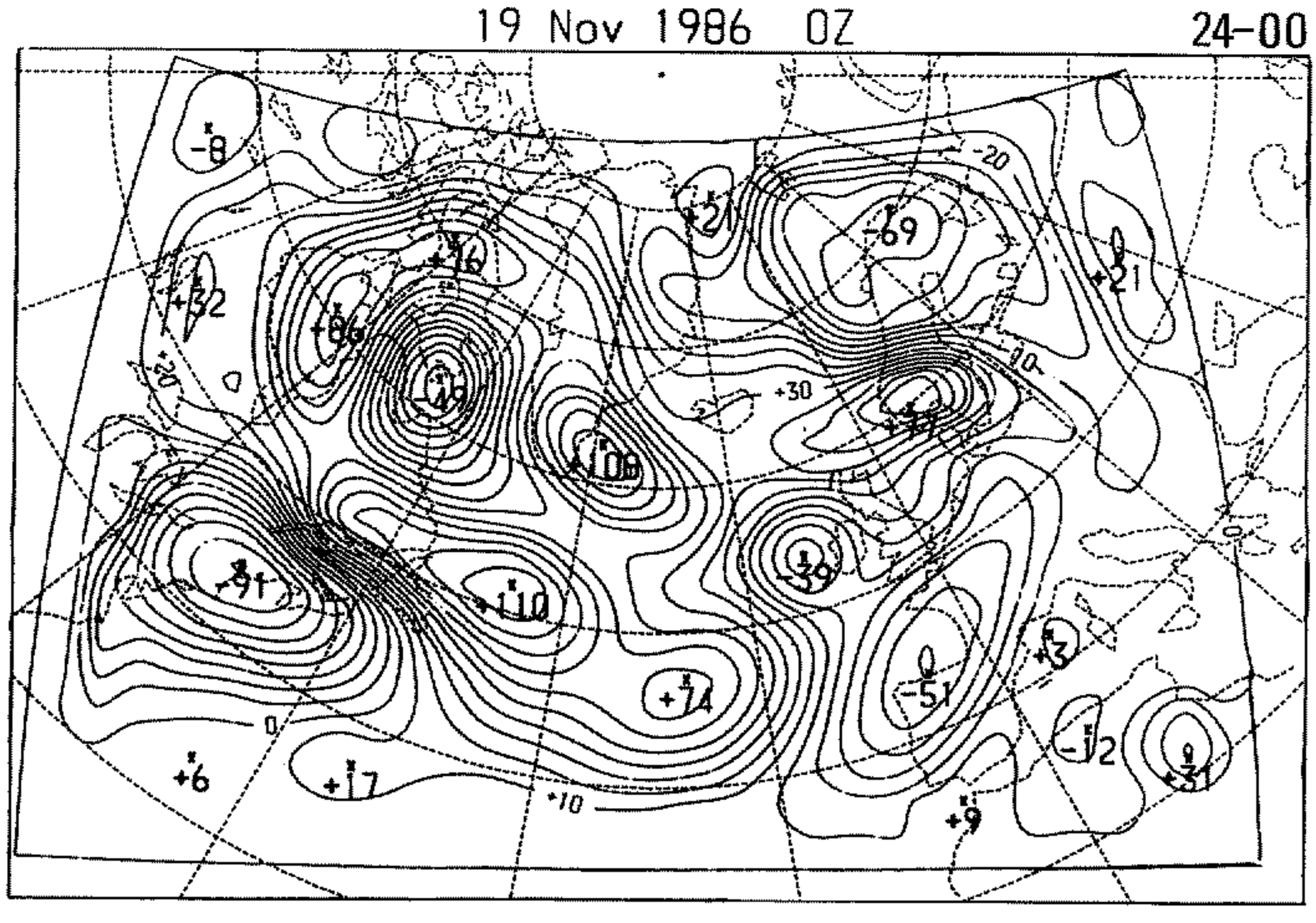

Figure 3. Forecast 500 hPa geopotential height change. This is the difference between the fields in Figs. 1 and 2.

