Numerical Methods in NWP Models: Spatial Aspects

By Fedor Mesinger and Zaviša I. Janjić

Department of Meteorology, University of Belgrade, Yugoslavia

(Manuscript received 9 December 1986)

Abstract

A review is presented of spatial aspects of the numerical methods used in weather prediction models. Considerations that have and are being made in distinguishing among various finite-difference techniques are summarized. Properties that have been examined in order to arrive at an estimate and/or at an understanding of the skill numerical schemes have in simulating atmospheric dynamical processes are listed and discussed. A number of points are illustrated by brief presentations of some as yet unpublished results of the authors. One of these concerns the difference in spectral properties of the fields simulated by two horizontal advection schemes. Another is addressed at error calculations done for a recent “$\theta$-conserving” sigma system pressure gradient force scheme. Finally, issues presently attracting interest in horizontal as well as in vertical differencing are reviewed, and some recent results are summarized.

1. Approach, and properties looked for

At the time of the preceding two “Tokyo symposia” the finite difference method was the only feasible approach to spatial discretization in numerical weather prediction models. Due to the invention of the transform technique, situation has dramatically changed. Most global (typically, medium range) models are today formulated using the spectral approach. In some cases (JMA, ECMWF) even limited area (typically, short range) models have been or are being developed using the spectral approach (Tatsumi, 1985; WMO, 1985).

Advantages of the spectral technique, in terms of the lack of the phase speed errors as well as of the polar problem, are well known. In medium range forecasts, most impressive progress has been associated with the use of spectral models; in particular with that of ECMWF (Bengtsson, 1985). On the negative side, most frequently the problem—if it is a problem—of “ripples” in representing steep mountains is mentioned. Its local effects have been stated as “in some cases undesirable” (Jarraud and Baede, 1985). Perhaps, one could here add the lack of flexibility; so that some approaches, recently introduced or gaining in popularity in grid point models, are not obviously applicable in spectral modeling.

In contrast to the spectral method, the problem with the grid point techniques in a way is that there are too many choices. Therefore, a matter of primary importance, in particular when it comes to finite-difference schemes, is to define criteria which a discretization scheme should satisfy. One typically observes several, often conflicting, of these which can be summarized as follows:

(1) Simulation of dynamical processes;
(2) Computational economy, measured by operation count;
(3) Programing considerations (such as vectorization/indexing, storage requirements, code portability, etc.).

The first criterion, obviously, deserves special attention. Various properties that have been examined in order to, hopefully,
optimally satisfy it, will be here briefly discussed.

a. Conservation properties (maintenance of integral properties of the continuous equations)

The conservation of integral properties of the continuous equations in the discretized systems has attracted attention of numerical modellers at an early stage of NWP development. The importance of energy conservation emerged almost immediately in connection with the difficulties to ensure stability of computations (Lilly, 1965; Lorenz, 1960). This constraint is now very frequently imposed in the grid-point models.

More elaborate techniques based on this approach have emerged shortly after. Most of the credit for their development should be given to Arakawa, who has attempted to control a wide range of processes in the discretized system by imposing constraints on conservation of integral properties. These include the control of the nonlinear energy cascade through conservation of energy and enstrophy in the case of nondivergent flow (Arakawa, 1966; Arakawa and Lamb, 1977; Janjić, 1977, 1984; Mesinger, 1981; Sadourny, 1975a, 1975b, also in Burridge and Haseler, 1977), control of the pressure gradient force errors in the sigma coordinate system by conservation of momentum (Arakawa, 1972; Arakawa and Lamb, 1977), control of statistical properties of the distribution of potential temperature (Arakawa, 1972; Arakawa and Lamb, 1977) etc.

