Optimizing Computations in Weather and Climate Prediction Models^{*}

F. BAER, BANGLIN ZHANG, AND BING ZHANG

Department of Meteorology, University of Maryland, College Park, MD USA

ABSTRACT

To meet the challenge of developing a comprehensive weather and climate prediction model which can give realistic scenarios for many time scales, more computer power than is currently available will be needed. One possibility for alleviating this shortcoming is to increase the integration timestep. We propose and test several methods which may prove useful. One procedure is an expansion of the model dependent variables in a Taylor series. Application of this method to simple models indicates acceptable increases in timestep by a factor of five. A multilevel approach which is less complex to apply gives comparable results and is more successful when high accuracy is desired. To bypass the limiting constraint of the CFL condition on gravity waves, an approach is suggested in which the prediction model is represented in its normal modes and the high frequency modes are balanced while the low frequency modes are predicted. Experiments with this procedure are described and in combination with the multi-level integration technique show substantial increases in integration timestep for acceptable integration results, both on the forecast and climate scale. Experiments are now underway applying this process to the NCAR/CCM3, a state-ofthe-art model.

1. Introduction

Global climate predictions currently made with coupled atmosphere and ocean components and sometimes with a biosphere included, are very complex and require so much computing power on available computing hardware using traditional numerical algorithms that only very few integrations can be accomplished with given resources. Yet despite their complexity, these models have difficulty producing agreement on more than a few general parameters of the system. Unfortunately, current models do not include a sufficient breadth of spatial scales to realistically predict the long time scales which are essential for climate assessment, nor are the details of some significant physical forcing yet sufficiently well understood. Since numerous scenarios must be computationally played out with the most comprehensive models in order to identify meaningful bounds on the range of influence of the many incorporated parametric variables, computing resources some orders of magnitude larger than presently available seem needed to solve the climate problem.

^{*} Contribution to "Riehl Memorial Volume of Meteorology and Atmospheric Physics".

Current climate prediction models are solved as a marching problem in time, and their computing requirements are determined by the number of steps needed to march into the future. Climate change requirements suggest that prediction times may reach beyond decades, including increased resolution to determine regional events. Given the many experiments needed to develop a successful model (GCM) on these time scales, some procedure to increase the speed of model output without loss of accuracy appears not only desirable but essential.

To understand how one might decrease the machine time needed to compute climate model integration results, define a "computing cycle" to be that time required to do once all the calculations which must systematically be repeated to complete the entire calculation. On a true serial machine which can handle just one computation at a time (non-vector), the computing cycle would be the time for that operation. On a massively parallel processor (MPP) with unlimited processors, the computing cycle would include all the calculations which would not need repetition by their dependence on previous calculations. For conventional marching problems, the minimum computing cycle could be one complete time step. Thus it is this computing cycle that one would attempt to approach as a limit for an integration. It is evident from this viewpoint that exploitation of MPPs is appropriate and desirable, and we shall presume that approach in the ensuing discussion.

Since there are numerous ways to construct the climate prediction systems in preparation for numerical computations, the choice should be geared to the computational device available. Consider a SIMD parallel processor. These machines were among the first of the MPPs designed and were comparatively simple as all the processors performed the same function, and thus could be economically manufactured even with many processors. If the climate prediction system could be constructed to take advantage of this computer design, integrations using such a machine with a very large number of processors could approach the ideal computing cycle discussed above.

We have explored this approach by applying the spectral representation first to the barotropic vorticity equation (BVE) and then to a two level baroclinic model, as simple test representations for the more complicated GCMs. If the procedure does not work for a simple model, it cannot be expected to work for more complex systems. Expanding the BVE model dependent variables in horizontal global functions (solid harmonics) and integrating over the global surface yields a set of coupled nonlinear ordinary differential equations which are first order in time and can be extrapolated. More detail on the development of these equations can be found in Baer (1964), and Baer and Platzman (1961) among others. The vector of the time dependent variable expansion coefficients (the weights of the global structure functions used in the expansion) appears in the prediction system as differentiated in time and must be evaluated as the sum of quadratic products of the time dependent coefficients themselves. The application to SIMD parallel processors becomes evident here when one notes that each quadratic product is multiplied by a constant coefficient (effectively a weight) denoted as an interaction coefficient (IC). The vector of ICs is very large. If the ICs are distributed each to a processor, each processor can perform the product of the quadratic product times its IC if only the two expansion coefficients have been delivered to it. Thus all processors perform the same task. A subsequent sweep over all the processors yields the sum which represents the time derivative field of the expansion coefficient vector. A suitable time stepping scheme then yields the new set of expansion coefficients. With a computer containing enough processors, this step can be accomplished in little over one machine cycle. Historically, this method for solving the prediction system proved unpopular because a technique denoted as the transform method, developed by Machenhauer and Rasmussen (1972) and Orszag (1970) was computationally more efficient on available serial processors. In the environment of SIMD parallel processors, this judgment might be reconsidered.

The process although simple and straightforward has the drawback that as resolution increases the number of computations needed grows as the cube of the expansion coefficient index and thus becomes increasingly larger than the number needed to solve the problem by the transform method. If the number of processors were unlimited this difficulty would not exist, but clearly that will never be the case. The other difficulty arises from limitations in communication to and from the processors. At the end of each iteration, new values must be communicated to the processors so that they can produce a new product. Although innovative programming is hastening this activity, the data distribution currently takes more computer time than the actual product calculation in each processor, thus limiting the benefits of parallel processing.

We have performed integrations with the BVE both in the IC format and using the transform method for intercomparison purposes on the CM-200 and CM-5 at Los Alamos National Laboratories (LANL). For various truncations the IC format has

3

run as fast or faster than the transform method, despite the limitations of communication. This result has encouraged us to program and run some experiments with the baroclinic model, a model similar to the one used by Baer and Alyea (1971) and Alyea (1972). Preliminary integrations suggest that on the CM-5 the IC method is comparable to the transform method in speed. Unfortunately, MPP developments have moved in the direction of MIMD machines and SIMD machines with very large numbers of processors (the largest had 64K) are no longer available. This approach has consequently been deferred.

