JOURNAL OF COMPUTATIONAL PHYSICS 50, 1–37 (1983)

Review Article

Current Progress and Prospects in Numerical Techniques for Weather Prediction Models

M. J. P. CULLEN

Forecasting Research Branch, Meteorological Office, Bracknell, Berkshire, England

Received May 4, 1982

Progress in numerical weather prediction since about 1977 is reviewed. The development of larger and faster computers now allows a global forecast to be made in three minutes per day using a 150 km grid in the horizontal, 15 levels in the vertical, and one of the very efficient integration schemes now available. The forecasts produced are useful guidance to forecasters to about four days on average, but there is a large difference in performance from case to case. In this review recent developments in understanding the solutions to the governing equations are discussed, some of the efficient integration schemes are described in detail, and an example of the current standard of forecasts is displayed.

1. INTRODUCTION

This paper reviews recent progress in numerical methods used in meteorology. It takes as a starting point the reviews by Kasahara and by Arakawa and Winninghof in [1], published in 1977, and the textbook by Haltiner and Williams [2], published in 1980. Since these reviews were prepared, larger and faster computers employing vector and parallel processing architecture, notably the CRAY-1 and CYBER 205, have come into use and the numerical models used in meteorology have expanded accordingly. Around 1975, models used for operational forecasts for up to three days ahead used grid lengths of 300–400 km and covered areas smaller than the Northern hemisphere. Today, models with grid lengths of 150 km covering the whole globe can be used operationally, taking only a few minutes of computing time per model day. Similar improvements can be made in other meteorological models. Limited area models used for rainfall prediction can be run operationally with grid lengths of 50 km and extensive research is being carried out with "mesoscale" models with grid lengths as small as 1 km. On a larger scale, climate simulation experiments requiring integration periods of several years can now be run using models of a size that could only be afforded for short range forecasts in 1975.

Associated with these increases in computing power have come further developments in numerical techniques and analysis. At the time of the review in [1] it

0021-9991/83 $3.00

was believed that the largest single cause of error (about 35%) in short range forecasting models was due to truncation error in the calculation of horizontal derivatives. Improvements in grid resolution since then have substantially reduced these errors. Thus while with a grid length of 400 km substantially better results are obtained using either staggering of the variables on the mesh [3] or fourth-order differencing in the advection terms [4], when these changes are made using a 200 km grid the effect on forecast errors is much less. Studies made at several institutes, e.g., [5, 6], show only a marginal improvement when the gridlength is reduced below 200 km. Similarly, the intercomparison of forecasts from different models of Cullen *et al.* [7] did not show markedly better skill than the earlier set of results from lower resolution models given by Baumhefner and Downey [8]. Though the models in [7] showed increased amplitudes for the weather systems, errors in position were not greatly altered and the overall mean error statistics reported in [7] are no better than in [8]. This is consistent with the analysis of the causes of errors in a 48 hour forecast by Robert [9]. His conclusion was that for second-order finite differencing on a 400 km grid, horizontal truncation errors accounted for 35% of the total. When the grid is refined to 200 km the errors from this cause should decrease by a factor of four, and only represent 12% of the total, while effects of inadequate initial data might rise from 18% to 25%. In addition, while sensitivity experiments using models with different resolutions and different formulations have shown diminishing returns in improving forecast skill, consolidation of and improvements to all parts of the models in the current generation of operational forecasts have resulted in useful improvements over the models available in 1975. For instance, recent skill scores from the European Centre for Medium Range Weather Forecasts (ECMWF) (Fig. 1) show a marked improvement in the forecast statistics over those quoted in [1]. In addition to improved statistics, these forecasts are assessed as useful guidance out to

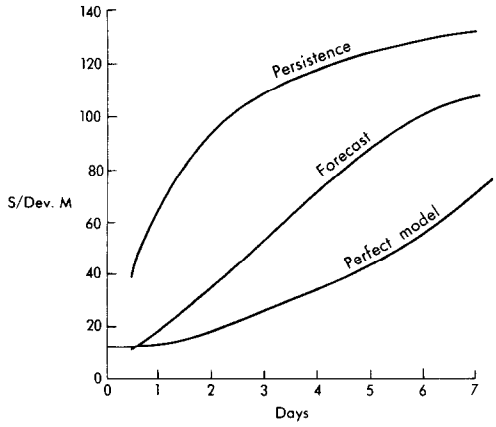

FIG. 1. Standard deviations of 500 mb error for European area (36°–72° N, 12° W–42° E) during 1980 from ECMWF (Source: ECMWF Forecast Reports for 1980).

four days on most occasions [10]. Examples of the current standard of prediction are shown in a later section.

It is now necessary, however, to reassess the current problems in numerical forecasting. The predominant causes of error may no longer be those in [9]. The first question to be answered is whether the limit of predictability will be reached as soon as the quality of the initial data is improved. This question can not yet be answered. If not, it is necessary to decide how further improvements can be made. Ideas about this are reviewed in later sections.

The plan of the rest of the paper is as follows: In Section 2 the equations in most common use are set out. The current state of theory on existence and uniqueness of solutions to such equations is reviewed. Though this full system has not been analysed, several other systems of equations describing fluid flow have been, and the results of these studies and their relevance to the atmosphere are reviewed here. Such studies also give guidance on the choice of appropriate numerical methods and these are summarised. It is well known that the flow in the atmosphere is in geostrophic balance almost everywhere, because of the Earth's rotation, with small regions which are highly ageostrophic. In other branches of physics similar situations are common, for instance high Reynolds number flows behave inviscidly almost everywhere but viscosity is dominant in small regions. In such a situation numerical methods should take this structure into account. Up to the present time little progress has been made in taking advantage of this behaviour in the atmosphere; the most relevant work is reviewed in this section.

In Section 3 the problem of constructing appropriate initial data is discussed. This has been a major topic of research since the review in [1]. The difficulty is that the governing equations can describe a wide range of motions, from sound waves propagating at $300\ \mathrm{m\ sec^{-1}}$ to large scale Rossby waves propagating at around $10\ \mathrm{m\ sec^{-1}}$. Much discussion has centred on the existence of a "slow manifold" of solutions [11] describing only meteorological motions rather than sound or gravity waves. This concept is highly questionable mathematically, but is a convenient way of discussing the initialisation problem. Raw observational data may not satisfy the necessary balance between wind and pressure observations, and so the data has to be "projected" onto this manifold. For instance, insertion of observed surface pressure values into a model results in exciting only fast-moving gravity waves unless geostrophically balanced winds are also inserted.

In Section 4 developments in the numerical solution of the forecast equations are discussed. Two very efficient integration schemes are described; the semi-implicit scheme used in the ECMWF model [12] and the split explicit scheme developed by Gadd [13–15] and used in the United Kingdom operational model. The use of spectral models has become more widespread since 1977; however, they are not discussed in detail here because the techniques differ little from those described by Bourke *et al.* [16]. Various other developments are also described, for instance in the treatment of the upper boundary and in the use of hybrid coordinate systems in the vertical [17].

As well as solutions of the dynamic equations, considerable effort has been put

into representing sub-grid-scale motions, diabatic effects such as radiation, and effects of moisture. This part of the development work is a major portion of the current effort going into weather prediction research. Certain parts of it are discussed in Section 5—for instance, the representation of turbulence. However, much of the development work is more a matter of understanding the basic physics than of computational techniques and therefore is not dealt with in this paper.

In Section 6 an example of the current standard of numerical forecasts is presented. This is taken from a large number of cases and it is difficult to be fair in making a selection. Therefore, to give a broader impression of the standard, tables showing the forecast assessments over a longer period are also shown. Much more detailed assessments of model performance have been carried out in practice: A detailed review of the performance of the ECMWF model, for instance, has been published by Bengtsson and Simmons [71].

2. The Mathematical Problem

2.1 *Equations of Motion*

The derivation of the various systems of equations in normal use is set out in [1, 2] and is therefore not repeated here. We describe only the most common system and mention alternatives briefly. The equations are written in spherical polar coordinates (λ, θ) for the flow parallel to the Earth's surface and in "sigma" coordinates for the flow perpendicular to the Earth's surface. The "vertical" coordinate σ is defined as p/p_*, where p is the pressure and p_* the surface pressure. Under this system $\sigma = 0$ is the top of the atmosphere and $\sigma = 1$ the Earth's surface. Various modifications or alternatives to this system are discussed later. The rest of the notation is as follows:

(u, v)	wind components parallel to Earth's surface in spherical polar coordinates
T	temperature
ϕ	geopotential
q	specific humidity
T_v	virtual temperature $= (1 + 0.61q)\,T$
(U, V)	$= (p_* u, p_* v)$
$\dot{\sigma}$	$= \mathbf{u} \cdot \nabla_h \sigma$, vertical velocity in σ coordinate
ω	$= \partial p/\partial t + \mathbf{u} \cdot \nabla_h p$, vertical velocity in p coordinate
∇_h	gradient on horizontal surface
$(\tau_\lambda, \tau_\theta)$	surface stress components
g	acceleration due to gravity
f	$= 2\Omega \sin\theta$ Coriolis parameter (Ω is angular velocity of Earth's rotation)
R	gas constant
C_p	specific heat at constant pressure
a	radius of Earth
κ	$= R/C_p$ for air
H	vertical heat flux

S diabatic heating
L latent heat of evaporation/fusion
P rate of precipitation of rain/snow
M vertical moisture flux
K coefficient of eddy diffusivity.

The equations of motion with the hydrostatic approximation made are then

$$\frac{\partial U}{\partial t}+\frac{1}{a\cos\theta}\left[\frac{\partial}{\partial\lambda}(Uu)+\frac{\partial}{\partial\theta}(Vu\cos\theta)\right]+\frac{\partial}{\partial\sigma}(p_*\dot{\sigma}u)$$
$$-V\left(f+\frac{u\tan\theta}{a}\right)+\frac{1}{a\cos\theta}\left(p_*\frac{\partial\phi}{\partial\lambda}+RT_v\frac{\partial p_*}{\partial\lambda}\right)=K\nabla\cdot p_*\nabla u+g\frac{\partial\tau_\lambda}{\partial\sigma}; \tag{2.1}$$

$$\frac{\partial V}{\partial t}+\frac{1}{a\cos\theta}\left[\frac{\partial}{\partial\lambda}(Uv)+\frac{\partial}{\partial\theta}(Vv\cos\theta)\right]+\frac{\partial}{\partial\sigma}(p_*\dot{\sigma}v)$$
$$+U\left(f+\frac{u\tan\theta}{a}\right)+\frac{1}{a}\left(p_*\frac{\partial\phi}{\partial\theta}+RT_v\frac{\partial p_*}{\partial\theta}\right)=K\nabla\cdot p_*\nabla v+g\frac{\partial\tau_\theta}{\partial\sigma}; \tag{2.2}$$

$$\frac{\partial\phi}{\partial\sigma}+\frac{RT_v}{\sigma}=0; \tag{2.3}$$

$$\frac{\partial p_*}{\partial t}+\frac{1}{a\cos\theta}\left[\frac{\partial U}{\partial\lambda}+\frac{\partial}{\partial\theta}(V\cos\theta)\right]+\frac{\partial}{\partial\sigma}(p_*\dot{\sigma})=0; \tag{2.4}$$

$$\frac{\partial}{\partial t}(p_*T)+\frac{1}{a\cos\theta}\left[\frac{\partial}{\partial\lambda}(UT)+\frac{\partial}{\partial\theta}(VT\cos\theta)\right]+\frac{\partial}{\partial\sigma}(p_*\dot{\sigma}T)-\frac{\kappa T\omega}{\sigma}$$
$$=K\nabla\cdot p_*^{\kappa+1}\nabla\left(\frac{T}{p_*}\right)+\frac{g}{C_p}\frac{\partial H}{\partial\sigma}+\frac{p_*}{C_p}[S+LP]; \tag{2.5}$$

$$\frac{\partial}{\partial t}(p_*q)+\frac{1}{a\cos\theta}\left[\frac{\partial}{\partial\lambda}(Uq)+\frac{\partial}{\partial\theta}(Vq\cos\theta)\right]+\frac{\partial}{\partial\sigma}(p_*\dot{\sigma}q)$$
$$=K\nabla\cdot p_*\nabla q+g\frac{\partial M}{\partial\sigma}-p_*P. \tag{2.6}$$

The term LP in (2.5) take a more complicated form in most models because of the effects of freezing/melting and reevaporation of precipitation. The top and bottom boundary conditions usually used are

$$\dot{\sigma}=0 \qquad \text{at} \quad \sigma=0,\,1.$$

Other possibilities are discussed later.

2.2 *Solutions to the Equations of Motion*

In this section we discuss the mathematical properties of (2.1)–(2.6) and related systems. The relevant theoretical work falls naturally into two parts. The first concerns solutions to the associated initial-boundary value problem—using techniques extended from linear theory to smooth solutions of nonlinear problems. The second is concerned with the formation of singularities in the solution of the inviscid form of the equations and how the correct solution to the physical problem with very small dissipation can be obtained.

The first part of the theory is described in detail by Oliger and Sundstrom, [18] and will only be summarized here. It describes under which boundary conditions a unique solution of (2.1)–(2.6) and related systems of equations exists. The proofs use ideas from linear equations with constant coefficients. These ideas have been generalised to apply to nonlinear equations, provided that the solutions stay smooth, by Kreiss [19]. The effect of this assumption of smoothness has been studied by Kreiss and Browning [20], who estimate for how long solutions remain smooth given smooth initial data.