However, a word of caution should be added here concerning the conservation of energy and enstrophy. Namely, the effectiveness of the energy and enstrophy conservation constraints may be related to the definition of quantities which are being conserved (Janjić, 1984). Therefore, this constraint cannot be applied automatically without further consideration. For example, if the semi-staggered grids B or E in Arakawa notation are used, with both velocity components defined at the same points, the most straightforward definition of vorticity does not provide efficient control over the nonlinear energy cascade because the eigenvalues of the finite-difference operator used to calculate vorticity from the streamfunction decrease, rather than increase, as the small-scale part of the admissible wave-number range is approached. Instead, vorticity should be defined in the same way as on the staggered grid C in Arakawa notation. In this case, the nonlinear energy cascade is controlled by the conservation of the C-grid-vorticity and E (or B)-grid rotational energy. The combination of these two constraints imposes more stringent control over the nonlinear energy cascade than that which can be achieved on the C grid.

b. Finite difference phenomenology

Many examples exist of schemes designed in order to optimize simulation of a given dynamical phenomenon. Most notable properties examined from that point of view are wave propagation properties, simulation of hydrodynamic instabilities and simulation of known and expected analytic solutions.

Both phase speed and group velocity properties of finite difference schemes have been looked into for a number of simple wave solutions. Solutions considered are gravity and gravity-inertia waves (Winninghoff, 1968, and Arakawa, 1970; Mesinger and Arakawa, 1976; Janjić and Mesinger, 1984), vertically propagating internal waves (Tokioka, 1978; Arakawa and Lamb, 1977), and Rossby waves (Mesinger, 1979; Bates, 1984). For applications in other fields of computational fluid dynamics, see the review by Trefethen (1982);

Consideration of local properties of wave propagation has also received considerable attention. An example of this type of approach is the analysis of the propagation of perturbations to neighboring grid points vs. propagation to points farther away (Mesinger, 1973; Janjić, 1974, 1979).

Growth rates of solutions to difference systems are of interest from the point of view of accuracy. Spurious instabilities have been discovered which, for obvious reasons, are not welcome (Nićković, 1979; Arakawa 1984).
Many examples exist of experiments aimed at simulating known linear solutions; such as those done studying the advection problem e.g., Takacs, 1985). With the solutions to nonlinear equations available only in very special cases, comparisons against expected properties of nonlinear solutions have also been made (e.g., Sadourny, 1975a; 1975b; Janjić, 1984; Rančić, 1986; Gavrilov and Janjić, 1987).

As an illustration, the rotational energy spectra obtained by Gavrilov and Janjić, (1987) in experiments with two E grid schemes are shown in Fig. 1. Both schemes conserved energy as defined on the E grid, and the only difference between the schemes was the definition of enstrophy which was being conserved. The integrations were performed up to 12000 time steps using the “rotating-flat-square-Earth” experiment design (Janjić, 1984). The diagrams shown in the figure represent the total rotational energies within prescribed wave number ranges after averaging over the last 2000 time steps. The natural logarithm scale was used for the ordinate axes. For comparison, the k⁻³ curves are plotted on both panels.

The scheme conserving enstrophy as defined on the C grid (lower panel) shows a very steep decrease of energy content as the small scales are approached. On the other hand, in the case of the scheme conserving enstrophy as defined on the E grid (upper panel), the energy content at the small scales in much higher, and decreases slowly with increasing intensity of the wave number vector. These results are in good qualitative agreement with what might be expected from the theory outlined at the end of the preceding subsection.

c. Estimates and/or minimization of the actual error

Other methods have also been used to estimate and/or minimize the actual (as vs. formal) error. One example is the sigma coordinate pressure gradient force error. Many calculations of this error have been made; as, for example, summarized in our review paper on that subject (Mesinger and Janjić, 1985).

Some more calculations that have been done since (Mesinger, 1987), emphasize the seriousness of the problem. We shall show results of one of these calculations here; they will, at the same time, illustrate the possibilities of the error calculation approach. Previously (Mesinger, 1982), error calculations have been performed for two schemes, and for two temperature profiles. One of the profiles (the “no inversion case”) was that of temperature a linear function of ln p, with temperature values of 273 K at 800 mb, and 283 K at 1000 mb. The other (the “inversion case”) differed from the no inversion case in having an inversion between 800 and 1000 mb, with temperature in that layer still
a linear function of ln $p$, and a value of 263 K at 1000 mb. As in earlier calculations, two neighboring surface pressure points were considered located along the direction of the $x$-axis at pressures of 1000 and 800 mb, respectively. Errors were calculated for a wind point located in between the two surface pressure points, at (or, approximately at) the level of $\sigma=0.9$. Calculations were performed for $\Delta \sigma=1/5$, $\Delta \sigma=1/(3 \times 5)$, $\Delta \sigma=1/(5 \times 5)$, etc., and in the limit as $\Delta \sigma \to 0$.