In keeping with the concept that many calculations can be achieved in one computing cycle on a MPP (even the MIMD machines have many processors; the latest machines have in the neighborhood of 1000), another approach is to incorporate as many time steps in one cycle as the computer will allow. Because of the initial value nature of the problem, the equations can be configured such that subsequent time steps can be related to the initial conditions if the calculations are spread out on many processors. Developing such formulas is straightforward for a linear system, but highly complex for nonlinear systems. We thus reverted to series expansion techniques in the expectation that including higher order terms in the time expansion would allow for longer time steps with equal accuracy, so long as the selected time step remained within the stability requirements of the model. Finally, one can reconstruct the prediction equations in a format that will allow for a larger time step without loss of accuracy. We have attempted this approach by expanding the dependent variables in normal modes derived from the linearized version of the nonlinear system, and by balancing the higher frequency modes, predicting only those modes which have a stability condition which allows the use of longer time steps.

We shall show in the subsequent discussion how the application of time expansion methods, including Taylor Series and multi-level techniques can allow for significant increases in time step when applied to simple models such as BVE and the shallow water equations (SWE). This success is encouraging insofar as it suggests that computing efficiency can be enhanced, and encourages us to proceed with applications to more state-of-the art GCMs. Additionally, we will describe experiments with the SWE in normal mode form involving the balancing of high frequency modes which indicate that sizable increases in time step can be achieved using this procedure with little loss of accuracy. This latter technique is not bound to an explicit time integration scheme and can be utilized with a Lagrangian approach, although we have not yet tested the latter.

2. Possible techniques

From a conceptual viewpoint, any initial value problem should be determinable from the initial conditions alone. This suggests that a formula which involves only functions of the initial conditions should be able to determine the ultimate solution without the requirement of solving a number of initial value problems in sequence. Details for developing such a formula clearly become exceptionally complex for even the simplest problem, but the concept is enlightening. Given a sufficiently powerful MPP, the computing cycle could be reduced to one time step for the entire prediction calculation. This hypothesis clearly ignores possible instabilities in the solutions and thus is only appropriate for stable methods or limited to the stability conditions of the selected scheme. Let us explore the possible applications.

Consider first the extension of the time step by use of a Taylor series (TS) which will raise the accuracy of each calculation and thus allow for an increase of the time increment when compared to use of a single first order term of the series. If one considers the following first order differential system in time,

$$\frac{d_{i}}{dt} = f_{i}(1, \cdots, n) \qquad 1 \le i \le n \tag{1}$$

where $_i$ represents any model dependent variable, higher order derivatives can be generated by repeated differentiation. The commonly used forward scheme presented as a Taylor series expansion takes the general form,

$$A t_{o} + t = A + \left(\frac{\partial A}{\partial t}\right)_{t_{o}} t + \left(\frac{\partial^{2} A}{\partial t^{2}}\right)_{t_{o}} \frac{t^{2}}{2!} + \cdots + \left(\frac{\partial^{n} A}{\partial t^{n}}\right)_{t_{o}} \frac{t^{n}}{n}$$

and the leapfrog scheme is,

$$A t_{o} + t = A t_{o} - t + 2\left(\frac{\partial A}{\partial t}\right)_{t_{o}} t + 2\left(\frac{\partial^{3} A}{\partial t^{3}}\right)_{t_{o}} \frac{t^{3}}{3!} + \dots + 2\left(\frac{\partial^{n} A}{\partial t^{n}}\right)_{t_{o}} \frac{t^{n}}{n!}$$

where A can be any of the $_i$. All the required derivatives can be calculated exactly from (1) using Mathematica (Wolfram, 1991), and since the calculations refer only to initial conditions, they can be spread over as many processors of a MPP as are available. If convergence of the Taylor series proves too slow, for more efficiency one can use a Pade´ approximation (see Press et al., 1992) by replacing each series coefficient with a rational function. The coefficients of the Pade´ approximant can be calculated from the Taylor series coefficients.

Since there are a limited number of processors available, and the calculation of the higher derivatives becomes somewhat unwieldy, we have also considered a multilevel approach in time to determine the series coefficients with a substantially smaller computing penalty. Based on the system (1), the following polynomial series can be substituted for the Taylor series at reference time t_o ;

$${}_{i}(t) = \sum_{j=0}^{j=m-1} a_{i,j} (t - t_{o})^{j}$$

$$\frac{d}{dt}{}_{i}(t) = \sum_{j=0}^{j=m-1} j a_{i,j} (t - t_{o})^{j-1}$$
(2).

Assuming that $_{i}$ and $\frac{d}{dt}$ are known from m previous time levels (t-m t), we generate 2m equations to solve for the m coefficients $a_{i,j}$. These equations are solved by the singular value decomposition (SVD, see Press et al., loc. cit.) method. The integration is advanced *l* timesteps (*l* t, 1 l L) with these coefficients and then the process is repeated until the final result is achieved. Both these techniques have been tested with several models and their efficacy will be demonstrated in the following discussion.

3. Models tested

We began our tests with the simplest models since calculations with them are of modest cost and lack of success would preclude more expensive calculations with more sophisticated models. These models fall into the category of two dimensional models either without divergence (the BVE is an example) or with divergence (the SWE fits this description).

a. BVE on a plane

This model is represented by the Lorenz (1960) low-order three component system. It is developed by the expansion of the BVE in doubly periodic Fourier series. The BVE has the following form,

$$\frac{\partial}{\partial t}\nabla^2 = -\hat{k} \cdot \nabla \times \nabla \nabla^2$$
(3)

and for doubly periodic flow, the stream function has the property that, $\left(x + \frac{2\pi}{k}, y + \frac{2\pi}{l}, t\right) = x, y, t'$, where *k* represents the planetary wave number, *l* is the latitudinal wave number, and x and y are the Cartesian coordinates in the longitudinal and latitudinal directions respectively. When the series is truncated to just three wave numbers, the minimum for nonlinear flows (see Platzman, 1962), the vorticity takes on the truncated form,

$\nabla^2 = A\cos ly + F\cos kx + 2G\sin ly\sin kx$

and the prediction system is represented by the following set of three ordinary differential equations in time;

$$\dot{A} = a \ k, l \ FG$$

$$\dot{F} = b \ k, l \ AG$$

$$\dot{G} = c \ k, l \ AF$$
(4).