Oliger and Sundstrom [18] discuss the existence and uniqueness of solutions to the meteorological equations in the context of the initial-boundary value problem. They write the equations in the general form

$$\mathbf{w}_t + \varepsilon A(\mathbf{w}, \mathbf{w}_x)\, \mathbf{w}_{xx} + B(\mathbf{w})\, \mathbf{w}_x + C(\mathbf{w}) = 0, \tag{2.7}$$

where A, B, and C are matrix operators and ε is a small parameter. When considering the boundary value problem, in which artificial lateral boundaries are imposed to restrict the area of computation, it is necessary to avoid the formulation of spurious computational boundary layers. The desired solution is that which would be obtained if the boundaries were absent. This means that it is necessary to find boundary conditions which yield a unique solution to the inviscid system

$$\mathbf{w}_t + B(\mathbf{w})\, \mathbf{w}_x + C(\mathbf{w}) = 0. \tag{2.8}$$

This problem is analysed by the methods originally developed for linear constant coefficient problems using the theories of Kreiss [19]. It can be shown that the undifferentiated terms do not affect the existence and uniqueness properties, so that on this count the presence of rotation does not affect the analysis. If the system

$$\mathbf{w}_t + B(\mathbf{w})\, \mathbf{w}_x = 0 \tag{2.9}$$

is first-order hyperbolic, then the required boundary conditions can be identified. Unfortunately, (2.1)–(2.6) is not hyperbolic in the inviscid case ($K = 0$ throughout). This is because of the hydrostatic approximation. The inviscid nonhydrostatic equations set out in [1] are hyperbolic and so the appropriate boundary conditions can be identified. This is also true for the shallow water equations, which can be regarded as an approximation to the vertical average of (2.1)–(2.6). It is simply necessary to identify the outgoing and incoming characteristics, and to supply a value for each variable on an incoming characteristic. Though these difficulties with the

hydrostatic equations apply only to the case of limited regions, they are also likely to cause problems in global integrations if there are implied computational boundaries to certain scales of motion due to variable resolution. Similar difficulties arise if the nonhydrostatic equations are used together with implicit time integration schemes to handle vertically propagating sound waves. The system of discrete equations obtained can be regarded as an approximation to a filtered system of equations which is not hyperbolic.

A limitation of this work is that it only proves the existence of a solution while it remains smooth. It is then necessary to find out whether and for how long it does stay smooth. This has been studied by Kreiss and Browning [20] for the shallow water equations. They use techniques extended from the linearised problem. If (2.1)–(2.6) is linearised about a state of uniform motion and the viscous and diabatic terms are omitted, two types of solution are obtained, horizontally propagating gravity waves with phase speeds up to 300 m $\sec^{-1}$ and waves propagating with the advection speed. In the atmosphere this is almost always less than 50 m $\sec^{-1}$. If the initial data is prepared so as to exclude gravity waves and contains waves with wavelength comparable to the Earth's radius, the solution remains smooth for the order of ten days. Estimates have also been made by Klainerman and Majda [21] for data including only gravity waves. The results of [20] are supported by computations by Sadourny [22] with the shallow water equations on a sphere, which show smooth solutions for about 15 days. After this time an "energy catastrophe" occurs in some of his integrations. It is known that solutions of the two-dimensional incompressible Euler equations

$$\partial \mathbf{u}/\partial t + \mathbf{u} \cdot \nabla \mathbf{u} + (1/\rho)\nabla p = 0, \qquad \nabla \cdot \mathbf{u} = 0 \tag{2.10}$$

stay as smooth as the initial data for all time [23]. This system is equivalent to the barotropic vorticity equation

$$\partial \zeta/\partial t + \mathbf{u} \cdot \nabla(\zeta + f) = 0, \qquad \mathbf{u} = \nabla \times \zeta \hat{\mathbf{k}}, \tag{2.11}$$

where ζ is the scalar vorticity and $\hat{\mathbf{k}}$ a unit vector in the third dimension. This system is the simplest model for atmospheric flows. Two similar results have been proved for three-dimensional quasi-geostrophic equations, written here in terms of a stream function ψ and Cartesian coordinates (x, y, z)

$$\begin{aligned} u &= -\psi_y, \qquad v = \psi_x, \\ Q &= \psi_{xx} + \psi_{yy} + \rho^{-1}(\rho\alpha\psi_z)_z, \\ DQ/Dt &= S, \\ \Theta &= \psi_z, \\ w &= -\alpha(D\Theta/Dt - H), \\ w &= 0 \qquad \text{at} \quad z = 0, h, \\ \mathbf{u} \cdot \mathbf{n} &= 0 \qquad \text{on lateral boundaries.} \end{aligned} \tag{2.12}$$

Here α is the specific volume and S and H are source terms for the potential vorticity Q and potential temperature Θ. The results are described by Bennett and Kloeden [24]. They show that solutions of (2.12) remain as smooth as the initial data for all time if the potential temperature is specified on $z = 0$ and h, and if the flow is contained in the x and y directions. It stays as smooth as the data for periodic boundary conditions in x and y with rigid boundaries at $z = 0$ and h for a finite time T inversely proportional to the gradients in the initial conditions. It has been shown by Williams [25] that discontinuities can form in the solutions of (2.12) with other choices of boundary conditions.

The next system that has been analysed is the semi-geostrophic system discussed by Hoskins [26]. In this system the momentum components U and V in (2.1) and (2.2) are replaced by their geostrophic values $-(fa)^{-1} p_*(\partial\phi/\partial\theta)$ and $(fa\cos\theta)^{-1} p_*(\partial\phi/\partial\lambda)$. It has been shown that this system, without viscous terms, forms discontinuities in velocity with associated infinite vorticities. The crucial difference is that a term $(\zeta + f)\hat{\mathbf{k}} \cdot \nabla\mathbf{u}$ appears in the semi-geostrophic vorticity equation, rather than just $f\hat{\mathbf{k}} \cdot \nabla\mathbf{u}$ as in the quasi-geostrophic system.

Once we move away to the full inviscid system, the evidence is that the solutions will still break down in a finite time. The remark in [18] that the presence of rotation does not affect whether the solutions stay smooth, but only the length of time that discontinuities take to form, means that we can consider various results obtained in the nonrotating case. The shallow water equations in one dimension form a standard example of a discontinuous solution, the hydraulic jump. In more than four dimensions, solutions of these equations with finite initial energy stay smooth for almost all data (Klainerman [27]). The important cases for meteorology are in two and three dimensions where least is known. John [28] has shown that discontinuities are likely to form in the three-dimensional case.

These results apply to purely hyperbolic systems of wave equations. However, (2.1)–(2.6) do not reduce to such a system because of the hydrostatic approximation, (2.3). Instead of (2.3), we can make the anelastic approximation

$$\nabla \cdot (\rho\mathbf{u}) = 0. \tag{2.13}$$

Since the Mach numbers are small, some aspects of the solutions may be similar to those of the three-dimensional incompressible Euler equations (2.10). It has been suggested by two separate numerical investigations, but not proved analytically, that the solution to these equations can give infinite vorticity in a finite time (Morf *et al.* [29], Chorin [30]. The physical effect is that vortex lines are stretched an infinite amount in a finite time. Chorin [31] has also shown how the solution can evolve before the blowup. In order for energy to be conserved, the vortex lines must become entangled so that there is cancellation of the induced velocity fields. This means that in the presence of viscosity, the dissipation is much more effective at preventing the blowup than a simple consideration of the balance between the stretching and viscous terms would suggest. This has to be taken into account when attempting to prove that

solutions of the three-dimensional Navier–Stokes equations stay smooth for all time for arbitrarily large Reynolds number.

Though it seems that the inviscid form of the atmospheric equations can yield infinite vorticities in a finite time, the physically dominant cause of breakdown is not yet clear. Different mechanisms are more important at different horizontal scales. The model described by Hoskins can predict discontinuities at the boundaries. The scaling assumptions in this model break down first where the Richardson number

$$\frac{g}{\Theta_0}\frac{\partial \Theta}{\partial z}\Big/\left(\frac{\partial v}{\partial z}\right)^2$$

becomes small. Under these conditions turbulent mixing occurs. This mixing may be well described by the incompressible Euler equations (2.10), which also have a tendency to generate infinite vorticity. At the smallest scales, viscous dissipation will act, enhanced by the tangling of the vortex tubes.

2.3 *Consequences for Numerical Solution of the Equations*

The problem with obtaining numerical solutions of (2.1)–(2.6) is that the scales on which mixing occurs are several orders of magnitude smaller than those that can be resolved. This is a standard situation in mathematical modelling. It is necessary to be able to show that the inviscid equations possess a discontinuous or singular generalised solution which can be proved to be the limit as the viscosity tends to zero of a viscous solution. This has been shown for a wide range of one-dimensional problems in gas dynamics, combustion, and oil reservoir engineering (see, for instance, Chorin and Marsden [32]). In more than one dimension little is known even for the simplest problems. In cases where it has been possible to construct a singular solution to the inviscid equations which corresponds to observed physical behaviour, it has been necessary to choose a form of the equations which conserves the correct physical quantities and to impose an "entropy" condition to make the solution unique. It is then possible to approximate it numerically, taking care that the approximate solution cannot converge to a solution which violates the entropy condition. This approach has only been fully developed for one-dimensional problems, for instance by Majda and Osher [33]. In the meteorological case, where the discontinuities are much more complex, it is not yet known whether a physically valid discontinuous solution to the inviscid equations can be constructed. In the case of the three-dimensional incompressible Euler equations studied by Chorin, it does not seem likely that this approach is possible. Under present circumstances there is no alternative to adding artificial viscosity to the solution and hoping that the resulting smooth solution approximates the right physics.

There are, however, certain precautions that must be taken with the numerical method. The breakdowns of the solutions to inviscid forms of the equations all take a definite time to occur from given data. A numerical approximation to the solution should not exhibit a faster breakdown. For instance, consider approximations to

$$\partial u/\partial t = u(\partial u/\partial x) \tag{2.14}$$

Using the standard notation

$$\bar{A}^x = \tfrac{1}{2}(A(x+\Delta x) + A(x-\Delta x)),$$
$$\Delta_x A = A(x+\tfrac{1}{2}\Delta x) - A(x-\tfrac{1}{2}\Delta x),$$
$$\delta_x A = \Delta x^{-1}(A(x+\tfrac{1}{2}\Delta x) - A(x-\tfrac{1}{2}\Delta x));$$

if (2.14) is approximated by

$$\delta_t \bar{u}^t = u\delta_x \bar{u}^x, \tag{2.15}$$

it is known that the truncation error contains terms which form an approximation to the equation

$$\partial u/\partial t = u^2, \tag{2.16}$$

which has a solution that becomes infinite at $t = 1$ for initial data $u = 1$ (Gary [34]). The solution of (2.14) merely becomes discontinuous. Thus using (2.15) can lead to spurious rapid error growth. This is avoided by using the scheme

$$\delta_t \bar{u}^t = \overline{u\delta_x u}^x.$$

Similar schemes have been reviewed by Morton [35] and shown only to produce error growth when $\partial u/\partial x > 0$. This is reasonable, since even a good approximation should produce error growth when there is compression. Once the discontinuity in the solution of (2.14) has formed, artificial viscosity has to be added, explicitly or implicitly, in any case.

When solving systems such as (2.11) or (2.12), where the solution stays as smooth as the data, it is important to ensure that the numerical solution also does so. For Eq. (2.11), the crucial property is that the absolute vorticity is convected by the motion and can only take on values present in the initial data. This is how the mathematical existence proof of Kato [23] is constructed. Numerically this can only be ensured by using a Lagrangian method. In an Eulerian scheme the best that can be done is to conserve the total enstrophy, the integral of the square of the absolute vorticity. Since the grid is finite, this will ensure that the vorticity stays bounded, though not necessarily by its initial values, and that the solution will stay smooth. Thus, for solving (2.11) it is appropriate to use the enstrophy conserving schemes of Sadourny [22], and for solving (2.12) the potential enstrophy conserving scheme of Arakawa [36]. These schemes can also be used for (2.1)–(2.6) as convenient ways of suppressing spurious error growth, though any other scheme which suppresses it until the true inviscid solution breaks down should be equally suitable.