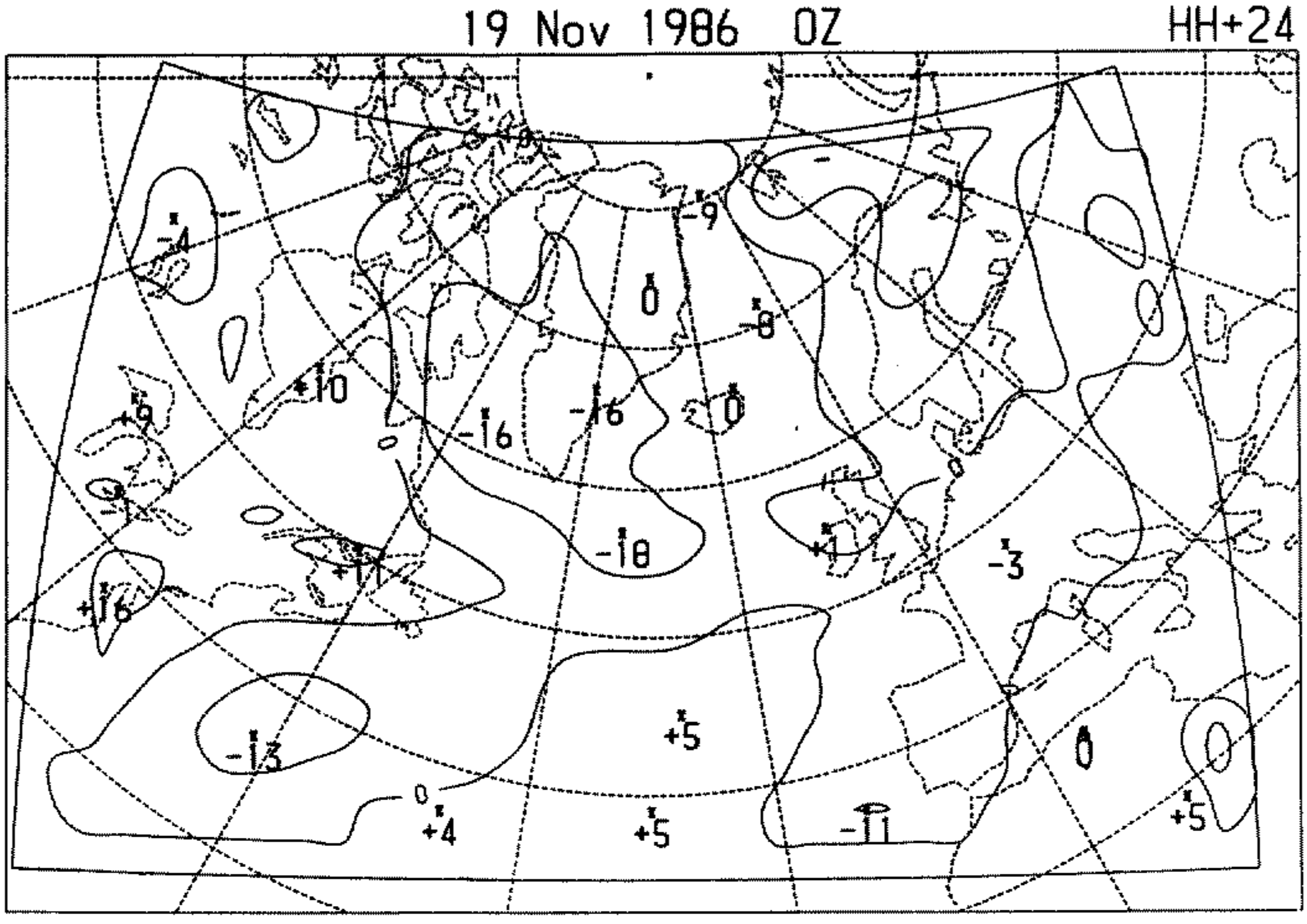

Figure 4. Difference between the 500 hPa height forecasts made with the primitive equation and slow equation models [PE] and [SE].

low frequency rotational components of the flow, which are of primary meteorological significance and interest. In the baroclinic case the separation between rotational and gravity components is less distinct and there are convergence problems associated with some initialization methods. Preliminary comparisons between baroclinic versions of the slow and primitive equation models have been made. These indicate that the slow equations are capable of accurately modelling baroclinic systems. However, the question of how these equations may simulate small-scale systems, such as fronts, is of fundamental importance and must be examined before the ultimate usefulness of the equations can be gauged.