As noted earlier, repeated differentiation of Eqs. 4 will lead to the following general formula (we present only one for the variable A) which can be calculated from the initial conditions by use of Mathematica;

$$\frac{\partial^{j} A}{\partial t^{j}} = \sum_{i}^{J} s_{i} a^{e_{1i}} b^{e_{2i}} c^{e_{3i}} A^{e_{4i}} F^{e_{5i}} G^{e_{6i}}$$

Note that the solutions will depend on the choice of wave numbers (l, k), the total energy of the system and the energy distribution amongst the modes; i.e., A, F, and G.

b. BVE on the sphere

The vorticity equation for this system is given by Eq. 3. However the expansion over the spherical surface is in associated harmonics which are a composite of Fourier series in longitude () and Legendre polynomials in latitude (). As with the Lorenz system, we select a low order system (see Baer, 1970a) which includes an arbitrary zonal flow and two components (n_{α}, n_{β}) of a planetary wave (l). All zonal components are linearly dependent on each other, thus only one need be included in the nonlinear system. We denote this component with the symbol n, its l index is zero, and it is described by the Legendre polynomial P_n (). The other two components are represented by two dimensional wave numbers and are conveniently represented as complex numbers; they are, $\alpha = n_{\alpha} + il$, $\beta = n_{\beta} + il$. The system equations in terms of the complex stream function coefficients () take the form (for details on this system and its solutions in terms of elliptic functions, see Baer, 1970b);

We selected a somewhat realistic but simple zonal structure in latitude based on the following formula,

G() =
$$(\sin^4 3 \cos^2 + \sin^4 2 - 0.2 \cos^2) \cos^2$$

For this system, the solutions depend on the wave numbers (n, α, β) , the total energy in the system, the energy distribution amongst the modes (n, α, β) , and the distribution of the zonal flow, G().

c. Shallow water equations on the sphere

This system has been thoroughly studied by Kasahara (1977) and we repeat only the basic system equations for completeness. For details the reader is referred to Kasahara's paper. The prediction equations (Navier-Stokes) for the wind components (u,v) and the free surface height (h) are,

$$\frac{du}{dt} - \left(f + \frac{u}{a}\tan\phi\right)v + \frac{g}{a\cos\phi}\frac{\partial h}{\partial\lambda} = 0$$

$$\frac{dv}{dt} + \left(f + \frac{u}{a}\tan\phi\right)u + \frac{g}{a}\frac{\partial h}{\partial\phi} = 0$$

$$\frac{dh}{dt} + \frac{h}{a\cos\phi}\left(\frac{\partial u}{\partial\lambda} + \frac{\partial v\cos\phi}{\partial\phi}\right) = 0$$
(6).

These equations may be written more conveniently in vector form if we assume that the vector $\mathbf{W} = (\mathbf{u} \mathbf{v} \mathbf{h})^{T}$; this yields the form,

$$\frac{\partial \mathbf{W}}{\partial t} + \mathbf{L}\mathbf{W} = i\mathbf{F}(\lambda, \phi, t)$$
(7).

The linear matrix operator **L** and the nonlinear vector **F** can clearly be determined from Eqs. 6. As with the low order barotropic systems, it is convenient to expand the dependent variables (**W**) in appropriate characteristic functions which in this case are the Hough modes $\mathbf{H}_{l}^{s}(,)$ (see Longuet-Higgins, 1968) and which depend on longitude and latitude. The nonzonal parts of **W** are expanded in these polynomials with time dependent coefficients $W_{l}^{s}(t)$, and the zonal components are expanded in Legendre polynomials. The result of the expansion of Eqs. 6 in Hough modes leads to the following prediction equation for the expansion coefficients,

$$\frac{dW_l^s(t)}{dt} + i\sigma_l^s W_l^s(t) = iF_l^s(t)$$
(8),

and the σ_I^s are the eigenvalues of the linear operator **L**. The coefficients of the nonlinear vector, $F_I^s(t)$, are determined by Gauss quadrature.

In the spectral domain, one predicts the evolution of the expansion coefficients. If the coefficient represents a high frequency mode such as a gravity mode, the stability condition could be severe and a higher order calculation would not be beneficial. Under such conditions other procedures have been advanced, the most common of which is to integrate these modes using an implicit time scheme. A more attractive alternative is to present the system in both high and low frequencies, predict the low frequency components and balance the high frequency components. If one combines the low frequency members (slow modes) of the set of coefficients $W_l^s(t)$ into a vector **Y** and the high frequency (fast) modes into the vector **X** by assessing the frequencies σ_l^s , one can write Eq. 8 as follows;

$$\dot{\mathbf{Z}} = i\Lambda_{z}\mathbf{Z} + \mathbf{N}_{z}(\mathbf{Y}, \mathbf{Z})$$
 fast modes,
 $\dot{\mathbf{Y}} = i\Lambda_{y}\mathbf{Y} + \mathbf{N}_{y}(\mathbf{Y}, \mathbf{Z})$ slow modes. (9)

The matrices Λ represent the linear coefficients of Eq. 8 and the vectors **N** are composed of the nonlinear terms. Details of this process may be found in Baer and Tribbia (1977) and there is a rich body of literature on this *initialization* methodology. In particular, although most schemes use the slow mode equation of (9) for prediction, a variety of schemes have been proposed to predict or balance the fast modes. A number of these were summarized by Daley (1980) and we list the five which we have tested. The vector $\mathbf{Z}_{\rm B}$ represents the fast modes which are to be balanced. If no prediction is specified for the slow modes (**Y**), the prediction follows Eq. (9).