3. Initial and Boundary Conditions

3.1 *Initialisation*

There have been substantial recent advances in the techniques used to calculate initial data for forecasting models. Two major areas of development are the nonlinear normal mode initialisation technique of Baer and Tribbia [37] and Machenhauer [38], implemented in a large model by Temperton and Williamson [72]; and the bounded derivative method of Kreiss [40], as described in detail by Browning *et al.* [39]. The background to these methods has been discussed in the previous section. A much more detailed description of the normal mode method is given in [2]. It is assumed that it is possible to find a solution of (2.1)–(2.6) which essentially only contains waves moving with the advection speed and does not contain gravity waves. This hypothesis is often discussed in terms of a "slow manifold" (Leith [11]). It is unlikely that such a solution exists mathematically, except possibly as a singular solution. However, if no attempt is made to eliminate gravity waves from the data, any forecast is contaminated by large unrealistic oscillations. It is not sufficient just to eliminate all gravity waves from the initial data by a linear normal mode technique, since they can still be generated by the nonlinear terms which were neglected in calculating the normal modes. Therefore it is necessary to eliminate this initial growth of gravity waves as well. This can be done by an iteration technique, as described by Baer and Tribbia [37]. Alternatively, since the most troublesome gravity waves have a high frequency, they can be excluded by making the initial time derivatives of the variables up to a certain order small. This is the method proposed in [40]. Suppose we have a pair of linear equations which admit wave solutions of two different speeds:

$$\frac{\partial \mathbf{u}}{\partial t} = A\frac{\partial \mathbf{u}}{\partial x} + B(\mathbf{u}, \mathbf{u}); \qquad \mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}. \tag{3.1}$$

We suppose A has two eigenvalues λ_1, λ_2 with associated eigenfunctions e_1, e_2. Suppose $|\lambda_1| \gg |\lambda_2|$ and that λ_2 is the eigenvalue of physical interest. An arbitrary initial state $\mathbf{u}$ can be written in the form $\mathbf{u} = u_1 + u_2$ with $u_1 = \alpha_1 e_1$, $u_2 = \alpha_2 e_2$. The desired initial state for the linear problem would be $\alpha_2 e_2$. Define projections P_1 and P_2 by setting $P_1 u = u_1$, $P_2 u = u_2$. Seek a solution containing essentially only the second eigenvector for nonlinear problem (3.1). The normal mode approach takes $\alpha_2 e_2$ as the first guess to the desired initial state. Then, for a general iterate set

$$P_1\left(A\frac{\partial u_n}{\partial x} + B(u^{n-1}, u^{n-1})\right) = 0, \tag{3.2}$$

so that

$$\lambda_1 (u_n)_1 + P_1 B(u^{n-1}, u^{n-1}) = 0. \tag{3.3}$$

The convergence properties of this iteration are discussed in [37]. In practice a single iteration is enough to make a useful improvement to the early stages of a forecast and further iterations have little effect.

The procedure of Browning *et al.* [39] is similar but more general and can also be used as a forecast method. The form of Eqs. (3.1) implicitly assumes that normal modes can be readily calculated. In other situations we can write the basic equations as follows:

$$\frac{\partial \mathbf{u}}{\partial t} + C(\mathbf{u}, \mathbf{u}) + \varepsilon^{-1} D \frac{\partial \mathbf{u}}{\partial x} = 0, \tag{3.4}$$

where ε is a small parameter. The term C governs the "slow" motions and D the "fast" motions. It has been assumed that all the nonlinear terms are contained in C. Then the "bounded derivative" approach, given an arbitrary initial state $\mathbf{u}$, is to require

$$\partial^n \mathbf{u}/\partial t^n = O(1) \qquad \text{at} \quad t = 0$$

up to any desired n. The condition for $n = 1$ gives

$$D(\partial \mathbf{u}/\partial x) = O(\varepsilon) \qquad \text{at} \quad t = 0. \tag{3.5}$$

Replace $\varepsilon^{-1} D(\partial \mathbf{u}/\partial x)$ by E, where E is a smooth $O(1)$ function at $t = 0$. In order to determine it, and hence the desired initial values of u, we set

$$\partial^2 \mathbf{u}/\partial t^2 = O(1) \qquad \text{at} \quad t = 0. \tag{3.6}$$

Since it is assumed that $\mathbf{u}$ is smooth at $t = 0$ and thus $\partial \mathbf{u}/\partial t$ and $\partial \mathbf{u}/\partial x$ are $O(1)$, (3.6) means that $\partial E/\partial t$ must be $O(1)$ at $t = 0$. However, we have

$$\frac{\partial E}{\partial t} = \varepsilon^{-1} D \frac{\partial}{\partial x}\left(\frac{\partial u}{\partial t}\right) = -\varepsilon^{-1} D \frac{\partial}{\partial x}(C(\mathbf{u}, \mathbf{u}) + \varepsilon^{-1} E)$$

so

$$\varepsilon \frac{\partial E}{\partial t} = -D \frac{\partial}{\partial x} C(\mathbf{u}, \mathbf{u}) - -\varepsilon^{-1} D \frac{\partial E}{\partial x}. \tag{3.7}$$

Since, by assumption, $\partial E/\partial x$ is $O(1)$ at $t = 0$, (3.7) can only be satisfied for a special choice of E which involves $C(\mathbf{u}, \mathbf{u})$ at $t = 0$. It is not possible to illustrate the procedure any further with the general equation (3.4). It is worked out fully for the shallow water equations in [39]. The mathematical theory continues by proving that if $\partial^n u/\partial t^n$ is $O(1)$ for all n up to a value N, then bounds can be established on the solution for a certain time T. The method depends on the assumption of initial smoothness of the data which may not be strictly correct but does nevertheless lead to a useful method of suppressing excessive "noise" in forecast integrations.

3.2 *Boundary Conditions*

Ideal lateral boundaries may be present in forecast models in order to limit the computational effort. However, as discussed earlier, primitive equations (2.1)–(2.6) are ill posed with any local choice of lateral boundary conditions. Boundary conditions for the nonhydrostatic equations can be chosen by the methods of [18]. Recent work by Gottlieb *et al.* [41] shows the importance of using the correct characteristic boundary conditions. In practice, however, most limited area models use enhanced artificial damping at the boundaries. Similar problems apply in nested grid models where a series of successively smaller grids of finer resolution are used nearer the area of interest. This is particularly important in tropical cyclone models where the smallest grid must be moved with the cyclone.

It is also important where the finest grid is not to be moved during the forecast but may be in a different place on different days according to the current weather situation. Many operational forecasting centres now have such flexible limited area models. Effective techniques have been developed for allowing information to pass through the boundaries in both directions without generating excessive small scale noise. Some examples are the mass, momentum, and internal energy conserving procedure of Kurihara [42], a boundary relaxation scheme to allow a grid-point model to be nested in a spectral model [6], and a system using three levels of grids which are integrated simultaneously with the necessary extra values created by interpolation [68]. These procedures, and the great variety of others that are used, are all adhoc devices, because of the impossibility of imposing correct lateral boundary conditions.

The problem with the upper boundary condition is that while the mass of a column of air is finite, its top is not well defined. If the equations are written in pressure or sigma (normalised pressure) coordinates, then the usual upper boundary condition is

$$Dp/Dt = 0 \quad \text{at} \quad p = 0 \qquad \text{or} \qquad D\sigma/Dt = 0 \quad \text{at} \quad \sigma = 0. \tag{3.8}$$

Though these conditions may appear superficially to be applied at infinity, they have the effect in practice of a rigid wall at a value of pressure halfway between zero and the top model level (Kirkwood and Derome [43]). It is therefore desirable to replace (3.8) with an absorbing upper boundary condition, perhaps in the manner of Enquist and Majda [44]. This possibility has been studied by Davies [45]. Assume that the governing equations can be separated into horizontal and vertical structure equations. The latter equation takes the form

$$\chi_{zz} + \lambda^2 \chi = 0, \tag{3.9}$$

where

$$\lambda^2 = (KH/h) - \tfrac{1}{4}, \qquad \chi = (p'/\rho_0)\exp(\tfrac{1}{2}z^*), \qquad z^* = z/H,$$

H is the atmospheric scale height, h is the separation constant, K the horizontal wave

number, p a perturbation pressure, and ρ_0 a basic state density. We have to determine λ as an eigenvalue of (3.9) subject to the boundary conditions

$$\begin{aligned} w' = 0; \quad & \text{i.e.,} \quad \left(\frac{\partial}{\partial z} + \left(\frac{1}{2} - K\right)\right)\chi = 0 \qquad \text{at} \quad z = 0, \\ w' = f(p), \quad & \text{i.e.,} \quad \frac{\partial}{\partial t}\left(\frac{\partial}{\partial z} + \left(\frac{1}{2} - K\right)\right)\chi = -\Gamma^*\chi \quad \text{at} \quad z = z_{\text{top}}, \end{aligned} \tag{3.10}$$

where Γ^* is an operator to be defined. This problem is studied in [45]. Consider waves with χ proportional to $\exp i(mx + ny - \tilde{\sigma}t - \lambda z)$. If an absorbing upper boundary condition is used for (3.9), its effectiveness is shown to depend on designing the operator Γ^* appearing in (3.10) such that $\Gamma^* \approx \tilde{\sigma}\lambda$ for the range of atmospheric wave motions that propagate energy vertically through the level $z = z_{\text{top}}$. The procedure of [44] would start from the nonlocal condition

$$\partial\chi/\partial z = -i\lambda\chi \qquad \text{at} \quad z = z_{\text{top}}$$

and derive a local approximation to it from an equation for λ. In [44] an expansion about normal incidence was used. This limit is inappropriate here, since it corresponds to nonpropagation in the vertical. It is therefore suggested that the appropriate expansion is in powers of $\partial/\partial x$, using (3.10) as the first local approximation.

Despite these theoretical developments, it is still common to use (3.8) as the upper boundary condition in large forecasting models. One reason for this is the limited impact on forecasts that results from inserting extra layers at the top (e.g., [46]) which suggests that spurious reflection of energy is not a primary source of error.

4. Numerical Solution Techniques

4.1 *Introduction*

Numerical techniques used in atmospheric models have traditionally been designed assuming that the equations have smooth solutions, and deal with any discontinuities that occur by means of artificial viscosity. The emphasis has been on developing efficient schemes, bearing in mind the wide variation of signal speeds in the equations, with little implicit damping and accurate phase speeds. These have been chosen so that perturbations cannot grow on a faster time scale than in the continuous equations. These considerations have led to the use of neutrally stable centred finite difference schemes by most workers. Spurious error growth is removed by using averaging techniques in calculating nonlinear terms, and artificial viscosity is added to deal with regions where the solution of the inviscid equations breaks down. Schemes satisfying these requirements are extensively reviewed in [1, 2]. A wide variety of techniques are used by different groups. The use of schemes with very efficient time integration procedures such as semi-implicit or split explicit methods is

gradually becoming more popular. Spectral models have become more widely used in recent years for global forecasting. The results are broadly similar to those of finite difference models. A detailed comparison of the ECMWF grid point and spectral models by Jarraud *et al.* [55] shows that the spectral model gave a small but distinct improvement in forecast skill. Recent advances in programming spectral models have made them as, or more, efficient than finite difference models. At ECMWF a spectral model with a maximum total wave number of 63 takes a similar amount of computing to a 192×96 latitude longitude grid-point model. The use of these models can thus be expected to spread. The techniques are described in detail in [2, 16], and it seems unnecessary to repeat them here.

Full descriptions of a semi-implicit and a split explicit difference model are given, as certain of the detailed ideas can be applied more widely. The vertical differencing schemes described can also be applied in models with a spectral representation for the horizontal fields. The scheme described in [50] for solving the implicit parts of the semi-implicit schemes in three-dimensional models by decoupling the equations into a set of two-dimensional problems can be applied to three-dimensional anelastic models used for smaller scale atmospheric modelling (Clark [66]). The same technique can be used for three-dimensional incompressible flows. The split explicit scheme is applied using the Lax–Wendroff method in [13], but can also be used in a leapfrog model, as in the fully compressible model of Klemp and Wilhelmson [67]. Thus it has been possible to adapt many of the great variety of atmospheric models to allow the use of efficient integration schemes. Even with faster computers, it is necessary to use such schemes if large models are to be implemented in real time. For instance, difficulties in converting the U.S. National Meteorological Center grid point model to use such a scheme led to the choice of a spectral model, and have delayed the operational implementation of the very flexible nested grid model of Phillips [68].