The error calculations were performed for three additional pressure gradient force schemes, and one additional temperature profile. Furthermore, errors have now been calculated for all of the sigma layers. Of most interest perhaps are the results obtained for the Arakawa-Suarez (1983) “$\theta$-conserving” scheme. This is a scheme with a “local” hydrostatic equation, free of the hydrostatic consistency problem of the earlier Arakawa (1972), Brown (1974)—Phillips (1974) scheme. For error calculations, the scheme had to be expanded to include horizontal differencing; a momentum conserving horizontal differencing was used (along the lines of, e.g., Eq. (291) of Arakawa and Lamb, 1977).

Of the three profiles, the largest errors have been obtained for the inversion case. These errors, for 5, 15 and 45 layers between the top of the model and the surface, are shown in Fig. 2. Values are given in increments of geopotential ($m^2 s^{-2}$) between the two neighboring grid points, along the direction of the increasing terrain elevations. Errors can be seen which are generally much greater than those of the two schemes considered earlier. This may be a result of the feature of the Arakawa-Suarez scheme of being optimized for the isentropic atmosphere case, a rather different profile from the one considered here. Nevertheless, the errors to us appear surprisingly large. Furthermore, there is little evidence, if any, of the errors being reduced with an increase in vertical resolution.

Identification of computational modes is an illuminating tool in anticipating the actual error, since, if admissible, they are prone to be excited by various forcings of the model.

![Inversion case 5 layers](image1)

![Inversion case 15 layers](image2)

![Inversion case 45 layers](image3)

Fig. 2 Errors of the Arakawa-Suarez (1983) “$\theta$-conserving” pressure gradient force scheme, expanded using straightforward momentum conserving horizontal differencing. See text for additional details.

Space computational modes can be typically identified by looking for stationary, or slowly moving, short wavelength counterparts of the physical modes.

d. Formal properties

Consistency and (acceptable) linear stability properties are of course mandatory requirements. In spite of many efforts, benefits from an increased formal accuracy may not have been sufficiently documented. Fourth-order accuracy schemes have typically been
introduced with little regard for other properties. It may be instructive to note that examples exist of increased formal accuracy leading to an increased actual error (Mesinger, 1982).

2. Horizontal differencing

It has been clear ever since the pioneering work of Winninghoff (1968) and Arakawa (1970; also Arakawa and Lamb, 1977) that various horizontal grids are not offering the same advantages for the simulation of large- and synoptic-scale atmospheric motions. Additional arguments have been summarized by Mesinger (1981), and by Janjić and Mesinger (1984). On balance, and with presently available finite-difference schemes, it would appear that there are overwhelming reasons against the use of the non-staggered, and also the staggered D grid.

Concerning the remaining two possibilities, the staggered C grid and the semi-staggered B/E grid, for some time now we have been using the semi-staggered (E) grid (Mesinger, 1973; Janjić, 1977; also, Ničković, 1982). With the benefit of hindsight, we are reassured in the feeling that this choice has not been a mistake.

In simulation of the geostrophic adjustment process, the B/E grid has a grid-separation problem with the external and the lower internal modes, for short waves. However, a technique has been developed (Mesinger, 1973; Janjić, 1974, 1979), which to a large extent overcomes the problem (Vasiljević, 1982; Cullen, 1983; Janjić and Mesinger, 1984). The C grid, on the other hand, has a difficulty with higher internal modes, for all wave lengths. These modes are present in models with currently used vertical resolutions and it is not known whether something can be done about this problem.

For simulation of the slowly changing quasi-geostrophic motion, horizontal advection schemes which control energy cascade towards smaller scales have been developed for both the C as well as for the E grid (Arakawa and Lamb, 1981; Janjić, 1984). However, not all of the conservation properties of the two schemes are the same. Among the differences, the C grid scheme conserves potential enstrophy, while the E grid scheme conserves momentum, and imposes a more stringent constraint on the false cascade of energy towards smaller scales. Specifically, within the nondivergent part of the flow, it completely prevents the false cascade of energy into the two-grid-interval wave. Thus, on balance, properties of the E grid scheme do not appear inferior, and may be superior, to those of the C grid scheme.