6. Initialization and the slow system

The data for a primitive equation model must be initialized if spurious gravity wave noise is to be avoided. In the [PE] model, the divergence damping reduces this noise after about 12 hours, but there is a severe initial shock and large unrealistic oscillations during the early forecast hours. The slow equations have no solutions corresponding to gravity waves, and are free from these problems. The diagnostic components of the system automatically ensure that a balance between the mass and wind fields is maintained at all times. In this sense, the slow equations are self-initializing. The [SE] model may also be used to initialize the data for the primitive equation model. However, the evolution of the flow in the [SE] forecast is smoother than in [PE], even when the latter starts from initialized data.

The different character of the flow evolution produced by the primitive and slow models is clearly shown in Fig. 5. These graphs show the height at a central point

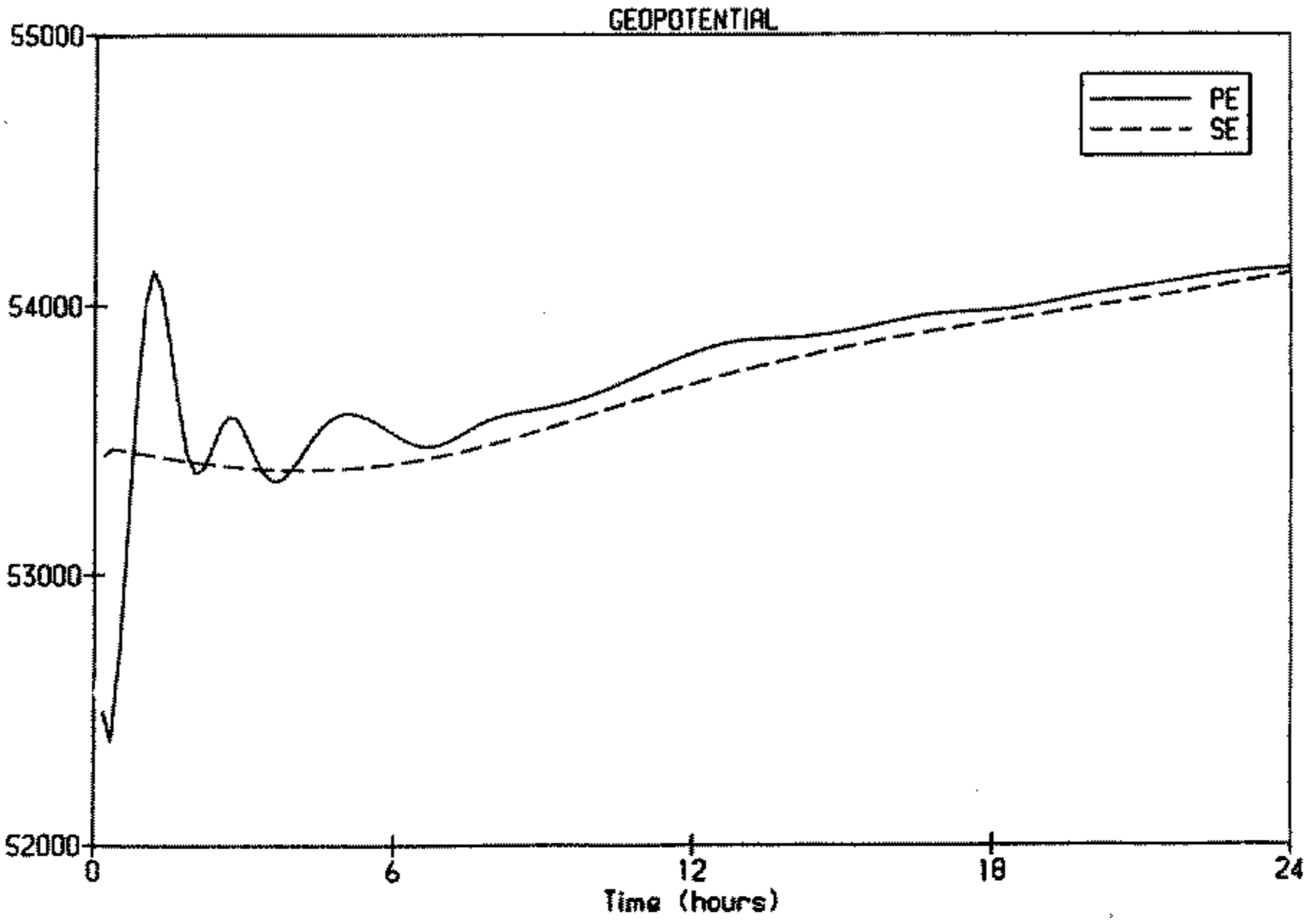

Figure 5. Geopotential height forecast for a central point (I=19, J=9) produced by the primitive equation model [PE] (solid line) and the slow equation model [SE] (dashed line).

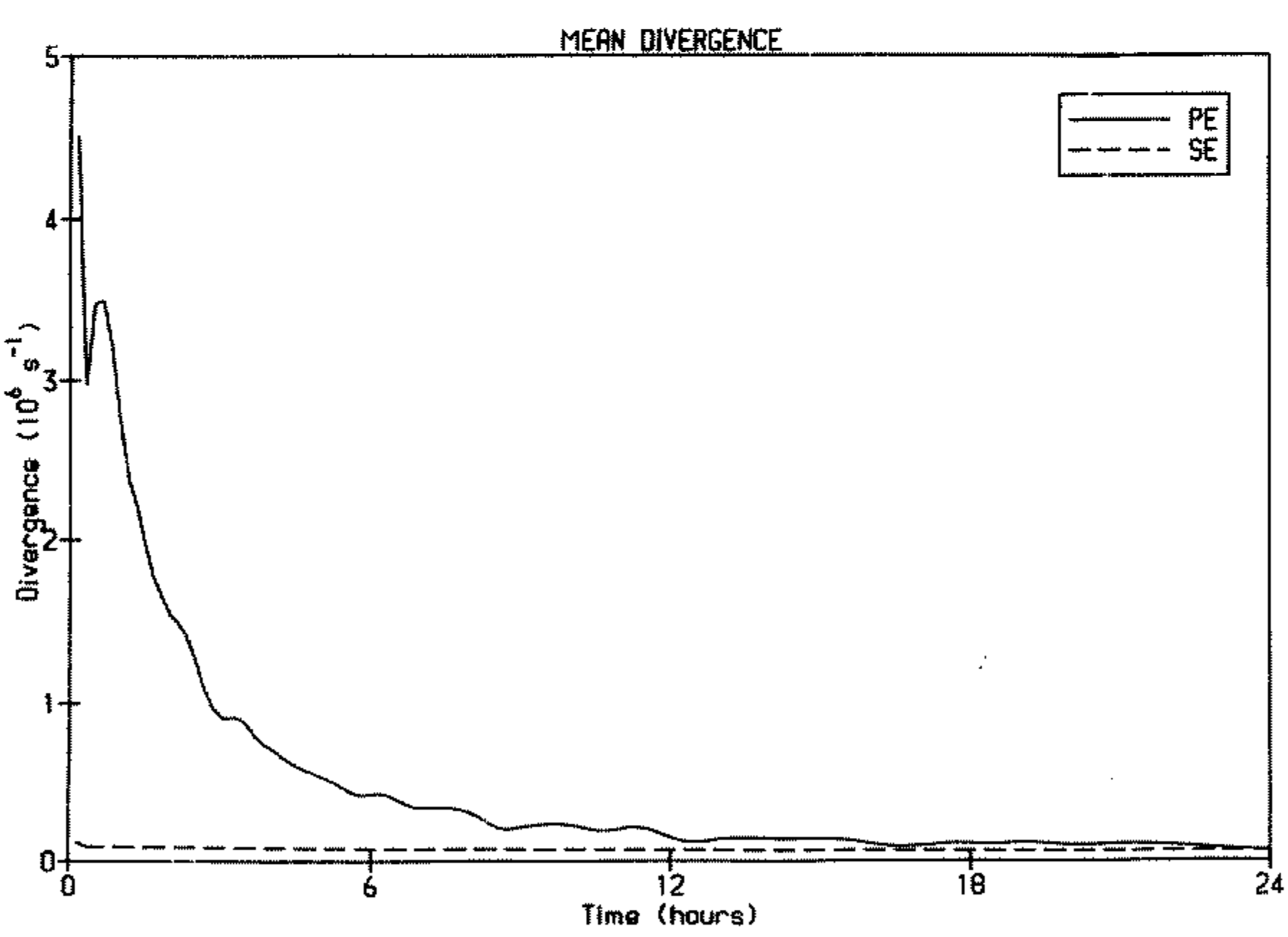

Figure 6. Mean absolute divergence for the forecast produced by the primitive equation model [PE] (solid line) and the slow equation model [SE] (dashed line).