4.2 *The ECMWF Grid-Point Model*

This model is described in detail in [12], which includes a full bibliography. References to detail of the formulation are therefore not given in this section. The model uses the standard governing equations (2.1)–(2.6) in sigma coordinates. The u momentum equation (2.1) is written in the form

$$\frac{\partial u}{\partial t} - \frac{1}{\cos\theta} Z p_* v \cos\theta + \frac{1}{a\cos\theta}\frac{\partial}{\partial\lambda}(\phi + E) + \frac{RT}{a\cos\theta}\frac{\partial}{\partial\lambda}(\ln p_*) + \dot{\sigma}\frac{\partial u}{\partial\sigma}$$
$$= \frac{1}{p_*}\left(K\nabla\cdot p_*\nabla u + g\frac{\partial\tau_\lambda}{\partial\sigma}\right) \equiv F_u, \tag{4.1}$$

with a similar equation for v. Here Z is the potential absolute vorticity

$$\frac{1}{p_*}\left(f + \frac{1}{a\cos\theta}\left(\frac{\partial v}{\partial\lambda} - \frac{\partial}{\partial\theta}(u\cos\theta)\right)\right) \tag{4.2}$$

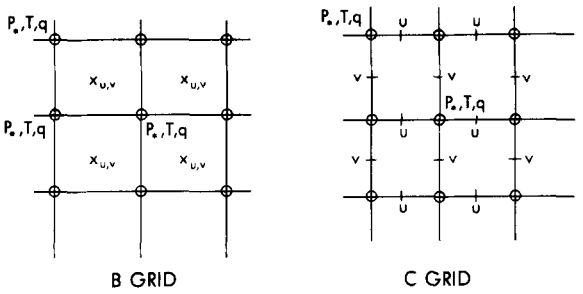

FIG. 2. Nomenclature for different staggered grids.

and E is the kinetic energy per unit mass

$$\frac{1}{2}\left(u^2 + \frac{1}{\cos\theta} v^2 \cos\theta\right). \tag{4.3}$$

Here F_u denotes all the friction terms. The equations are solved on a latitude–longitude grid using 192 points round each line of latitude and 96 points between the north and south poles. There are 15 levels in the vertical at values of σ of (0.996, 0.967, 0.914, 0.845, 0.765, 0.678, 0.589, 0.500, 0.415, 0.334, 0.260, 0.193, 0.132, 0.077, 0.025). The unequal spacing is designed to give high resolution in the boundary layer and near the level of maximum wind ($\sigma = 0.1$ to 0.3). The variables on the grid are staggered. The horizontal wind components, temperature, and humidity are kept on the main levels, and the vertical velocity and geopotential at intermediate levels. The horizontal grid arrangement is the C grid (Fig. 2) as described in the article by Arakawa and Winninghof in [1]. The finite difference scheme is summarised below using the standard notation defined in Section 2. Special care is required in choosing forms for the vertical differencing because the grid is nonuniform in σ. Equation (2.3) is approximed by

$$\phi_{k+1/2} = \phi_* + \sum_{l=k+1}^{K} RT_l(\Delta_\sigma \ln \sigma)_l, \tag{4.4}$$

where ϕ_* is the value of geopotential at the Earth's surface. Equation (2.4) is approximated by

$$\frac{\partial p_*}{\partial t} + \frac{1}{a\cos\theta}(\delta_\lambda U + \delta_\theta(V\cos\theta)) + (\Delta_\sigma\sigma)^{-1}(\Delta_\sigma p_* \dot{\sigma}) = 0. \tag{4.5}$$

Vertically integrated forms of (4.5) are used to derive $\dot{\sigma}$ at each level and a vertical integral from $\sigma = 0$ to 1 is used to give $\partial p_*/\partial t$ independently of $\dot{\sigma}$. This form ensures

mass conservation. The u momentum equation (4.1) is approximated, except at the poles, by

$$\frac{\partial u}{\partial t} - \frac{1}{\cos\theta}(\overline{V\cos\theta}^{\lambda\theta}\,\overline{Z}^{\theta}) + \frac{1}{a\cos\theta}\delta_\lambda(\overline{\phi}^{\sigma} + E) + \frac{R\overline{T}^{\lambda}}{a\cos\theta}\delta_\lambda(\ln p_*) + (\overline{p_*}^{\lambda}\Delta_\sigma\sigma)^{-1}\,\overline{\overline{p_*\dot{\sigma}}^{\lambda}\,\delta_\sigma u}^{\sigma} = F_u, \tag{4.6}$$

where

$$Z = (\overline{p_* a\cos\theta}^{\lambda\theta})^{-1}(af\overline{\cos\theta}^{\theta} + \delta_\lambda v - \delta_\theta(u\cos\theta)), \tag{4.7}$$

$$E = \tfrac{1}{2}(\overline{u^2}^{\lambda} + (1/\cos\theta)\,\overline{v^2\cos\theta}^{\theta}). \tag{4.8}$$

An equivalent form is used for the v momentum equation. This scheme conserves potential absolute enstrophy ($p_* Z^2$) and potential vorticity under the effect of the horizontal advection terms. It does not conserve energy. An alternative version which also conserves energy was found to have excessive truncation error and gave unrealistic results. This problem was first analysed by Hollingsworth and Kållberg [47] and later by Mesinger [48], who shows how it can be avoided. The temperature equation is approximated in advective form rather than flux form (2.5).

$$\frac{\partial T}{\partial t} + \frac{1}{p_*}\left(\frac{1}{a\cos\theta}(\overline{U\delta_\lambda T}^{\lambda} + \overline{V\cos\theta\,\delta_\theta T}^{\theta})\right) + (\Delta_\sigma\sigma)^{-1}\,p_*\overline{\dot{\sigma}\Delta_\sigma T}^{\sigma} - \left(\frac{\kappa T\omega}{\sigma}\right) = F_T, \tag{4.9}$$

where

$$\frac{\kappa T\omega}{\sigma} = \frac{\kappa T}{\sigma}\left(\overline{\sigma\frac{\partial p_*}{\partial t} + p_*\dot{\sigma}}^{\sigma}\right) + \frac{\kappa}{a\cos\theta}(\overline{U\overline{T}^{\lambda}\delta_\lambda(\ln p_*)}^{\lambda}) + \overline{V\cos\theta\,\overline{T}^{\theta}\,\delta_\theta(\ln p_*)}^{\theta} \tag{4.10}$$

and F_T includes all the frictional and diabatic terms appearing on the right of Eq. (2.5). The form of (4.10) ensures that transformations between kinetic and total potential energy exactly cancel if we define

$$1/\sigma = (\Delta_\sigma\sigma)^{-1}\Delta_\sigma\ln\sigma. \tag{4.11}$$

Moisture equation (2.6) is approximated in a similar way to (4.9). These schemes ensure that $p_* T^2$, $p_* q^2$, q, and total potential energy are individually conserved under advective processes.

At the poles the schemes have to be modified. The temperatures are kept at the poles. The v component of velocity is kept a half grid length away. Equation (4.5) is modified so that p_* at the pole changes as a result of the total mass flux through the innermost latitude circle. Thus

$$\frac{\partial p_*}{\partial t} + (\text{SIGN})\frac{a\,\Delta\lambda}{\varepsilon}\sum_{i=1}^{N}(V\cos\theta)_i + (\delta_\sigma\sigma)^{-1}\,\delta_\sigma(p_*\dot{\sigma}) = 0, \tag{4.12}$$

where N is the number of points where V is defined and

$$\varepsilon = N\left(\frac{1}{2}a\,\Delta\lambda\right)\cos\left(\frac{\pi}{2}-\frac{\Delta\theta}{2}\right)\frac{a\,\Delta\theta}{2}.$$

Here SIGN $= +1$ for the North pole and -1 for the South pole. Vertical integrals of this equation are used to find $\dot{\sigma}$ and $\partial p_*/\partial t$ individually as before. The u component of velocity at the poles is meaningless and must be replaced by a polar zonal mass flux U_i so that the standard finite difference equation for V at the rows adjacent to the pole can be used. Define the kinetic energy per unit mass E at the poles by

$$E = \left(\frac{1}{2}N\cos\left(\frac{\pi}{2}-\frac{\Delta\theta}{2}\right)\right)^{-1}\sum_{i=1}^{N}\frac{1}{2}v_i^2\cos\left(\frac{\pi}{2}-\frac{\Delta\theta}{2}\right). \tag{4.13}$$

The zonal mass flux U is not a true prognostic variable and is determined from a representation of the polar boundary conditions;

$$\begin{aligned}\tfrac{1}{2}a\,\Delta\theta(U_{i+1/2}-U_{i-1/2}) + (\text{SIGN})\,V_i\cos\theta a\,\Delta\lambda \\ = \frac{1}{N}(\text{SIGN})\sum_{i=1}^{N}V_i\cos\theta a\,\Delta\lambda,\end{aligned} \tag{4.14}$$

$$\sum_{i=1}^{N}U_{i+1/2} = 0.$$

The advection term in (4.9) is modified in the same way as in (4.12).

The time integration scheme is carried out by the semi-implicit scheme developed by Robert *et al.* [49]. The three-dimensional version is more fully described in Hoskins and Simmons [50]. Only linearised terms representing gravity waves are treated implicitly. The scheme can be written as follows:

$$\begin{aligned}&\delta_t\overline{(\ln p_*)}^{t} + (p_* a\cos\theta)^{-1}\left(u\,\overline{\delta_\lambda p_*}^{\lambda} + v\cos\theta\,\overline{\delta_\theta p_*}^{\theta}\right)\\ &\quad + (\delta_\sigma\sigma)^{-1}\left(\delta_\sigma\bar{\dot{\sigma}}^{2t}\right) + (a\cos\theta)^{-1}\left(\delta_\lambda\bar{u}^{2t} + \delta_\theta\bar{v}^{2t}\cos\theta\right) = 0,\end{aligned} \tag{4.15}$$

$$\begin{aligned}&\delta_t\bar{u}^{t} + (a\cos\theta)^{-1}\left(\delta_\lambda\overline{\overline{\phi}^{\sigma}}^{2t} + RT_0\,\delta_\lambda\overline{(\ln p_*)}^{2t}\right)\\ &\quad + (a\cos\theta)^{-1}R(\overline{T}_\lambda - T_0)\,\delta_\lambda(\ln p_*) = a_u,\end{aligned} \tag{4.16}$$

$$\delta_t \overline{T}^t + (\Delta_\sigma \sigma)^{-1} (\overline{\dot{\sigma}^{2t} \Delta_\sigma T_0}^{\sigma}) - \frac{\kappa T_0}{\sigma} (\overline{\dot{\sigma}^{2t} + \sigma \delta_t (\overline{\ln p_*}^t)}^{\sigma})$$

$$+ (\Delta_\sigma \sigma)^{-1} (\overline{\dot{\sigma} \Delta_\sigma (T - T_0)}^{\sigma}) - \sigma^{-1} \kappa (T - T_0)(\dot{\sigma} + \sigma (\overline{\delta_t \ln p_*}^t)_{\neq}^{\sigma}) = a_T, \tag{4.17}$$

where

$$(\overline{\delta_t \ln p_*}^t)_{\neq} + (p_* a \cos\theta)^{-1} (\delta_\lambda U + \delta_\theta V \cos\theta) + (\Delta_\sigma \sigma)^{-1} \Delta \dot{\sigma} = 0. \tag{4.18}$$

There is also an equation for v. The terms a_u, a_T represent all the other terms in the equations. The temperature T_0 about which the system is linearised is a constant 300° K. It has been shown (Simmons *et al.* [51]) that if (4.15)–(4.18) are linearised about too cold a basic state, the system becomes unconditionally unstable.

The linearised system is solved by eliminating all variables at time $t + \Delta t$ except the discrete divergence defined as

$$d = (a \cos\theta)^{-1} \delta_\lambda u + (a \cos\theta)^{-1} \delta_\theta v \cos\theta.$$

The equation for this takes the form

$$\overline{\delta_t d}^t - \Delta t^2 \mathbf{G} (a \cos\theta)^{-1} (\delta_\lambda ((a \cos\theta)^{-1} \delta_\lambda \delta_{tt}(d) + \delta_\theta (a^{-1} \cos\theta \delta_\theta \delta_{tt}(d))) = 0, \tag{4.19}$$

where $\mathbf{G}$ is an $n \times n$ matrix, n is the number of vertical layers in the model. It expresses the vertical coupling of the equations due to the vertical integrations of hydrostatic equation (4.4) and continuity equation (4.5).

Equation (4.19) is solved by diagonalising the matrix $\mathbf{G}$ by multiplying by the matrix $\mathbf{E}$ of eigenfunctions of $\mathbf{G}$. The resulting discrete approximation to a two-dimensional Helmholz equation is solved by Fourier transform in the longitudinal direction and Gaussian elimination in the latitudinal direction.

The remaining problems in making the scheme efficient are to overcome the time-step restriction caused by the convergence of meridians at the poles and to add an artificial viscosity term to stabilise the solution. These problems are both solved by using an implicit fourth-order diffusion term. Given a provisional value $u^*_{t+\Delta t}$ from (4.16), set

$$u_{t+\Delta t} = u^*_{t+\Delta t} - \left(\frac{\kappa_\lambda}{a^4 \cos^4\theta} \delta_\lambda^4 u^{t+\Delta t} + \frac{\kappa_\theta}{a^4} \delta_\theta^4 u^{t-\Delta t} \right). \tag{4.20}$$

A similar procedure is carried out in the equation for V, T, and q. Operationally,

$$\kappa_\lambda = (a \Delta\lambda)^3.$$

Note that the longitudinal term δ_λ^4 is treated implicitly. The implicit equation is solved by a Fourier transform. The latitudinal term is added on using a forward time

step. This scheme avoids the need for any further smoothing in high latitudes. Equation (4.20) is evaluated near the poles using a reflective boundary condition.

This completes the description of the numerical method used to solve (2.1)–(2.6) in the ECMWF model. The techniques used to represent the other processes, surface exchanges, convection, and radiation are very elaborate but are outside the scope of this review. These are subject to frequent change, details can be obtained from the current internal reports. The model uses a 15 minute time step and takes 20 CPU minutes to integrate one simulated day on the CRAY-1.

4.3 *The Meteorological Office Model*

The other numerical scheme that we describe is that used in the current Meteorological Office model. This model is required for a different purpose than the ECMWF model, in that its forecasts have to be available within 4 hours of data time. The ECMWF model is not run until 10 hours after data time so that observations are available from the whole of the world and a high-quality global data set can be constructed.

The model uses a splitting technique developed and analysed by Gadd [13–15] and discussed further by Carpenter [52]. The equations are solved on a latitude–longitude grid with the same longitudinal resolution as the ECMWF model (192 points) but with 60 points between equator and pole. The model is implemented on a CYBER 205, and a major design feature of this vector computer is that the significant relatively modest increases in I/O speed. Thus the model is chosen to be entirely contained in memory in forecast mode. This gives the split explicit scheme an advantage over that used in the ECMWF model because it requires only one time level of data storage. For some applications the model is not run for the whole globe but only for the region north of 30° S. In this case, a high diffusion zone is used near the southern boundary to prevent reflection of energy.