Method A: (see Machenhauer, 1977)

$$\mathbf{Z}_{B}^{(\prime)} t = i\Lambda_{z}^{(\prime)} \mathbf{N}_{z} \mathbf{Y} t^{(\prime)} \mathbf{Z}_{B} t^{(\prime)}$$

$$\dot{\mathbf{Y}} t = i\Lambda_{y} \mathbf{Y} t + \mathbf{N}_{y} \mathbf{Y} t, \mathbf{Z}_{B} t$$

Method B: Each predicted fast mode is represented by the symbol z.

$$zt + t = zt - te^{-2i\lambda_z t} + \frac{1 - e^{-2i\lambda_z t}}{i\lambda_z} \mathbf{N}_z \mathbf{Y} t, \mathbf{Z} t$$

Method C:

$$\mathbf{Z}_B t = i \Lambda_z^{-1} \mathbf{N}_z \mathbf{Y} t, \mathbf{Z}_B t - t$$

Method D: (see Tribbia, 1979)

$$\mathbf{Z}_{B} t = i \Lambda_{z}^{-1} \mathbf{N}_{z} \mathbf{Y} t , 0$$

$$\dot{\mathbf{Y}} = i \Lambda_{y} \mathbf{Y} t + \mathbf{N}_{y} \mathbf{Y} t , \mathbf{Z}_{B} t$$

Method E:

$$\mathbf{Z}_{B} t = i \Lambda_{z}^{-1} \mathbf{N}_{z} \mathbf{Y} t, \mathbf{Z}_{B} t - t$$

$$\dot{\mathbf{Y}} = i \Lambda_{y} \mathbf{Y} t + \mathbf{N}_{y} \mathbf{Y} t, \mathbf{Z}_{B} t - t$$

The solutions depend on the spectral truncation limits for the Hough modes, the zonal profile chosen, and the initial energy distributions.

d. SWE in normal mode form

For this model representation, we use the Hough equations (8) and integrate them as a control run to generate a model climate. We extract the mean flow from this climate. We then linearize Eqs. 8 but include a linear interaction with the mean flow which we have taken from the climate run. Thus the resulting equations will look identical to Eqs. 8 except that we must replace the nonlinear term on the right hand side; this will look as follows,

$$iF_l^s(t) \to i\sum_{s'} \alpha_l^{s'} W_l^{s'}(t) \tag{10}$$

where the $\alpha_l^{s'}$ are the coefficients which come about because of the linear interaction with the specified climate zonal flow. Note that (8) can now be rewritten as

$$\frac{\partial \mathbf{W}}{\partial t} + \hat{\mathbf{L}} \mathbf{W} = 0$$

$$\hat{\mathbf{L}} = \mathbf{L} - i \, \alpha_I^s$$
(11)

and α_l^s is the matrix of the coefficients α_l^s . If we decompose $\hat{\mathbf{L}}$ into its principal

components (say X (\mathbf{W})) and define them as the normal modes, they satisfy the linear system,

$$\frac{\partial \mathbf{X}}{\partial t} + \mathbf{A}_{\hat{L}} \mathbf{X} = 0$$
(12).

We now expand the SWE (Eqs. 7) in terms of these modes as distinct from the Hough mode expansion (Eqs. 8), generating a new spectral system. Using the eigenvalues of \hat{L} , we separate the modes of X into fast and slow components, developing a system similar to (9).

Test integrations were performed with this system using both the standard leap-frog and the multi-level schemes, wherein the slow modes were predicted and the fast modes were balanced using Method C described above. The solutions will depend on the truncation of the system, the cutoff point between fast and slow modes and on the initial distribution of dependent variables.

4. Results

We have tested the above models with a variety of techniques to establish a speedup in computing without loss of accuracy.

a. Taylor series with Lorenz model

Using the Taylor series expansion for the dependent variables of (4), we have integrated the equations under a variety of conditions. We considered two different maximum zonal winds (15 and 20 meters per second), three wave ratios (k/l = 2, 6, 20) representing shorter to the longest planetary waves respectively, and three different initial energy distributions amongst the waves determining the relative magnitudes of the coefficients A, F, and G. These variations gave a total of 18 experiments from which to draw our conclusions for this model.

For each experiment we established a reference integration which was made with a three minute timestep; we established that this time increment gave results which was exceptionally close to the analytic solution. Subsequent integrations were performed over extended times ranging from 100 to 1200 hours, the limit determined from the known oscillation frequency, since we required the nonlinear oscillation to go through a number of cycles. The experiments were then run using various numbers of terms of the Taylor series. We considered three features of the integrations to measure the accuracy of the predictions, in all cases comparing the results to the reference integration. We chose the time mean of the variables, their nonlinear period, and the amplitude of their oscillation. Finally we defined a maximum allowable timestep (t_{max}) as that integration interval which gave results which did not exceed a 10 percent error norm of the three measures defined above when compared to the exact solution. Because there are three variables, the norm involves nine components.

The specific value of t_{max} is not consequential, but the relative value for each of the various experiments is. We therefore normalize the t_{max} values to the case which uses the regular leapfrog method; i. e., with only the first term of the Taylor series. Thus for example a value of five would indicate that one could use a timestep five times larger than the conventional timestep used when applying the usual leapfrog scheme and achieve results within ten percent of the exact solution. The results vary depending on the specific experimental conditions chosen, so we have summarized them for all 18 experiments, including a standard deviation so that the variability among the experiments may be noted. Figure 1 presents these results, showing the increase in time step achievable for the allowed error discussed above and normalized as stated (ordinate) as a function of the number of terms in the Taylor series used (abscissa). Note that both the leapfrog scheme and the forward scheme have been tested.

We see for this system that savings in computational time can be substantial; up to a factor of 6.5 if one is willing to use 15 terms of the Taylor series. Comparable savings can also be achieved using the forward time scheme. For studies using this model these benefits are dramatic, but one also gains some insight into the potential computational advantages for less simple models. Utilizing MPP machines effectively, the penalty for using many terms of the Taylor series may be negligible.

b. Taylor series with spherical model

In this series of experiments we expand the dependent variables of (5) in Taylor series and integrate for a variety of conditions similar to the process described above for the BVE in a channel. Here we also consider two energy levels determined from the maximum initial zonal wind speed. Additionally we chose three distinct planetary waves (l = 3, 5, 10), and five distinct combinations of (n_{α} , n_{β}). Finally, we selected two zonal/wave initial energy combinations yielding a total of 20 experiments. The reference integration was established in an identical fashion to the channel problem, and the experiments were run in the same way, with varying numbers of terms of the Taylor series. The conditions on t_{max} were the same as in the previous model, as were the error bounds, but we used the amplitudes of the two wave energies and the total energy as the measures for testing the prediction accuracy and errors, again comparing to the reference (exact) solution.