Finally, a very recent study of Dragosavac and Janjić (1986) shows that with currently used horizontal resolutions the linear amplitude response of the centered B/E grid schemes to forcing by topography may be generally more accurate than that of the C grid schemes. Thus, according to this study, the B/E grid may be advantageous for the simulation of the steady solutions induced by topography, even on the synoptic scale.

In an earlier review paper of ours we have pointed out an indexing problem of the E grid when it is used for a limited area model (Janjić and Mesinger, 1984). Since that time, we have solved this problem by the use of a suitable one dimensional horizontal indexing. It achieves storage of all horizontal elements of an array in ascending memory locations, so that code vectorization on a vector computer is very efficient (Mesinger et al., 1987). On a CYBER 205 computer, with no special FORTRAN used, for a “minimum physics” model this has led to a speed-up of more than a factor of ten compared to the previous only partly vectorizable code. Similar increase in efficiency has apparently been achieved by the use of one dimensional horizontal indexing in several other grid point models.

3. Vertical differencing

One of the recent trends in vertical discretization is the use of finite-element schemes (e.g., Staniforth, 1985). Currently work on the use of finite-element schemes in the vertical is in progress also at ECMWF (e.g., Burridge and Steppeler, 1985). However, in vertical as well as in horizontal, a finite-
element scheme is of course no different than a finite-difference scheme, except for the approach used to arrive at its algorithm. Judged on the basis of their properties as discussed in the preceding section, one would expect their performance to be similar to that of high formal accuracy schemes.

There has also been a renewed interest in isentropic coordinate modeling (e.g., Bleck, 1984). Use of the isentropic coordinate enables an error-free simulation of an atmosphere at rest, in presence of orography not possible when using the customary terrain-following (sigma) coordinate.

Another approach recently advanced to overcome the sigma system problems is the use of the so called step-mountain coordinate (Mesinger, 1983). With this coordinate, mountains are formed of grid-blocks, with vertical sides. This is achieved with coordinate surfaces prescribed to remain at fixed elevations at places where they touch (and define) or intersect the ground surface. Thus, the coordinate surfaces are quasi-horizontal, and the long list of sigma system problems (e.g., Sadourny et al., 1981) is not present any longer.

An attractive feature of a step-mountain ("eta") model is that it can easily be run as a sigma system model, the only difference being the definition of ground surface grid point values of the vertical coordinate. This permits a comparison of the sigma vs. eta formulation. We have made two experiments of this kind, with a model including realistic steep mountains (steps at 290, 1112 and 2433 m, and, at one point a vertical edge rising from sea level through all three mountain layers, that is, up to the height of 2433 m; Mesinger et al., 1987). Both of these experiments have revealed a substantial amount of noise resulting from the sigma (as compared to eta) formulation.

One of our two experiments, especially as its step-mountain result, gave a rather successful simulation of the perhaps difficult "historic" Buzzi-Tibaldi (1978) case of Alpine ("Genoa") lee cyclogenesis. A parallel experiment showed that, starting with the same initial data, no cyclogenesis occurs without mountains. Still, the mountains experiment did simulate the accompanying mid-tropospheric cutoff, a phenomenon that apparently has not been reproduced in previous simulations of mountain-induced Genoa lee cyclogenesis.

For a North American region, experimental step-mountain simulations were performed for a case of March 1984, involving development of a secondary storm southeast of the Appalachians. Neither the then operational U.S. National Meteorological Center's Limited Area Forecast Model (LFM), nor the recently introduced Nested Grid Model (NGM; Hoke et al., 1985), used in a number of sensitivity experiments, were successful in simulating the redevelopment (Collins and Tracton, 1985). The step-mountain model, on the other hand, with a space resolution set up to mimic that of NGM, successfully simulated the ridging indicating the redevelopment (Mesinger et al., 1987).

References


———, 1987: Some more calculations of the sigma system pressure gradient force errors. To be published.


Phillips, N. A., 1974: Application of Arakawa’s
energy-conserving layer model to operational numerical weather prediction. Office Note 104, National Meteorological Center, NWS/NOAA, 40 pp.