The equations are solved in sigma coordinates as in the ECMWF model. Fifteen levels are used, but are slightly differently spaced because precise information is required at certain standard levels for aviation. Those used are (0.997, 0.975, 0.935, 0.870, 0.790, 0.690, 0.590, 0.490, 0.390, 0.310, 0.250, 0.190, 0.125, 0.065, 0.025). The variables are staggered on the grid differently from the ECMWF model. The geopotential is calculated at a full level, though the vertical velocity is calculated at an intermediate level. In the horizontal, the *B* grid (Fig. 2) is used. This is because it has been found that the time truncation error due to the splitting is less damaging in practice with this arrangement [13].

The equations are split into two parts, referred to below as adjustment and advective parts. Each major time step of 15 minutes consists of three 5 minute adjustment steps, carried out by a forward–backward scheme, and one advection step, carried out by a two-step Lax–Wendroff method. The advective rather than the flux form of the Eqs. (2.1)–(2.6) is used. Potential temperature Θ is used as a variable instead of T. The splitting is as follows:

Adjustment:

$$\frac{\partial u}{\partial t}+\dot{\sigma}\frac{\partial u}{\partial \sigma}-fv+\frac{1}{a\cos\theta}\left(\frac{\partial \phi}{\partial \lambda}+C_p\sigma^{\kappa}\Theta\frac{\partial}{\partial \lambda}\ln p_*\right)=0, \tag{4.21}$$

$$\frac{\partial \Theta}{\partial t}+\dot{\sigma}\frac{\partial \Theta}{\partial \sigma}=0, \tag{4.22}$$

$$\frac{\partial p_*}{\partial t}+\frac{1}{a\cos\theta}\left(\frac{\partial U}{\partial \lambda}+\frac{\partial}{\partial \theta}(V\cos\theta)\right)+\frac{\partial}{\partial \sigma}(p_*\dot{\sigma})=0, \tag{4.23}$$

$$\frac{\partial \phi}{\partial \sigma}+\frac{RT}{\sigma}=0, \tag{4.24}$$

together with equivalent parts of the equations for v and q.

Advection and other terms:

$$\frac{\partial u}{\partial t}+\frac{u}{a\cos\theta}\frac{\partial u}{\partial \lambda}+\frac{v}{a}\frac{\partial u}{\partial \theta}-\frac{uv\tan\theta}{a}=F_u, \tag{4.25}$$

$$\frac{\partial \Theta}{\partial t}+\frac{u}{a\cos\theta}\frac{\partial \Theta}{\partial \lambda}+\frac{v}{a}\frac{\partial \Theta}{\partial \theta}=F_{\Theta}, \tag{4.26}$$

together with equivalent parts of the equations for v and q.

The adjustment terms are solved by the forward–backward scheme as follows:

$$p_*^{t+\Delta t}-p_*^{t}=-\frac{\Delta t}{a\cos\theta}\left[\delta_\lambda\left(\overline{\overline{p_*}^{\lambda\theta}\sum_{n=1}^{N}u_n\Delta\sigma_n}^{\theta}\right)+\delta_\theta\left(\overline{\cos\theta\,\overline{p_*}^{\lambda\theta}\sum_{n=1}^{N}v_n\Delta\sigma_n}^{\lambda}\right)\right]^{t} \tag{4.27}$$

with a similar forward step for (4.22) and the q equation.

$$u^{t+\Delta t}-u^{t}=\Delta t\left[\frac{1}{2}f(v^{t}+v^{t+\Delta t})-\overline{\dot{\sigma}^{t}}^{\lambda\theta}\delta_\sigma u^{t}-\frac{1}{a\cos\theta}\left(\overline{\delta_\lambda\phi^{t+\Delta t}}^{\theta}+C_p\sigma^{\kappa}\,\overline{\Theta^{t+\Delta t}}^{\lambda}\,\delta_\lambda(\ln p_*^{t+\Delta t})^{\theta}\right)\right] \tag{4.28}$$

with a similar equation for v. Note that Eq. (4.28) is not fully backward, and the pair (4.27), (4.28), if applied indefinitely, would be unconditionally unstable because of the treatment of the vertical advection term. However, the stability of the entire scheme depends on the eigenvalues of the operator defined by three applications of the adjustment equations followed by one application of the advection equations. In practice, it is found that the small potential computational instability from this term is compensated for elsewhere, perhaps by the damping implicit in the advection

scheme. The centred implicit treatment of the Coriolis term in (4.28) is required for successful integrations of more than a few days, because errors from the slow computational instability of an explicit representation tend to accumulate in time.

The vertical differences that appear in (4.22) and (4.24) require care because of the unequal spacing of levels, as in the ECMWF model. They are calculated as follows to ensure balancing of the energy conversions:

$$\dot{\sigma}(\partial\Theta/\partial\sigma) = (\alpha_k \varepsilon_{k+1/2} + \beta_k \varepsilon_{k-1/2})(\alpha_k + \beta_k)^{-1}, \tag{4.29}$$

where

$$\begin{aligned} \alpha_k &= \sigma_k - \sigma_{k-1/2}, \\ \beta_k &= \sigma_{k+1/2} - \sigma_k, \\ \varepsilon_{k+1/2} &= \dot{\sigma}_{k+1/2}(\Theta_{k+1} - \Theta_k)(\sigma_{k+1} - \sigma_k)^{-1}, \\ \phi_k &= \phi_* + \sum_{s=1}^{k-1} RT_s \log\left(\frac{\sigma_{s-1/2}}{\sigma_{s+1/2}}\right) + RT_k \log\left(\frac{\sigma_k}{\sigma_{k+1/2}}\right) \end{aligned} \tag{4.30}$$

The horizontal differencing scheme is prone to a weak grid separation effect since the B grid is well known to support two independent solutions. This is overcome by a method devised by Mesinger [53] and extended by Janjic [54]. A term

$$w(\Delta t^2/4)(\nabla_+^2 - \nabla_\times^2)\,\Pi \tag{4.31}$$

is added to (4.27), where

$$\begin{aligned} \nabla_+^2 \Pi = {} & \frac{1}{a^2 \cos^2\theta}(\delta_\lambda(\overline{p_*}^{\lambda}\delta_\lambda\hat{\phi}) + \delta_\lambda(\overline{\hat{\Theta}}^{\lambda}\overline{p_*}^{\lambda}\delta_\lambda \ln p_*)) \\ & + \frac{1}{a^2\cos^2\theta}(\delta_\theta(\cos\theta\overline{p_*}^{\theta}\delta_\theta\hat{\phi}) + \delta_\theta(\cos\theta\overline{\hat{\Theta}}^{\theta}\overline{p_*}^{\theta}\delta_\theta \ln p_*)), \\ \nabla_\times^2 \Pi = {} & \frac{1}{a^2\cos^2\theta}(\delta_\lambda(\overline{\overline{p_*}\,\overline{\delta_\lambda\hat{\phi}}^{\theta}}^{\theta}) + \delta_\lambda(\overline{\overline{p_*}\,\overline{\hat{\theta}}^{\lambda}\delta_\lambda \ln p_*}^{\theta}\,)) \\ & + \frac{1}{a^2\cos\theta}(\delta_\theta(\overline{\cos\theta p_*\,\overline{\delta_\theta\hat{\phi}}^{\lambda}}^{\lambda} + \delta_\theta(\overline{\cos\theta\overline{p_*}\,\overline{\hat{\theta}}^{\theta}\delta_\theta \ln p_*}^{\lambda}\,) \end{aligned}$$

with

$$\hat{\phi} = \int_0^1 \phi\, d\sigma, \qquad \hat{\Theta} = C_p \int_0^1 \sigma^\kappa \Theta\, d\sigma.$$

The maximum value of w that can be used is 0.25, and this value is used at present.

The advection terms (4.25), (4.26) are approximated by the two step Lax–Wendroff scheme as follows:

$$u^{t+(1/2)\Delta t} = \overline{u^t}^{\lambda\theta} - \frac{1}{2}\Delta t \left(\frac{\overline{u}^{\lambda\theta}}{a\cos\theta}\delta_\lambda \overline{u}^{\theta} + \frac{\overline{v}^{\lambda\theta}}{a}\delta_\theta \overline{u}^{\lambda} \right)^n, \tag{4.32}$$

$$u^{t+\Delta t} = u^t - \Delta t \left(\frac{\overline{u}^{\lambda\theta}}{a\cos\theta}((1+c)\,\overline{\delta_\lambda u}^{\theta} - c[(1-b)\,\delta_{3\lambda}\overline{u}^{\theta} + b\delta_{3\lambda}\overline{u}^{3\theta}]) \right.$$
$$\left. + \frac{\overline{v}^{\lambda\theta}}{a}((1+c)\,\overline{\delta_\theta u}^{\lambda} - c[(1-b)\,\delta_{3\theta}\overline{u}^{\lambda} + b\delta_{3\theta}\overline{u}^{3\lambda}]) \right). \tag{4.33}$$

The constants b and c are set to $\frac{1}{3}$ and $\frac{3}{4}(1-\mu^2)$, where the Courant number μ is

$$\left(\frac{u^2\Delta t^2}{a^2\cos^2\theta\,\Delta\lambda^2} + \frac{v^2\Delta t^2}{a^2\Delta\theta^2} \right)^{1/2} \tag{4.34}$$

As discussed in [52], a value of $c = \frac{1}{2}(1-\mu^2)$ would give fourth-order phase speeds. Due to error cancellation and the effect of errors in scales closer to the grid length, the larger value of c was preferred. However, the choice is found to make almost no visible difference with the horizontal resolution used in this model. At the poles the temperature and pressure are stored, but not predicted independently. At each time step the value is set to the mean of the surrounding points. Equation (4.33) is simplified near the pole with c set equal to zero. Otherwise the difference schemes are unchanged.

This difference scheme does not exactly conserve any quadratic quantities, unlike the ECMWF scheme. The form used for the vertical differencing ensures that the conversions between kinetic and potential energy balance. Practical tests of the scheme suggest that, for forecasts for a few days ahead, no serious deficiencies in conservative properties arise. The smoothing terms $K\nabla \cdot p_* \nabla u$ that appear in (2.1)–(2.6) are approximated by a special scheme. The model is written in sigma coordinates, and sigma surfaces slope steeply near regions of steep orography, so that a diffusion operator on a σ surface does not carry out horizontal smoothing. Because the atmosphere is strongly stratified in the vertical this can lead to errors, particularly in the temperature. Therefore the formula

$$\nabla_p^2 u \simeq \nabla_\sigma^2 u - 2\nabla_\sigma \left(\sigma \frac{\partial u}{\partial \sigma} \right) \cdot \nabla \ln p_* \tag{4.35}$$

is used, where the suffixes p and σ refer to differences taken on p and σ surfaces. Equation (4.35) only contains the largest necessary correction term, namely, that due to the local slope of the σ surface, and the neglect of the more complicated curvature terms may give rise to errors where there are rapid changes in the slopes of the mountains.

In order to overcome the problem of the convergence of the meridians near the pole, Fourier filtering is used. A discrete Fourier transform is carried out on the time increments at each line of latitude and the amplitudes of all waves whose wavelengths are shorter than a preset limit are set to zero. Separate filtering is carried out for increments from the adjustment and advection terms. If a given wave is filtered from the adjustment increments, it must also be filtered from the advective increments. The normal area filtered extends to 45° N for advection and 60° N for adjustment, but can easily be varied.

This completes the description of the numerical methods used in the model. The schemes used to represent the additional physical effects are very elaborate, but outside the scope of this article. Full details of these procedures will be published at a later date. The version extending to 30° S at present takes $2\frac{1}{2}$ minutes of CPU time per simulated day on a CYBER 205.

4.4 *Further Developments in Numerical Techniques*

In this section we briefly review some other current developments in numerical techniques.

The use of semi-implicit or split explicit time integration schemes as described above has led to considerable increases in efficiency and allows time steps of up to 15 minutes to be taken with a horizontal grid length of 150 km. Robert [56] has devised a method for the shallow water equations which allows even longer time steps by using a Lagrangian–rezoning technique for the vorticity advection and the existing implicit treatment of the divergence terms. It is possible that such methods could be extended to treat the advective terms of the full equations with a splitting technique. Many advection schemes developed for other physical applications, however, have been unsuccessful when used for the meteorological problem because they damp or distort sinusoidal waves with a wavelength of less than about ten grid lengths. While some damping is required in practice, the amount that has to be added to a neutrally stable centred scheme like that in the ECMWF model is much less than that present in schemes designed for step function data. The amount present in the staggered two-step Lax–Wendroff scheme appears to be the most that can be tolerated, the unstaggered version tested by Grammeltvedt [57] is not accurate enough.

There are rather more current developments in the vertical approximation scheme. The sigma coordinate system provides a very convenient way of handing complex terrain, but may also lead to truncation errors. A recent discussion of these is given by Mesinger [74]. This may be particularly serious at high levels in the atmosphere where the true flow is horizontal. There may also be errors caused by the averaging of topographic heights over grid squares. This results in lowering the maximum heights and the blocking effect of ridges may be underestimated. An alternative method, using a rigid wall to represent the mountain, has been studied by Egger [58]. The effect of any given mountain range may well depend on whether the wind blows across or parallel to the ridges, and a more complete representation than just an average height may be required [59].