Figure 2 shows the summary results of the 20 experiments in a representation identical to Fig. 1. Note however that we made the calculations for only a maximum of seven terms of the Taylor series and thus were only able to experience a benefit factor of slightly above four. If one were to extrapolate to 15 Taylor series terms, the results would be comparable to Fig. 1. The development of the higher derivatives of the Taylor series for this model using Mathematica was considerably more involved and computer intensive, and for this reason we terminated the calculation. On a MPP with many processors, this increased computing penalty would become irrelevant. Nevertheless, this model also indicates substantial savings in computing requirements using the Taylor series approach.

c. Taylor series with shallow water Hough model

The shallow water system presents a greater challenge to computing efficiency than the BVE insofar as it allows for gravity motions which propagate at relatively high frequencies and violating the CFL condition with explicit time schemes will lead to instability. Moreover this system does not have analytic solutions, and like other non-externally forced nonlinear systems may exhibit bifurcation tendencies. To avoid the latter problem, we have tested the system only on the forecast time scale and limited our integrations to six days. To reduce potential instability from gravity motions, we have balanced a subset of gravity modes.

We have investigated the shallow water model in Hough mode form (Eqs. 8) by separation into fast and slow modes (Eqs. 9). Following Kasahara (loc. cit.), we include 20 planetary waves; Hough modes 1 l 60, including 20 Rossby modes, 20 eastward gravity modes and 20 westward gravity modes for each planetary wave.

The equivalent depth chosen was 10 km and the initial distribution was a Haurwitz wave (see Phillips, 1959) with wave number six (for details see Kasahara). For testing purposes a control run was made using the leapfrog scheme with a time step of three minutes and including the explicit prediction of all modes (Eqs. 9) for six days.

In the experiments to be described, the dependent variables (\mathbf{Y}, \mathbf{Z}) of Eqs. 9 were expanded in Taylor series and integrations were performed using the first order terms only (standard leapfrog scheme) and using both the first and third order terms. Various calculations were also made in which all Rossby modes were predicted explicitly, but different subsets of gravity modes were balanced. It is evident that if more gravity modes are balanced, a longer timestep could be used.

We established a number of experiments to test the effect of balancing on the allowed time increment by assessing the linear frequency of the gravity modes from the first of Eqs. 9. Ten categories were set up, each with a subset of balanced gravity modes which had a stability time scale smaller than the selected time step whereas those gravity modes explicitly predicted had a frequency which was stable for the selected time step. The first category included all gravity waves in the subset of balanced modes and the time step selected was 57 min. The tenth category had a time step of 12 minutes and included only a small fraction of the gravity modes in the balanced subset. Table 1 lists the time steps used for cutoff of the balanced modes in the various categories. Note that one can integrate with a smaller time step than the critical one for any category, since all the explicitly predicted modes for any category are linearly stable for the critical time step listed. Indeed choosing a shorter time step improved the integration results as might be expected and will be demonstrated.

Table 1: Critical time steps selected for each category of balanced gravity wavesubsets.

Category	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10
t _{crit} (min)	57	57	51	36	27	21	18	15	12	12

Let us first consider the effects of the various balancing schemes (Methods A-E of Section 3 above). Our observation from a number of experiments with the various methods suggests that no one method stands out as significantly superior. Figure 3a demonstrates the RMS vector wind difference (when compared to the control run) developed during the six days of integration with Category D1 and using the critical time step of 57 min., clearly the worst case since all the gravity modes are balanced. Note that the errors grow substantially with time for all the methods. We noted that as we changed category and balanced fewer gravity modes using the appropriate critical timestep, the error was lower and grew less rapidly, but there was no particular preference demonstrated for any one method. However, when Category D7 was reached, the error growth was minimal and all methods gave very similar results (see Figure 3c). If the time step for any Category is reduced from the critical value, the error was reduced. This can be seen from Figure 3b which is the same as Fig. 3a but describes results based on a time step of 30 minutes.

The impact of balancing fewer modes (with the penalty of using a shorter timestep) on the forecast quality can be seen on Figure 4. Based on the results shown above with the various methods, we describe only those with the use of method A. The RMS vector wind error is shown for experiments with all ten categories using t_{crit} . As indicated above, we see on Fig. 4a that errors grow substantially until one reaches Category D7 after which the errors remain small.

Using more terms of the Taylor series expansion is demonstrated on Fig. 4b where we use the first and third terms of the Taylor series. We note that the errors are substantially reduced from those seen on Fig 4a (the ordinary leapfrog scheme) but is most apparent and significant for those categories greater than D6. Because of the complexity in calculating the higher derivatives for the Taylor series expansion, we did not include higher order terms, but the tendency for improvement in predictions using them seems evident.

d. Multi-level scheme with Lorenz model

Because the derivatives as coefficients of the Taylor series become more difficult to evaluate as their order increases and/or the model complexity expands, we tested the effects on integration time using the multi-level scheme as a substitute for the Taylor series coefficients (see Eq. 2). For simplicity, we applied the technique to the Lorenz model discussed above in subsection *a*. We performed the same 18 experiments and used the three minute timestep integrations for reference. However to establish better statistics we ran each case for 150-200 periodic cycles, considerably more than for the Taylor series experiments. Moreover, we determined error statistics at various accuracy levels in addition to the ten percent level chosen previously, ranging from 0.1% to 10%. A number of variations of the multi-level

scheme were also considered; for the number of terms in the series, we varied m in the range, 1 m 70, and for the number of time steps with the same coefficients, we varied L in the range 1 L 5. Finally to test for convergence we ran some experiments using the Pade' approximation.

At the ten percent accuracy level, the results using the multi-level scheme are similar to those using the Taylor series method with a number of terms. However, as the accuracy requirements are increased, significant improvement is seen with the multi-level scheme, whereas no such benefit is noted using the Taylor series scheme. Indeed when the accuracy is set at 0.1 percent, the multi-level scheme gives very good results for substantially longer time steps than the Taylor series scheme. A demonstration of this is seen on Figure 5. Here we show on the ordinate the ratio of the allowed time step to that of the simple leapfrog scheme for the stated accuracy, and the number of terms in the series (m) on the abscissa. The number of iterations for a given set of coefficients for these experiments is L=3. The bars on the one percent curve represent the variance in the mean value for the 18 experiments and give a measure of the variability in the experiments.