($I = 19, J = 9$) plotted against time. The noisy character of the [PE] forecast is abundantly clear, and is in sharp contrast to the smooth curve resulting from the [SE] integration. A similar plot of mean divergence, shown in Fig. 6, confirms the noisy character of the [PE] run, and further demonstrates how the slow system produces a smooth evolution, free from the initial shock and subsequent noise.

The slow equations may be used to initialize the data for the [PE] model. This is done by integrating the initial data, using [SE], for one or more timesteps, but with $\Delta t = 0$, i.e. omitting the semi-Lagrangian step. This is completely equivalent to the implicit normal mode method of initialization (Temperton 1985; Juvanon du Vachat 1986). The number of 'zero' timesteps is equivalent to the number of iterations of the normal mode technique. In practice, two iterations are sufficient to initialize the data. Forecasts from uninitialized data are denoted by NIL, and those from data after two zero timesteps are indicated by NL2.

The evolutions of the height field at a central point (I=26, J=12) for the [PE] model *without* and *with* initialization are shown in Fig. 7. The absence of an initial shock in the NL2 run is clear. However, there is evidence that some high frequency oscillations remain. In Fig. 8 the NL2 run is replotted with an expanded vertical scale (it is denoted PRIM), together with the values forecast by the [SE] model (denoted SLOW). It is clear that the latter produces a smoother evolution. The residual noise in [PE] was not removed by further iterations of the initialization. Use of a higher order scheme (Tribbia 1984) might well remove this noise. However, there seems to be an inherent tendency for the [PE] model to produce some high frequency components.

In Table 1 the differences between the uninitialized and initialized runs, at various times during the forecasts, are presented. The differences after 24 hours are significantly smaller than the initial changes. These results are broadly comparable to the results using the Laplace transform technique of initialization (Lynch 1985, Table 2).

The complete absence of high frequency noise from the [SE] forecast may be of great advantage if the model is to be used for data assimilation. The insertion of observational data which are not completely compatible with the forecast state may give rise to large oscillations in a primitive equation model. The resulting noise causes problems for subsequent data induction cycles. The slow equation model should be free from these problems. A number of experiments in the assimilation of inserted data, using the [PE] and [SE] models, are discussed in detail in Lynch (1987).