The problem of the coordinate surfaces not being horizontal at the higher levels is

easier to deal with. Simmons and Burridge [17] have developed a hybrid sigma–pressure coordinate system for the ECMWF model. This has been tested by Simmons and Strufing [60]. At the same time they have implemented changes to the vertical finite difference scheme in order to achieve energy and angular momentum conservation with the new coordinates. These changes can be summarized as follows: Consider a general terrain-following vertical coordinate η which is a monotonic function of pressure p and dependent on the surface pressure p_*. Equations (2.3) and (2.4) become

$$\frac{\partial}{\partial \eta}\left(\frac{\partial p}{\partial t}\right)+\frac{1}{a\cos\theta}\left[\frac{\partial}{\partial \lambda}\left(u\frac{\partial p}{\partial \eta}\right)+\frac{\partial}{\partial \theta}\left(v\cos\theta\frac{\partial p}{\partial \eta}\right)\right]+\frac{\partial}{\partial \eta}\left(\dot{\eta}\frac{\partial p}{\partial \eta}\right)=0, \quad (4.36)$$

$$\frac{\partial \phi}{\partial \eta}=-\frac{RT}{p}\frac{\partial p}{\partial \eta}. \quad (4.37)$$

The model variables u, v, and T are carried at unequally spaced levels denoted by an integral suffix, and values of η and thus the pressure are calculated at intermediate half levels. As with sigma coordinates, care is required because of the unequal spacing, and it is necessary to make a special choice of the values of η at the half levels. In order to conserve energy, the vertical advection terms in (4.36) and the other equations of motion are approximated by

$$\left(\dot{\eta}\frac{\partial T}{\partial \eta}\right)_k=\frac{1}{2\Delta p_k}\left(\left(\dot{\eta}\frac{\partial p}{\partial \eta}\right)_{k+1/2}(T_{k+1}-T_k)+\left(\dot{\eta}\frac{\partial p}{\partial \eta}\right)_{k-1/2}(T_k-T_{k-1})\right), \quad (4.38)$$

where Δp_k is $(p_{k+1/2}-p_{k-1/2})$. Equation (4.37) is approximated by

$$\phi_{k+1/2}=\phi_*+\sum_{l=k+1}^{K} RT_l \ln\left(\frac{p_{l+1/2}}{p_{l-1/2}}\right) \quad (4.39)$$

which is in a similar form to (4.4) in σ coordinates. Angular momentum conservation is achieved by setting

$$\left(\frac{RT}{p}\nabla p\right)_k=\frac{RT_k}{\Delta p_k}\left(\left(\ln\frac{p_{k+1/2}}{p_{k-1/2}}\right)\nabla p_{k-1/2}+\alpha_k\nabla(\Delta p_k)\right) \quad (4.40)$$

in the pressure gradient term in the momentum equation. The parameter α_k is chosen to be

$$1-(p_{k-1/2}/\Delta p_k)\ln(p_{k+1/2}/p_{k-1/2}), \quad (4.41)$$

which ensures that (4.40) reduces to (4.6) in sigma coordinates. The rest of the scheme required to achieve exact balance of the energy conversion terms and accurate calculation of pressure gradients over steep topography is set out in [60].

All the finite difference expressions used involve values of pressure only at half levels. The values of p at full levels can be defined, for instance, by

$$p_k = \frac{\Delta p_k}{\ln(p_{k+1/2}/p_{k-1/2})}, \qquad k > 1,$$
$$p_1 = 2p_{1/2}. \tag{4.42}$$

Other options are possible. This finite difference scheme can be implemented in a semi-implicit model by constructing a linearised system similar to (4.19). Some of the hydrostatic equation now has to be evaluated explicitly, since (4.37) is nonlinear.

The vertical coordinate η to be used in practice is selected by experiment, usually to give the upper levels of the model on or near constant pressure surfaces, and smoothly varying layer thicknesses near high topography.

The upper boundary condition used with this scheme is $\dot{\eta} = 0$ at $\eta = 0$. This is similar to the condition normally used in sigma coordinates. It would therefore be of interest to combine this scheme with the absorbing upper boundary conditions discussed in the previous section.

5. Representation of Small Scale Processes

It is common in computational fluid dynamics to be unable to resolve the viscous dissipation scale. In weather prediction, however, this scale is not resolved by many orders of magnitude, and a great many features of the observed flow are not resolved either. It is therefore necessary both to add artificial viscosity to the basic inviscid equations and to insert extra terms to represent the mean effect of unresolved motions. It has to be assumed that the results do not depend on the detail of the unresolved scales, and this assumption is one of the limiting factors on predictive skill.

It is clearly important to determine whether the solutions derived by adding artificial viscosity to the equations and solving on a coarse grid actually converge to the inviscid limit of Eqs. (2.1)–(2.6), and whether such a limit describes the observed behaviour of the atmosphere. These questions have not yet been answered, or even studied very much. In this respect meteorology lags well behind some other branches of computational physics. Some work has been carried out on studying forecasts using very fine grids. These studies inevitably have had to be carried out on limited areas, for instance the U.S.A., and the results can be affected by boundary conditions and by the lack of high resolution data. It is also impossible to assess the forecasts for more than a limited period, since weather systems only stay in the fine mesh for a short time. It has been found in general that 24 hour surface pressure forecasts do not change greatly when the resolution is reduced below 100 km, but the intensity of precipitation continues to increase. More details can be found in papers by Miyakoda and Rosati [61], Phillips [68], Anthes and Keyser [62], and Bosart [63], which include results from grid lengths as short as 34 km. There have been problems,

however, which may indicate difficulties in achieving the desired convergence of the solution. Inevitably, these tend to be unpublished, but it was found in early versions of both the ECMWF and U.K. models discussed in Section 4 that overdevelopment of depressions occured in many cases. This had to be dealt with by making adjustments to the way sub-grid-scale frictional processes were represented, in particular the vertical exchange of momentum between the lower layers of the model and the Earth's surface. In development of the U.K. fine mesh model, which uses essentially the same scheme as set out in Section 4 but on a limited area, it has been found that precipitation amounts increase as the grid length is decreased from 100 km through 75 km to 50 km, in agreement with other authors, but that when checked against observations the extra precipitation in the 50 km model is sometimes incorrect and the 75 km model more accurate. Further development of the methods used to simulate sub-grid-scale vertical convective mixing may be required in order to resolve this problem.

Another recent piece of work which may indicate convergence problems with the numerical method is a study by Williamson *et al.* [64]. They show that the initial data obtained by applying normal mode initialisation, as discussed in Section 3, to real data sets disagrees with the analysed observations by more than the observational error. This implies that the analysis derived from observations does not represent a balanced model state. This can result from deficiencies in the analysis or interpolation method, errors in or absence of observed data, or differences between the forecast model and the atmosphere. Since the errors occur even where there is plenty of data and there are larger than normal observational errors, there is a possibility that they indicate problems with the forecast model.

One of the usual explanations for the problems discussed above are deficiencies in the representation of sub-grid-scale processes. Research on these representations is probably more extensive than that on numerical solution of the governing equations and a review of it is out of the scope of this paper. Here we discuss a few aspects of it relevant to the mathematical problem. It has to be assumed that the effect of the sub-grid-scale processes on the results can be expressed in terms of information which is predicted explicitly, as in the standard problem of turbulent closure. However, the scales involved are probably too large for closures relevant to three-dimensional isotropic turbulence to be relevant, even if they could be found. It has been proposed by Leith [69] and Basdevant *et al.* [73] that two-dimensional turbulence theory be used. Any such theory leads to the formulation of a form for artificial viscous terms to be added to the equations. As discussed in [70], however, the form chosen for these terms has little impact on the results. The suggestion by Chorin [30] and others that three-dimensional turbulence is intermittent rather than homogeneous could lead to radically different closures.

There has been more development work in vertical mixing processes. These are included in models as parts of routines for calculating precipitation and interactions with the surface of the Earth. Reviews of these representations are given in [1, 2]. The vertical mixing between the ground and the lowest layers of the atmosphere is derived from observational studies. The mixing between layers in the model is usually

restricted to temperature and humidity so that in areas where $\partial\Theta_\omega/\partial p > 0$, the model relaxes rapidly towards the neutral state $\partial\Theta_\omega/\partial p = 0$. The amount of mixing is minimised in regions where $\partial\Theta_\omega/\partial p < 0$. (Here Θ_ω is a potential temperature allowing for the effect of latent heat release as well as for adiabatic temperature changes.) It is not usual to include vertical mixing of momentum except close to the Earth's surface. In some models attempts have been made to represent the vertical exchange of momentum due to the presence of sub-grid-scale lee waves set off by the flow of the air over mountains.

These approaches all assume that the best practical approximation to the solution of (2.1)–(2.6) is produced by adding artificial mixing terms to ensure that the numerical solution stays smooth and hoping that the large scale predictions will correspond to reality. Since so much is unresolved, this procedure is clearly an act of faith. An alternative strategy would be to seek a piecewise smooth solution as can be done in certain other branches of physics. It is not clear whether such a description would be appropriate for atmospheric motions on scales larger than 50 km, and any such description would have to be essentially three-dimensional. The work described by Hoskins [26] gives some suggestive starting points. The control exerted on the solution by conservation laws for potential vorticity, mass, and potential temperature is so strong that a great deal can be inferred about the solution by almost qualitative methods. Further progress in this direction can be expected.

6. Results from Numerical Forecasts

In this section we show some of the numerical forecasts produced by the models described in Section 4. The first is a sample three-day forecast from the 26th October 1981. Results are shown for the height of the 500 mb constant pressure surface, for the temperature at 850 mb (about 5000 feet), and for the surface pressure. The weather fronts on the surface pressure charts have been inferred from the pressure and temperature charts. This case was chosen to illustrate how it is sometimes possible to forecast quite large changes from the initial state correctly. An impression of the overall standard over a large number of cases is given later.

The observed 500 mb charts for 26 and 29 October are given in Figs. 3 and 4, with three-day forecasts valid on the 29th October from the ECMWF model and the Meteorological Office model in Figs. 5 and 6. Results are only shown for the European and North Atlantic areas, though one model was integrated globally and the other to 30° S. Consider first the observed changes at 500 mb over the three-day period. There has been considerable progression of the troughs and ridges on the

25° W has moved to 5° W. The centre at 5° E has relaxed north-eastwards and lost its identity, and the centre at 25° E has moved east to 55° E with an intensifying high centre to the north of it. The ECMWF forecast (Fig. 5) predicts all these changes. Detailed comparison shows some faults; for instance, the ridge in mid Atlantic is about 10° too far east and is not intense enough, and the height values over the Baltic

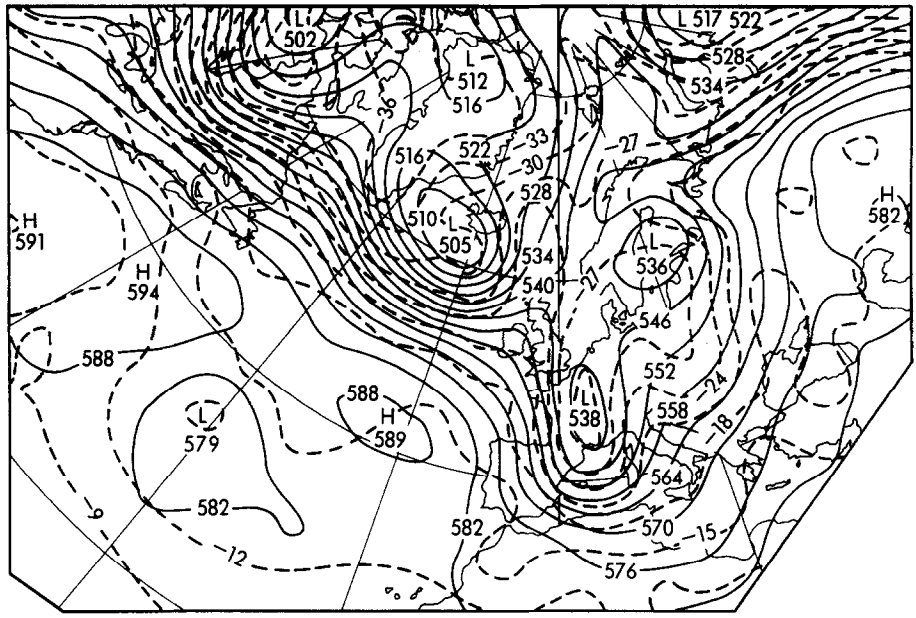

FIG. 3. Height analysis at 500 mb for North Atlantic/Europe valid 12Z on 26 October 1981.