Changing L does not appear to affect the results systematically. For some experiments a smaller value gave better results whereas for others a larger value improved the results. The figure shows how the results are affected by the number of series terms used. In all cases, using more than 20 terms had a negative effect and in general the optimum benefit was found at less than 20 terms. Finally, the application of the Pade' approximation did not increase the maximum allowed time step significantly for any of the experiments we ran.

e. Normal mode form of the SWE model

The previous experiments were designed to establish the advantages of a series expansion in optimizing the time step. In this study with the SWE expanded in normal modes, the emphasis is on enlarging the time step based on the methodology of splitting the modes and balancing those of high frequency. For these experiments we included both forcing and dissipation, distributed uniformly in scale. Additionally these functions were energetically balanced so that no significant changes to the energetics of the system was notable. The total number of modes selected was 2520; the slowest 500 were predicted as described in subsection 3.*d* above and the remainder were balanced. The critical period at cutoff for this choice (and others)

may be seen from Figure 7. For integration purposes both the simple leapfrog and the multi-level schemes were tested. For the multi-level scheme, the parameters were set to L = 10 and m = 10. To compare the experimental integrations to the expected solutions of the system, control runs were made for ten years with t = 3 minutes, yielding robust solutions. Note that for this nonlinear system exact solutions do not exist. Three separate experiments were made using the initial conditions specified by Kasahara (loc. cit.) but using wave five instead of six. In addition to the case for which the energy level was specified by Kasahara (Case a), we ran two additional cases in which the energy level of the system was halved (Case b) and doubled (Case c). The initial fields of u,v,h are shown on Figure 6.

The time steps we selected to run these experiments were t = 45 min. for case a, t = 48 min. for Case b, and t = 36 min. for Case c. These times were selected based on tests which indicated stable integrations and were somewhat less than the critical values determined from the cutoff frequency of the modes based on linear analysis. Figure 7 shows the CFL limit for a subset of modes used in the integrations. To establish meaningful statistics on integration periods of one month, we selected every 60th day of the control runs as an initial data set and ran the experiments for 30 days. We thus created 50 predictions for each experiment. These predictions were averaged to yield much more stable results.

The error growth of the global RMS vector wind difference over the 30 days of integration may be seen on Figure 8. We show here the 50 realization average difference between the case and the control normalized by the standard deviation of the daily vector wind from the control. The normalization factor determines the random variability of the 50 realizations. Note that the variability of the experiments relative to the control are much less than this variability. Although the error grows during the period it never exceed 15 percent for any of the cases which indicates a good forecast. Although there is some variability from case to case, it is not substantial. We do observe that the multi-level scheme outperforms the simple leapfrog scheme for all cases and at all times, especially as the integration time increases. Similar results were noted for the geopotential height differences and the errors were even smaller, never reaching 12 percent.

F. Normal mode form of the SWE model; Climate scale

Having seen the benefits to using the modal approach on the forecast scale, we applied it on the climate scale. We used the same model and initial conditions as in the previous study for monthly forecasts. In this experiment we integrated for ten years using the same time increments for the cases as in the previous experiment, and extracted monthly averages of relevant variables for the last nine years of data, yielding 108 realizations (note that there is no seasonality in the model). These results were compared to the output of the control runs which used the three minute timestep.

Using zonally averaged data, Figure 9 shows the average difference of the zonal wind of each case from the corresponding control for the entire sample of realizations (108) plotted against northern hemisphere latitude. The differences are quite small indicating a satisfactory prediction, but it is noteworthy that the multi-level scheme outperforms the simple leapfrog scheme significantly. To establish the variability of the experimental results relative to the model variability determined from the 108 control realizations, Figure 10 --using a format similar to Fig. 9-- depicts the ratio (in percent) of the differences of the standard deviation of the experimental zonal wind from the control (based on 108 realizations) to the standard deviation of the zonal wind of the control. This measure shows the errors of the experiments relative to the variability of the model itself, determined from the control run. We see that the errors are small everywhere relative to the variability of the control run, but that the multi-level scheme gives far more accurate results than the simple leapfrog scheme. This is true for all three cases.

A more precise determination on the performance of the experiments can be seen from details of the modes themselves as contrasted to the gross statistics seen on Fig. 10. We compared the 108 realization average amplitudes of the 500 predicted modes of the three cases and noted that they were quite similar to the control values, but more interestingly, that only a handful had sizable amplitudes. Thus these few modes represented most of the profile of the predicted variables and their variation. We selected the eight most dominant modes from each experiment (which appear in pairs as expected) and from the Hough mode analysis represent all three of the model variables, both wind components and the geopotential height. The results of the analysis are represented on Table 2. The selected modes from each experiment are listed as they appear sequentially in the modal set. The average amplitudes of the control run are shown in parenthesis followed by the means of the differences in amplitude between both the simple leapfrog and multi-level integrations from the control respectively, both differences normalized by the control values and presented in percent. The data corroborate the results shown on Fig. 9 and indicate that no significant mode varies far from the control. As to the variability of the differences within the 108 realizations, we calculated the standard deviations of the differences, normalized them by the standard deviation of the model variability as determined from the control run, and presented these values in percent for the eight modes and all cases. These results correspond to those presented on Fig. 10. Again we see that no mode has variability in differences to the control in excess of about 25-30 percent of the model variability and for most modes, considerably less. Finally, and as noted also from Fig. 10, in almost all cases the multi-level scheme out performs the simple leapfrog scheme by a significant factor.

Experiments with other cutoff points were carried out with the smallest point at mode 50. Although this allowed for a considerably longer integration time step, see Fig 7, the errors were unacceptably large.

5. Conclusions

In the research presented herein, we have attempted to identify procedures which could help speed up calculations of atmospheric and/or oceanic prediction models with the intent of achieving more model runs with limited computing resources. The need for such speedup was detailed in the introduction. Our initial effort was to use a series expansion of the model dependent variables in time, calculate the coefficients of these series from initial conditions and test the methodology to determine if comparable results to conventional time integration methods could be gained with a longer time step.