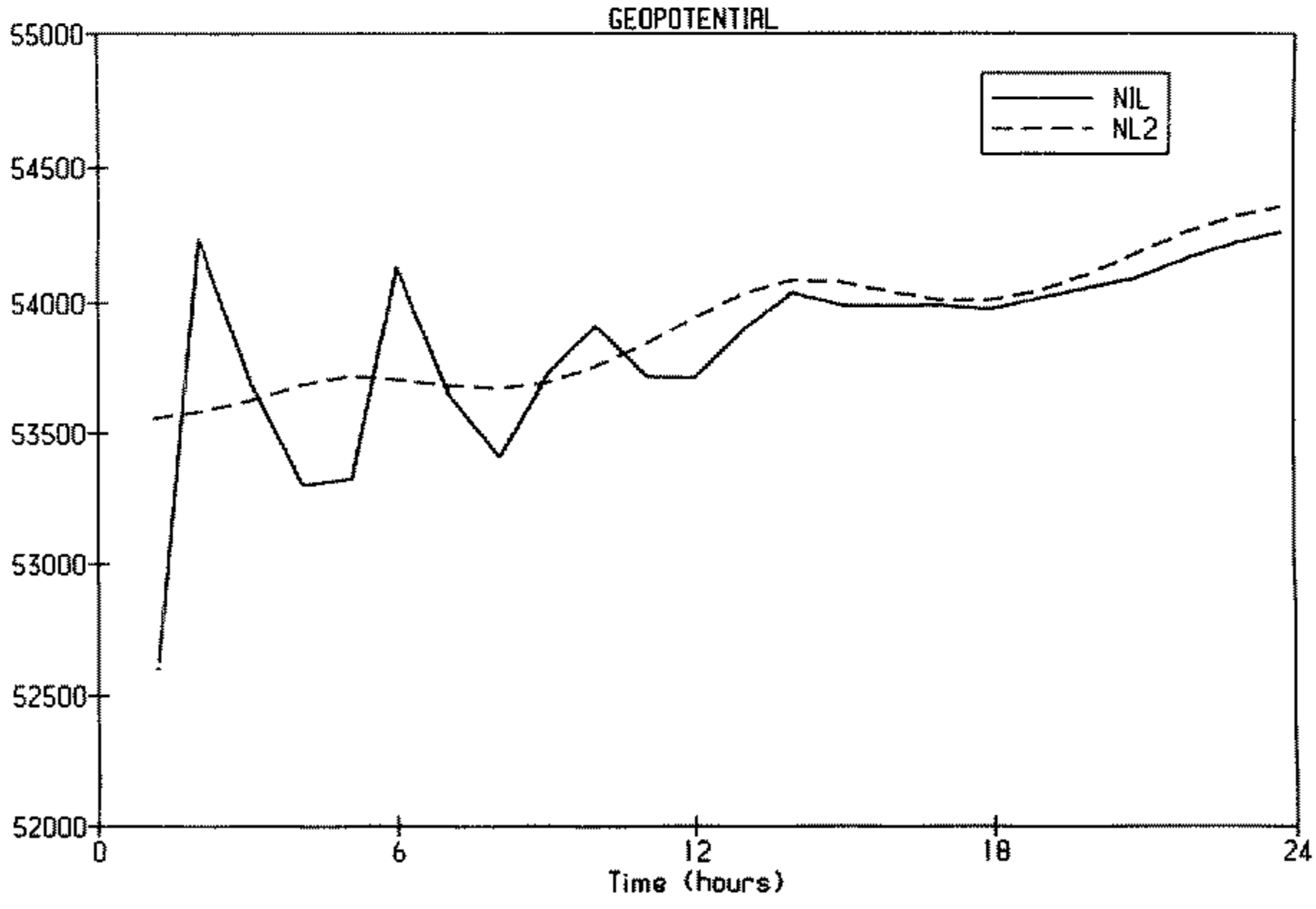

Figure 7. Geopotential height forecast for a central point (I=26, J=12) for the primitive equation model [PE] starting from uninitialized data (solid line) and data initialized with two nonlinear iterations (dashed line).