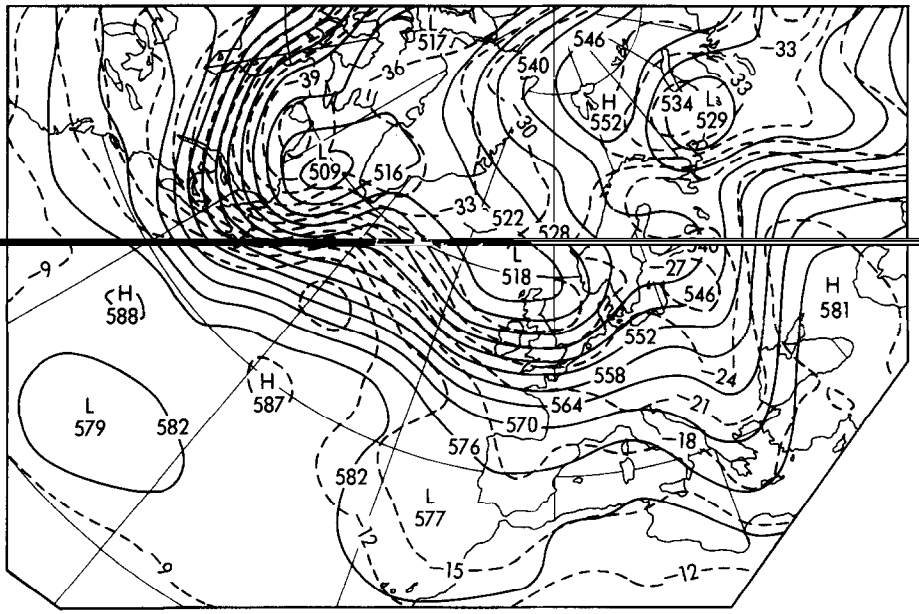

FIG. 4. As Fig. 3, valid 12Z on 29 October 1981.

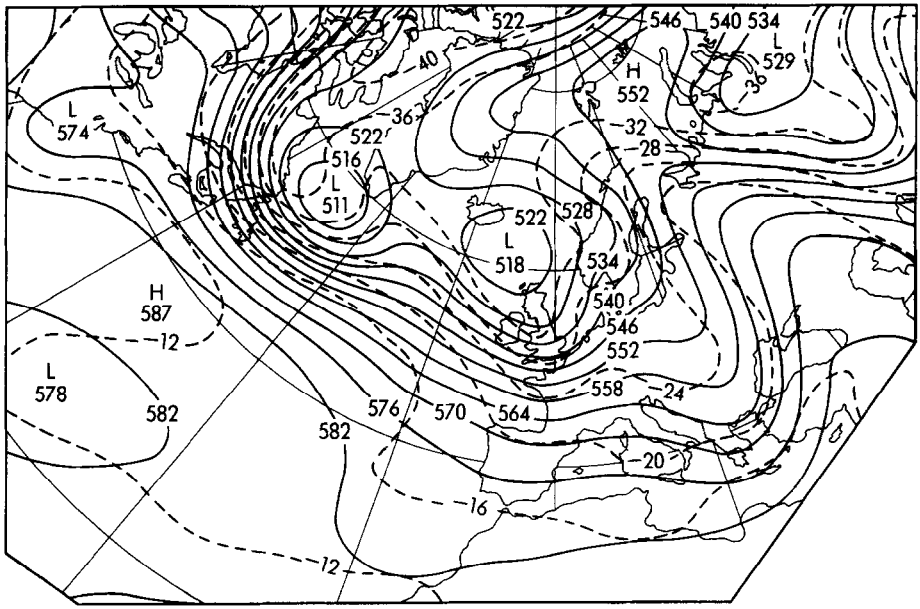

FIG. 5. Three day 500 mb height forecast valid 12Z 29 October 1981, ECMWF model.

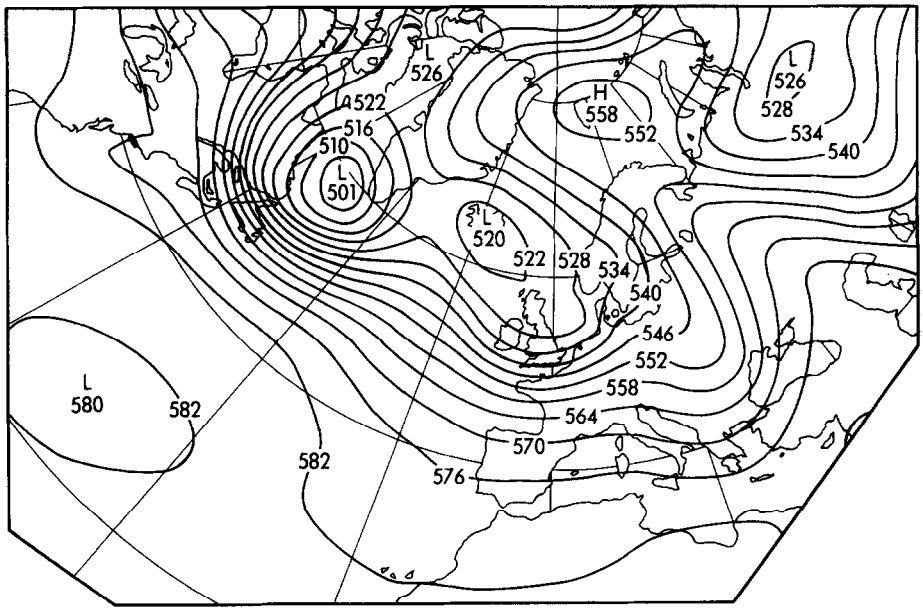

FIG. 6. Three day 500 mb height forecast valid 12Z 29 October 1981, U.K. model.

have risen too much. A spurious trough is also left over the East Coast of the USA at 35° N. The Meteorological Office model (Fig. 6) also predicts all the major changes. The predicted central values of the high and low centres are less accurate than in the ECMWF forecast. Most features are too intense. The amplitudes of the waves in the height pattern is also generally greater, regardless of the truth. In this particular case, the amplitudes of the ridge at 30° W and the trough at 25° E are nearer the truth. The positional errors are present in both forecasts.

The surface pressure and 850 mb temperature charts for 26 and 29 October are shown in Figs. 7 and 8. During the period the very deep depression near Iceland has moved east to Northern Scotland and filled. The systems originally at 90° W at 35° N and 65° N have almost merged into a complex depression at 50° W with a southern centre at 35° W associated with further fronts. An intense region of high pressure has formed behind them at 70° W. The fronts originally over the U.K. have moved southeast and lost their identity. The depression at 10° E 40° N on the 26th has moved north-east to 60° N 30° E together with its fronts. The wind over the U.K. has backed from north-west to west with fronts moving east into Germany, rather than south-east into Spain.

The ECMWF model (Fig. 9) has forecast the main pressure centres quite well. The depression near Scotland is too shallow and too far west. The developments in the west Atlantic are correctly forecast except for the development of a spurious centre at 35° N 70° W. Over Europe the forecast is less accurate. The ridge of warm air at 40° E is over-intense and the depression over the Baltic region is not accurately forecast; the main trough is forecast to lie north-south rather than east-west.

The Meteorological Office model (Fig. 10) produces a similar forecast. The main difference is near Scotland, where three separate centres are shown. The average of the three positions is about where the main centre should be. The forecast over Europe is similar to ECMWF. All the features are somewhat more intense than in the ECMWF forecast, again regardless of the truth.

This comparison shows that both models gave essentially similar forecasts and predicted most of the observed changes over three days in this case. The Meteorological Office model gave more intense and somewhat less accurate high and low centres. At present this is believed to be because it was run from an analysis produced for the ECMWF model, which was consequently inconsistent with some of the model formulation. Some over-development takes place while the data in the bottom boundary layer, which controls the effective friction, adjusts to suit the Met Office model.

These differences are typical of those produced in sensitivity experiments, when one feature of a model formulation is altered. Such changes may affect the overall intensity of the systems produced, but not affect the major differences between the forecasts and the observations.

In order to gain an overall impression of the current standard of numerical forecasts we can use either a statistical measure, as already shown in Fig. 1, or subjective assessments. In the latter case the numerical forecasts are assessed according to whether they provide useful guidance to the forecaster. A summary of

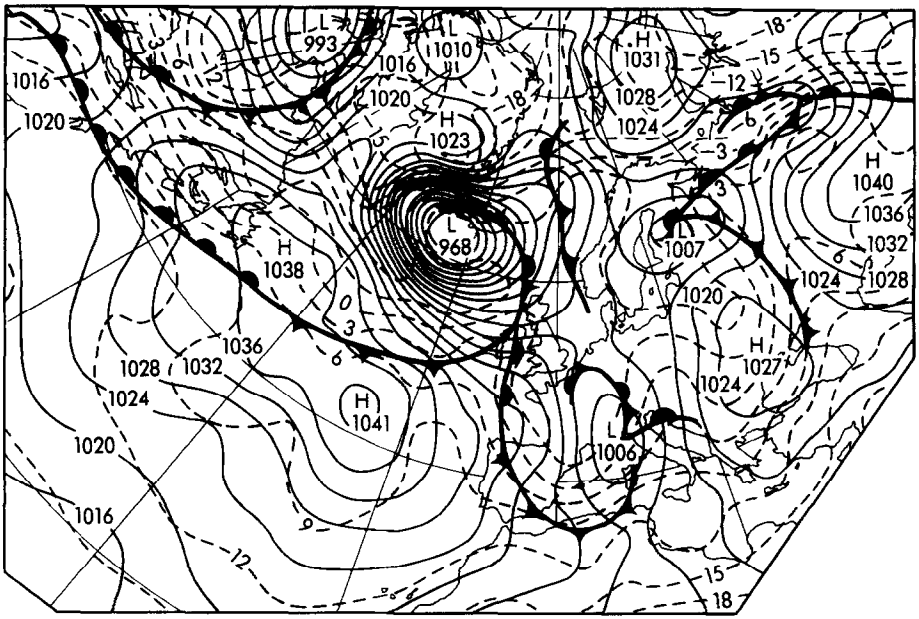

FIG. 7. Surface pressure and 850 mb temperature analysis valid 12Z on 26 October 1981.

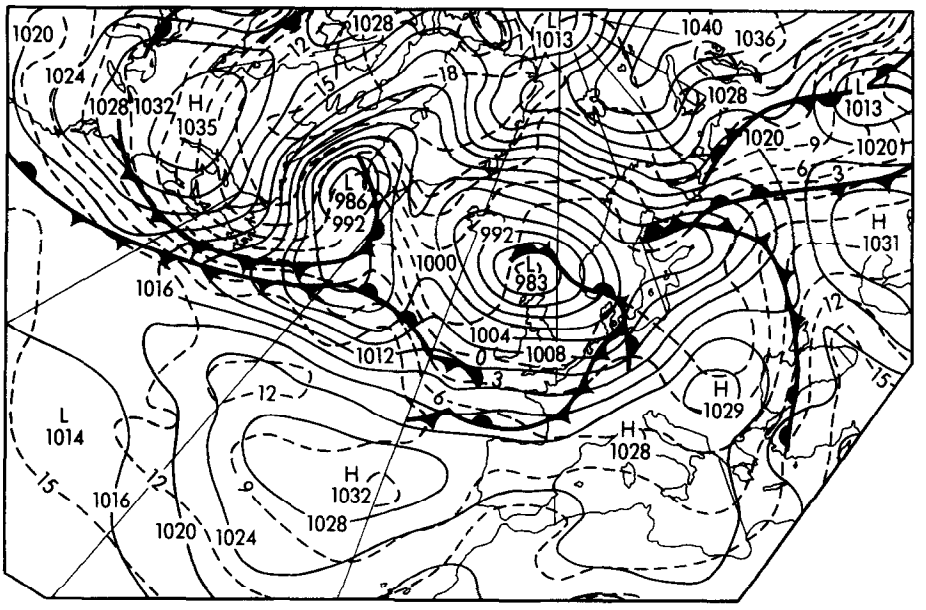

FIG. 8. As Fig. 7, valid 12Z 29 October 1981.

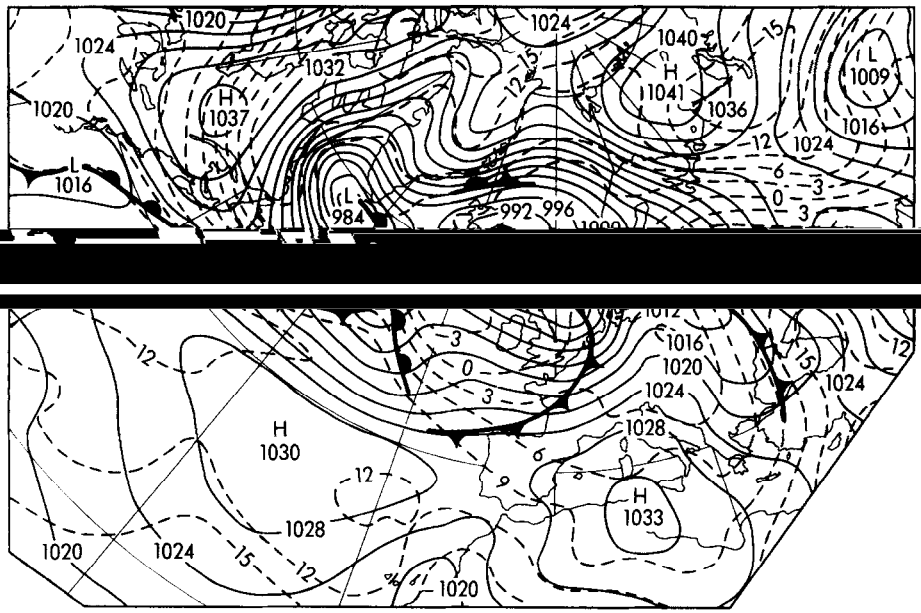

FIG. 9. Three day surface pressure and 850 mb temperature forecast valid 12Z 29 October 1981, ECMWF model.