We explored using Taylor series since for first order systems in time, the higher derivatives could be determined exactly using sophisticated computer programs such as Mathematica. We applied the process to the barotropic vorticity equation (BVE) on both the plane and on the sphere and noted benefits up to a factor of five in time step if sufficient terms of the series were used. When a more complex system was tested (the shallow water equations (SWE)) a benefit was also noted, but the computing complexity for generating the coefficients prompted us to explore other avenues. Nevertheless, this procedure could potentially serve the prediction community well, particularly those who use simple models for their studies.

To overcome the limitations noted above, we explored an alternate series form using a multi-level scheme. In this case the coefficients are determined from a set of initial conditions determined by calculations with the model itself, in addition to the observed initial conditions. On application to the BVE, results were comparable to those from the Taylor series experiments; however when very high accuracy requirements were imposed, the method performed substantially better. The technique does not require the intense computations which limited our studies with the Taylor series technique, and it can be applied to any model without limitation. Indeed we tested the method with the SWE model and observed very little computing penalty.

For most models other than the most simple, gravity wave propagation will severely limit the time truncation interval if the integration method is explicit (the CFL criterion). In working with the SWE in spectral form we noted that balancing the high frequency modes clearly allowed for significant increases in the time step, and in some cases gave satisfactory results with much larger time steps than would be allowed if all the modes were predicted. Working with this knowledge, we expanded the SWE system to include a climatological mean state and generated the normal modes to this system. We then integrated this system in its modal form balancing the fastest modes and numerically integrating the slower modes. From the experiments described in the text, we discovered that this method gave satisfactory results with sizable increases in time step when compared to control integrations. The experiments included both 30 day integrations (forecast scale) and monthly averages of climate scale integrations. Moreover, tests with the multi-level scheme together with this method gave results superior to the simple leapfrog scheme.

Indeed we were so pleased with the results of these pioneering experiments that we turned to a state-of-the-art model, the NCAR/CCM3, as an application. Since the method does not require rebuilding of the model, we constructed a framework surrounding the model and have to date performed one integration. Our results when compared to the archives of the CCM3 suggest that we can achieve a threefold advantage to the time increment used by that model with comparable prediction skill. This experiment requires further study before it can be published, but we comment on it because of the potentially encouraging results which this method might yield. Moreover, the method is not bound to an explicit time integration scheme and can therefore be utilized with a Lagrangian prediction procedure.

Acknowledgments. The research reported herein has been supported by the U. S. Department of Energy Office of Energy Research, the CHAMMP Program, through Grant No. DEFG0295ER62022 to the University of Maryland College Park. Many of the computations were performed at and with assistance from NERSC, EOL Berkeley National Laboratory, Berkeley, CA.

REFERENCES

- Alyea, F. M., 1972: Numerical simulation of an ice age paleoclimate. Ph. D. Dissertation to the Dept. of Atmos. Sci., Colorado State University, Ft. Collins, CO, 120pp.
- Baer, F., 1964: Integration with the spectral vorticity equation. *J. Atmos. Sci.*, **21**, 260-276.
- Baer, F., 1970a: Dependence of the highly truncated spectral vorticity equation on initial conditions. *J. Atmos. Sci.*, **27**, 987-999.
- Baer, F., 1970b: Analytical solutions to low-order spectral systems. Archiv. Met. Geoph. Biokl., Ser. A., **19**. No. 3, 255-285.
- Baer, F., Alyea, F. M., 1971: Effects of spectral truncation on general circulation and long-range prediction. *J. Atmos. Sci.*, **28**, 457-480.
- Baer, F., Platzman, G. W., 1961: A procedure for numerical integration of the spectral vorticity equation. *J. Meteor.*, **18**, 393-401.
- Baer, F., J. J. Tribbia, 1977: On complete filtering of gravity of modes through nonlinear initialization. *Mon. Wea. Rev.*, **105**, 1536-1539.
- Daley, R., 1980: The development of efficient time integration schemes using model normal modes. *Mon. Wea. Rev.*, **108**, 100-110.
- Kasahara, A., 1977: Numerical integration of the global barotropic primitive equations with Hough harmonic expansion. *J. Atmos. Sci.*, **34**, 687-701.
- Longuet-Higgins, M. S., 1968: The eigenfunctions of Laplace's tidal equations over the sphere. *Phil. Trans. Roy. Soc. London*, **A262**, 511-607.
- Lorenz, E. N., 1960: Maximum simplification of the dynamic equations. *Tellus*, **12**, 243-254.
- Machenhauer, B. and E. Rasmussen, 1972: On the integration of the spectral hydrodynamical equations by a transform method. Rep. No. 3, Inst. Theoret. Meteorology, Copenhagen University, 44pp.
- Machenhauer, B. 1977: On the dynamics of gravity oscillations in a shallow water model with applications to normal mode initialization. *Beit. Phys. Atmos.*, **50**, 253-271.
- Orszag, S. A., 1970: Transform method for the calculation of vector coupled sums: Application to the spectral form of the vorticity equation. *J. Atmos. Sci.*, **27**, 890-895.
- Phillips, N. A., 1959: Numerical integration of the primitive equations on the hemisphere. *Mon. Wea. Rev.*, **87**, 333-345.

- Platzman, G. W., 1962: The analytical dynamics of the spectral vorticity equation. *J. Atmos. Sci.*, **19**, 313-328.
- Press, W. H., et al., 1992: Numerical recipes in fortran. Cambridge University Press, New York, ISBN-0-521-43064-X, 963pp.
- Tribbia, J. J., 1979: Non-linear initialization on an equatorial beta-plane. *Mon. Wea. Rev.*, **107**, 704-713.
- Wolfram, S., 1991: Mathematica. Addison Wesley publishing Company, New York, ISBN-0-201-51502-4, 961pp.