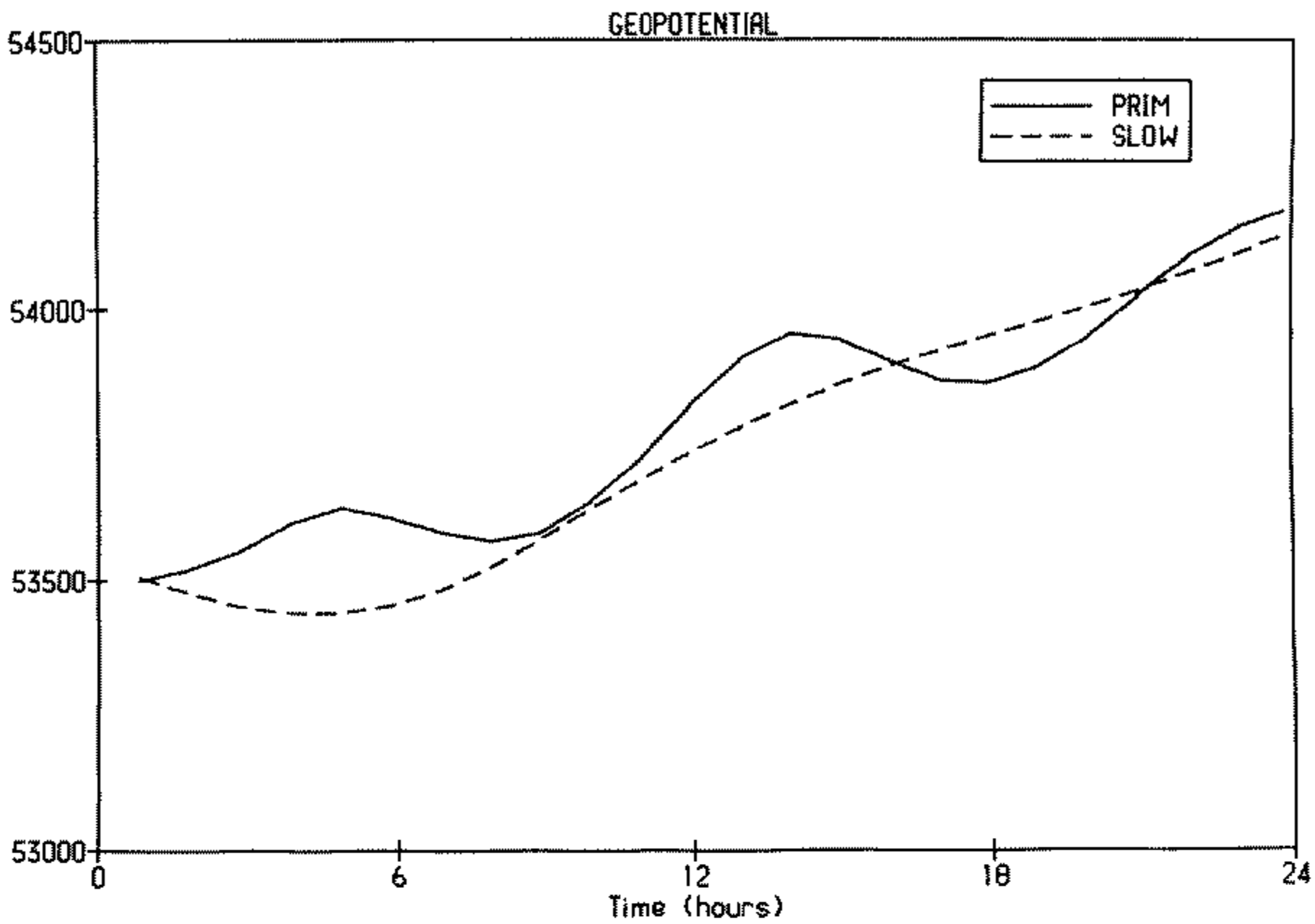

Figure 8. Geopotential height forecast for a central point (I=26, J=12) for the primitive equation model [PE] starting from data initialized with two nonlinear iterations (solid line) and for the slow equation model [SE] with uninitialized data (dashed line).

TABLE 1. ROOT-MEAN-SQUARE (AND MAXIMUM) DIFFERENCES IN HEIGHT AND WIND BETWEEN THE ORIGINAL AND INITIALIZED FIELDS (NL2 – NIL) AND BETWEEN THE 12- AND 24-HOUR FORECASTS RESULTING FROM THESE FIELDS.

Forecast	z (m)	u ($m\,s^{-1}$)	v ($m\,s^{-1}$)
HH+00	5·36 (32·12)	2·11 (12·35)	2·08 (13·77)
HH+12	4·29 (11·78)	0·75 (4·52)	0·64 (5·10)
HH+24	2·34 (8·52)	0·54 (3·12)	0·56 (5·10)

7. CONCLUSIONS

A filtered system, the slow equation system, has been derived using the ideas of normal mode initialization. The prognostic element of the system is an equation expressing conservation of potential vorticity. The remaining components, the balance and imbalance equations, are diagnostic. The system models the low frequency rotational motions and is free from gravity wave solutions.

The slow equations may also be derived by using scaling assumptions similar to those used to obtain the general balance system. There is a difference at $O(Ro^2)$ between the systems. The balance system has spurious unphysical solutions; the slow equations are free from this problem.

The slow system is equivalent to a filtering procedure originally proposed by Hinkelmann (1969) in which diagnostic equations are derived by omission of the first and second time derivatives of the divergence.

The slow system is ideally suited to the application of the semi-Lagrangian method, since the only prognostic element is a pure conservation equation (the adjustment process is diagnostic). Comparison runs show that barotropic integrations with the primitive and slow equations give very similar results. Thus, the approximations made in deriving the slow system result in negligible error.