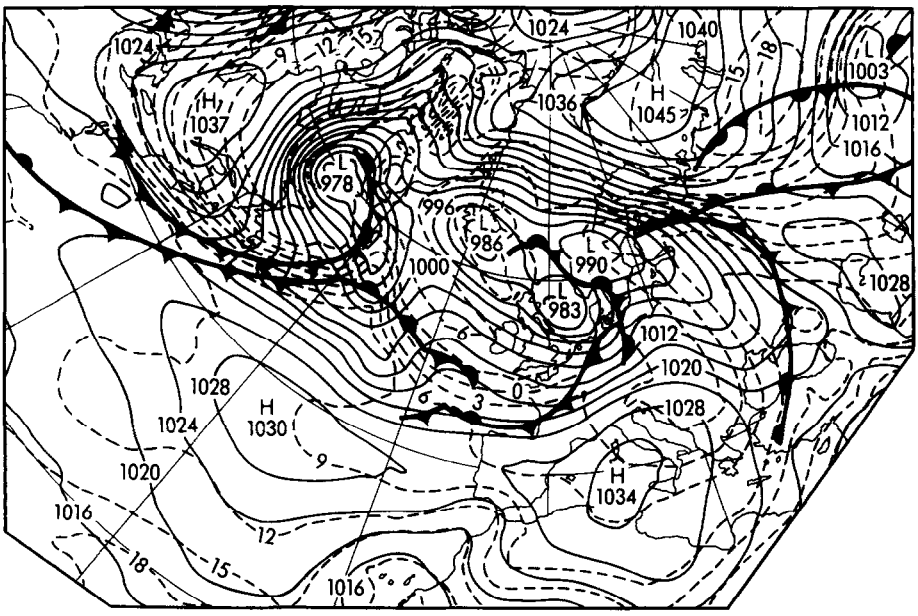

FIG. 10. As Fig. 9, U.K. model.

TABLE I

Assessment of ECMWF Model in Period 3/3/80 to 9/3/81 (52 Cases)

	Day						
	1	2	3	4	5	6	7
N. Atlantic/Europe							
Surface							
A	50	44	21	10	2	2	1
B	2	8	25	21	17	7	7
C	0	0	6	21	33	43	44
500 mb							
A	52	48	30	13	2	2	1
B	0	4	19	25	25	15	8
C	0	0	3	14	25	35	43
U.K. area							
Surface							
A	49	37	30	17	9	5	2
B	3	15	16	19	19	18	14
C	0	0	6	16	24	29	36
500 mb							
A	51	47	37	21	16	10	4
B	1	5	12	19	19	17	11
C	0	0	3	12	17	25	37

Note. Assessment scheme: A, good guidance; B, no major error; C, misleading in some important respect.

such assessments for the ECMWF model is given in Table I. For further details of this assessment see [10].

While inevitably subjective, this table indicates that very few major errors of significance to a medium range forecaster occur by day 3, but that half or more forecasts are misleading beyond day 5. This assessment was carried out independently by all 15 member states of ECMWF, and the conclusions of all of them were in rough agreement as to how long the forecasts were useful.

7. Discussion

This paper illustrates that considerable progress has been made in numerical weather prediction in the last ten years and useful results are being obtained up to four days ahead. However, whereas ten years ago it was relatively easy to identify the major sources of error and rectify them, as set out in [9], it is now no longer so clear what further developments are required. It is well known that the atmosphere is inherently unpredictable and that the length of time for which it is predictable varies

markedly from case to case (Lorenz [65]). It is therefore not clear how much more can be done. There is an obvious need to generate better initial data sets, both by using more observations and by improving assimilation techniques.

Other errors can be studied in the context of their effect on the climatology of the model. In [71] the systematic forecast errors in the ECMWF model are considered in this way. The discussion in Section 5 illustrated some possible areas of further work. In particular, more effort is needed in understanding the solutions to the inviscid equations. This is likely to become even more important as emphasis is placed on mesoscale modelling with gridlengths of 10–50 km. It may well be that in the next ten years it will be this aspect of weather prediction which makes the greatest advance.

ACKNOWLEDGMENTS

The author would like to thank Professor H.-O. Kreiss, Dr. A. J. Simmons, and various colleagues at the Meteorological Office for comments on earlier versions of this manuscript.

REFERENCES

1. J. CHANG (Ed.), "Methods in Computational Physics," Vol. 17, Academic Press, New York, 1977.
2. G. J. HALTINER AND R. T. WILLIAMS, "Numerical Prediction and Dynamic Meteorology," 2nd ed., Wiley, New York, 1980.
3. D. J. GAUNTLETT, D. M. BURRIDGE, AND K. ARPE, ECMWF Research Dept. Internal Report No. 6 (1977), Reading, U.K.
4. E. KALNAY-RIVAS AND D. HOITSMA, *in* "Proceedings, AMS Conference on NWP, Silver Spring, Md., Nov. 1979," preprint, pp. 302–312.
5. D. L. WILLIAMSON, *Mon. Weather Rev.* **106** (1978), 69.
6. L M. LESLIE, G. A. MILLS, AND D. J. GAUNTLETT, *Quart. J. Roy. Meteorol. Soc.* **107** (1981), 629.
7. M. J. P. CULLEN, S. J. FOREMAN, J. W. PRINCE, A. M. RADFORD, AND D. R. ROSKILLY, *Mon. Weather Rev.* **109** (1981), 422.
8. D. BAUMHEFNER AND P. DOWNEY, *Mon. Weather Rev.* **106** (1978), 1245.
9. A. J. ROBERT, *Montreal McGill Univ. Dept. Met. Publ. Met.* **114** (1976), 68.
10. B. A. HALL, Met. O. 11 Tech. Note No. 147, U.K. Meteorological Office, Bracknell, 1981.
11. C. E. LEITH, *J. Atmos. Sci.* **37** (1980), 954.
12. D. M. BURRIDGE AND J. HASELER, ECMWF Technical Report No. 4, 1977.
13. A. J. GADD, *Quart. J. Roy. Meteorol. Soc.* **104** (1978), 569.
14. A. J. GADD, *Quart. J. Roy. Meteorol. Soc.* **104** (1978), 583.
15. A. J. GADD, *Quart. J. Roy. Meteorol. Soc.* **106** (1980), 215.
16. W. BOURKE, B. MCAVANEY, K. PURI, AND P. THURLING, *in* "Methods in Computational Physics," Vol. 17, pp. 267–324, Academic Press, New York, 1977.
17. A. J. SIMMONS AND D. M. BURRIDGE, *Mon. Weather Rev.* **109** (1981), 758.
18. J. OLIGER AND A. SUNDSTROM, *SIAM J. Appl. Math.* **35** (1978), 419.
19. H.-O. KREISS, *Comm. Pure Appl. Math.* **23** (1970), 277.
20. H.-O. KREISS AND G. BROWNING, *SIAM J. Appl. Math.* **42** (1982), 704.
21. S. KLAINERMAN AND A. MAJDA, *Comm. Pure Appl. Math.* **34** (1981), 481.
22. R. SADOURNY, *J. Atmos. Sci.* **32** (1975), 680.
23. T. KATO, *Arch. Ration. Mech. Anal.* **25** (1967), 188.

24. A. F. BENNETT AND P. E. KLOEDEN, *Quart. J. Roy. Meteorol. Soc.* **107** (1981), 121.
25. R. T. WILLIAMS, *J. Atmos. Sci.* **29** (1972), 3.
26. B. J. HOSKINS, *Ann. Rev. Fluid Mech.* **14** (1982), 131.
27. S. KLAINERMAN, *Comm. Pure Appl. Math.* **33** (1980), 43.
28. F. JOHN, *Comm. Pure Appl. Math.* **34** (1981), 29.
29. R. H. MORF, S. A. ORSZAG, AND U. FRISCH, *Phys. Rev. Lett.* **44** (1980), 572.
30. A. J. CHORIN, *Comm. Pure Appl. Math.* **34** (1981), 853.
31. A. J. CHORIN, *Comm. Pure Appl. Math.* (1982), to appear.
32. A. J. CHORIN AND J. E. MARSDEN, "A Mathematical Introduction to Fluid Mechanics," Springer–Verlag, Berlin/New York, 1979.
33. A. MAJDA AND S. OSHER, *Comm. Pure Appl. Math.* **32** (1979), 797.
34. J. GARY, *in* "Numerical Methods Used in Atmospheric Models," Vol. II, pp. 474–499, WMO GARP Publ. Series No. 17, Geneva, 1979.
35. K. W. MORTON, *in* "The State of the Art in Numerical Analysis, Proc. IMA Conference" (D. Jacobs, Ed.), pp. 699–756, Academic Press, New York, 1977.
36. A. ARAKAWA AND V. R. LAMB, *Mon. Weather Rev.* **109** (1981), 18.
37. F. BAER AND J. J. TRIBBIA, *Mon. Weather Rev.* **105** (1977), 1536.
38. B. MACHENHAUER, *Beitr. Phys. Atmos.* **50** (1977), 253.
39. G. BROWNING, A. KASAHARA, AND H.-O. KREISS, *J. Atmos. Sci.* **37** (1980), 1424.
40. H.-O. KREISS, *Comm. Pure Appl. Math.* **33** (1980), 399.
41. D. GOTTLIEB, M. GUNZBURGER, AND E. TURKEL, *SIAM J. Numer. Anal.* **19** (1982), 671.
42. Y. KURIHARA, G. J. TRIPOLI, AND M. A. BENDER, *Mon. Weather Rev.* **107** (1979), 239.
43. E. KIRKWOOD AND J. DEROME, *Mon. Weather Rev.* **105** (1977), 1239.
44. B. ENQUIST AND A. MAJDA, *Comm. Pure Appl. Math.* **32** (1979), 313.
45. H. C. DAVIES, Met. O. 11 Tech. Note No. 142, U.K. Meteorological Office, Bracknell, 1980.
46. A. J. SIMMONS, *in* "WMO/GARP WGNE Report No. 21" (I. D. Rutherford, Ed.), pp. 57–61, Geneva 1980.
47. A. HOLLINGSWORTH AND P. KÅLLBERG, ECMWF Research Dept. Internal Report No. 22, Reading, U.K., 1979.
48. F. MESINGER, *Mon. Weather Rev.* **109** (1981), 467.
49. A. J. ROBERT, J. HENDERSON, AND C. TURNBULL, *Mon. Weather Rev.* **100** (1972), 329.
50. B. J. HOSKINS AND A. J. SIMMONS, *Quart. J. Roy. Meteorol. Soc.* **101** (1975), 637.
51. A. J. SIMMONS, B. J. HOSKINS, AND D. M. BURRIDGE, *Mon. Weather Rev.* **106** (1978), 405.
52. K. M. CARPENTER, *Quart. J. Roy. Meteorol. Soc.* **107** (1981), 467.
53. F. MESINGER, *Tellus* **25** (1973), 444.
54. Z. I. JANJIC, *Beitr. Phys. Atmos.* **52** (1979), 69.
55. M. JARRAUD, C. GIRARD, AND U. CUBASCH, ECMWF Tech. Report No. 23, 1981.
56. A. J. ROBERT, *Atmos. Ocean* **19** (1981), 35.
57. A. GRAMMELTVEDT, *Mon. Weather Rev.* **97** (1969), 384.
58. J. EGGER, *Tellus* **24** (1972), 324.
59. W. N. C. KEEPING, *in* "WMO/GARP WGNE Report No. 2" (I. D. Rutherford, Ed.), 5.1–5.2, Geneva, 1981.
60. A. J. SIMMONS AND R. STRÜFING, ECMWF Tech. Report No. 28, 1982.
61. K. MIYAKODA AND ROSATI, *Mon. Weather Rev.* **105** (1977), 1092.
62. R. ANTHES AND D. KEYSER, *Mon. Weather Rev.* **107** (1979), 963.
63. L. F. BOSART, *Mon. Weather Rev.* **108** (1980), 1087.
64. D. L. WILLIAMSON, R. W. DALEY, AND T. W. SCHLATTER, *Mon. Weather Rev.* **109** (1981), 2357.
65. E. N. LORENZ, *Tellus* **21** (1969), 289.
66. T. L. CLARK, *J. Comput. Phys.* **24** (1977), 186.
67. J. B. KLEMP AND R. B. WILHELMSON, *J. Atmos. Sci.* **35** (1978), 1070.
68. N. A. PHILLIPS, N.O.A.A. Tech. Report NWS 22, U.S. Dept. Commerce, Rockville, Md., 1979.
69. C. E. LEITH, *J. Atmos. Sci.* **28** (1971), 145.

70. M. J. P. CULLEN, *Quart. J. Roy. Meteorol. Soc.* **102** (1976), 77.
71. L. BENGTSSON AND A. J. SIMMONS, *in* "Large Scale Dynamical Processes in the Atmosphere" (B. J. Hoskins and R. P. Pearce, Eds.), Academic Press, New York, to appear.
72. C. TEMPERTON AND D. L. WILLIAMSON, *Mon. Weather Rev.* **109** (1981), 729.
73. C. BASDEVANT, B. LEGRAS, R. SADOURNY, AND M. BELAND, *J. Atmos. Sci.* **38** (1981), 2305.
74. F. MESINGER, *Geoph. Astr. Fluid Dyn.* **19** (1982), 105.