FIGURE CAPTIONS

- Figure 1. Ratio of acceptable time step using a number of terms in the Taylor series to using a single term when applied to the Lorenz low-order system. Both leapfrog and forward schemes are presented and standard deviations from 18 experiments are shown.
- Figure 2. Ratio of acceptable time step using a number of terms in the Taylor series to using a single term when applied to the low-order system on the sphere.Both leapfrog and forward schemes are presented and standard deviations from 20 experiments are shown.
- Figure 3. RMS vector wind difference between experiments and the control run evolving with integration time using the SWE with balancing of the high frequency modes. Results using two cutoff points (D1 and D7), three time steps (in minutes) and five balancing schemes (A-E) are displayed.
- Figure 4. RMS vector wind difference between experiments and the control run evolving with integration time using the SWE with balancing of the high frequency modes. All ten cutoff frequencies tested (Dx) are shown with their limiting time steps in minutes. The lower panel describes the effect of adding the third order Taylor series term. Balance scheme A is used throughout.
- Figure 5. Ratio of acceptable time step using a number of terms in the multi-level scheme (with L=3) to using a single term when applied to the Lorenz low-order system with leapfrog integration. Three acceptability conditions representing allowed errors in percent from the control are presented and standard deviations from 18 experiments are shown.
- Figure 6. Initial conditions (u, v, h) for the experiments using the SWE with a climatological zonal field. Cases a, b and c are shown.

- Figure 7. The CFL critical timestep in minutes for the first 1000 modes of the SWE model with a climatological zonal field expressed in normal mode form for the three cases considered. The table gives the values for three modal indices.
- Figure 8. Fifty case averaged RMS vector wind difference from control normalized by the standard deviation of the daily control vector wind for the normal mode SWE model shown evolving in time. All three cases using both the multilevel and leapfrog schemes are presented.
- Figure 9. Zonally averaged difference from control of the nine year monthly mean zonal wind (u) as a function of latitude wind for the normal mode SWE model. All three cases using both the multi-level and leapfrog schemes are presented.
- Figure 10. Ratio (in percent) of the differences of the standard deviation of the experimental mean zonal wind from the control to the standard deviation of the mean zonal wind of the control taken from the 108 monthly realizations, presented as a function of latitude for all three cases and both the multi-level and leapfrog schemes.

Table 2: Percentage differences of mean normal mode amplitudes from control and their standard deviations normalized by control values. For the control experiments, the normalized amplitudes are given in parentheses. The eight largest modes are presented.

mode #	control	leap-frog mean	m-level mean	leap-frog std.	m-level std.							
Case (a)												
107	(0.1970)	19.35	8.14	10.40	5.10							
108	(0.1912)	15.90	4.21	11.44	1.46							
115	(0.3200)	17.62	4.57	10.88	5.95							
116	(0.3110)	14.96	4.43	21.12	16.89							
117	(0.3152)	15.79	4.31	3.60	8.36							
118	(0.3213)	15.17	5.31	9.32	4.70							
145	(0.1043)	12.90	2.69	10.64	2.95							
146	(0.1056)	15.79	5.55	10.56	7.79							
Case (b)												
141	(0.1844)	10.39	10.39	19.75	18.75							
142	(0.2023)	17.74	17.74	17.68	10.83							
169	(0.1559)	6.06	6.06	5.25	26.64							
170	(0.1702)	12.67	12.67	17.61	14.61							
223	(0.1602)	21.01	21.01	28.93	12.54							
224	(0.1524)	16.25	16.25	30.50	17.03							
317	(0.4585)	26.74	26.74	35.30	22.67							
318	(0.4493)	25.59	25.59	34.62	21.26							
Case (c)												
71	(0.3815)	43.83	12.51	29.24	9.76							
72	(0.3876)	44.95	13.19	12.87	5.47							
99	(0.1958)	2.79	4.00	26.53	11.64							
100	(0.1936)	3.59	3.64	21.41	10.26							
111	(0.2238)	51.44	7.31	5.55	27.20							
112	(0.2489)	54.59	11.48	18.28	8.25							
371	(0.2002)	27.68	8.76	8.82	1.46							
372	(0.2002)	30.93	8.32	7.83	0.54							

Figure 1. Ratio of acceptable time step using a number of terms in the Taylor series to using a single term when applied to the Lorenz low-order system. Both leapfrog and forward schemes are presented and standard deviations from 18 experiments are shown.



Figure 2. Ratio of acceptable time step using a number of terms in the Taylor series to using a single term when applied to the low-order system on the sphere. Both leapfrog and forward schemes are presented and standard deviations from 20 experiments are shown.



Figure 3. RMS vector wind difference between experiments and the control run evolving with integration time using the SWE with balancing of the high frequency modes. Results using two cutoff points (D1 and D7), three time steps (in minutes) and five balancing schemes (A-E) are displayed.



Figure 4. RMS vector wind difference between experiments and the control run evolving with integration time using the SWE with balancing of the high frequency modes. All ten cutoff frequencies tested (Dx) are shown with their limiting time steps in minutes. The lower panel describes the effect of adding the third order Taylor series term. Balance scheme A is used throughout.



Figure 5. Ratio of acceptable time step using a number of terms in the multi-level scheme (with L=3) to using a single term when applied to the Lorenz low-order system with leapfrog integration. Three acceptability conditions representing allowed errors in percent from the control are presented and standard deviations from 18 experiments are shown.



Figure 6. Initial conditions (u, v, h) for the experiments using the SWE with a climatological zonal field. Cases a, b and c are shown.



Figure 7. The CFL critical timestep in minutes for the first 1000 modes of the SWE model with a climatological zonal field expressed in normal mode form for the three cases considered. The table gives the values for three modal indices.



Figure 8. Fifty case averaged RMS vector wind difference from control normalized by the standard deviation of the daily control vector wind for the normal mode SWE model shown evolving in time. All three cases using both the multi-level and leapfrog schemes are presented.



Figure 9. Zonally averaged difference from control of the nine year monthly mean zonal wind (u) as a function of latitude wind for the normal mode SWE model. All three cases using both the multi-level and leapfrog schemes are presented.



Figure 10. Ratio (in percent) of the differences of the standard deviation of the experimental mean zonal wind from the control to the standard deviation of the mean zonal wind of the control taken from the 108 monthly realizations, presented as a function of latitude for all three cases and both the multi-level and leapfrog schemes.