The slow system is self-initializing. It may also be used to initialize a primitive equation model. This method is identical to the implicit normal mode method (Temperton 1985; Juvanon du Vachat 1986).

Since the slow system is noise-free, it may be used to advantage for data assimilation. The balance equation is an inherent component of the slow system, and may conveniently be used to ensure that inserted data project onto the rotational modes; the imbalance equation ensures an appropriate divergence field. Data insertion experiments are discussed in detail in Lynch (1987).

Acknowledgments

I am grateful to my colleagues Ray Bates, Jim Hamilton and Aidan McDonald for numerous illuminating discussions, and to Jim Logue for a perceptive review of this paper.

References

Bates, J. R. and McDonald, A. 1982 Multiply-upstream semi-Lagrangian advective schemes: analysis and application to a multi-level primitive equation model. *Mon. Weather Rev.*, **110,** 1831–1842

Bourke, W. and McGregor, J. L. 1983 A nonlinear vertical mode initialization scheme for a limited area prediction model. *ibid.*, **111,** 2285–2297

Browning, G, L. and Kreiss, H.-O. 1987 Reduced systems for the shallow water equations. *J. Atmos. Sci.*, **44,** 2813–2822

Burger, A. P. 1958 Scale consideration of planetary motions of the atmosphere. *Tellus*, **10,** 195–205

Charney, J. G. 1962 'Integration of the primitive and balance equations'. Pp. 131–152 in *Proc. Intl. Symp. on NWP*. Japanese Meteorological Agency

1973 'Planetary fluid dynamics'. Pp. 97–351 in *Dynamic meteorology*. Ed. P. Morel, D. Reidel

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

Daley, R. 1980 The development of efficient time integration schemes using normal mode models. *Mon. Weather Rev.*, **108,** 100–110

1982 A non-iterative procedure for the time integration of the balance equations. *ibid.*, **110,** 1821–1830

Hinkelmann, K. H. 1969 'Primitive equations'. Pp. 306–375 in *Lectures in numerical short-range weather prediction.* Regional Training Seminar, Moscow. WMO No. 297

Juvanon du Vachat, R. 1986 A general formulation of normal modes for limited area models; application to initialization. *Mon. Weather Rev.*, **114,** 2478–2487

Lynch, P. 1985 Initialization of a barotropic limited-area model using the Laplace transform technique. *ibid.*, **113,** 1338–1344

1987 'The slow equations. Part I: Derivation and properties of the system. Part II: Application to continuous data assimilation' Tech. Note No. 50, Irish Meteorological Service, Dublin

1988 Deducing the wind from the vorticity and divergence. *Mon. Weather Rev.*, **116,** 86–93

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

McDonald, A. 1986 A semi-Lagrangian and semi-implicit two-time-level integration scheme. *Mon. Weather Rev.*, **114,** 824–830

McDonald, A. and Bates, J. R. 1987 Improving the estimate of the departure point in a two-time-level semi-Lagrangian and semi-implicit scheme. *ibid.*, **115,** 737–739

Moura, A. D. 1976 The eigensolutions of linearized balance equations over a sphere. *J. Atmos. Sci.*, **33,** 877–907

Richardson, L. F. 1922 *Weather prediction by numerical process.* Cambridge Univ. Press, London (Reprinted by Dover, New York, 1965)

Sangster, W. E. 1960 A method of representing the horizontal pressure force without reduction of pressures to sea level. *J. Meteorol.*, **17,** 166–176

Temperton, C. 1985 'Applications of a new principle for normal mode initialization'. Pp. 105–107 in Preprint volume, Seventh conference on numerical weather prediction, June 17–20 1985, Montreal, P.Q., American Meteorological Society

1988 Implicit normal mode initialization. *Mon. Weather Rev.*, **116,** 1013–1031

Thompson, P. D. 1980 A short-range prediction scheme based on conservation principles and the generalized balance condition. *Beitr. Phys. Atmos.*, **53,** 256–263

Tribbia, J. J. 1984 A simple scheme for high-order nonlinear normal mode initialization. *Mon. Weather Rev.*, **112,** 278–284