A Numerov–Galerkin Technique Applied to a Finite-Element Shallow-Water Equations Model with Enforced Conservation of Integral Invariants and Selective Lumping

I. M. NAVON

National Research Institute for Mathematical Sciences, CSIR,
P. O. Box 395, Pretoria 0001, South Africa

A new two-stage Numerov–Galerkin method is presented and applied, following a suggestion by Cullen and Morton [1] to achieve higher accuracy by combining the Galerkin product with a high-order compact (Numerov) difference approximation to derivatives in the nonlinear advection operator of the shallow-water equations. The practical applicability of the new method is supported by computational examples involving medium (10 days) and long-term (20 days) integrations of the nonlinear shallow-water equations. Small-scale noise was eliminated by periodical application of a Shuman filter to only one component of the velocity. Conservation of integral invariants is ensured by using a new augmented Lagrangian multiplier–penalty method in the Numerov–Galerkin method. Experiments were performed using the new method with differently weighted selective lumping schemes for the mass matrix. The new method exhibits a consistently higher accuracy than the single-stage Galerkin method and requires much less computational effort. A hypothesis is offered to explain the increased accuracy obtained by selective lumping schemes in the long-term numerical integrations (10–20 days) as compared to the consistent mass approach in Numerov–Galerkin schemes.

1. INTRODUCTION

In the last few years the application of finite-element methods to large-scale weather forecasting has become somewhat more extended [2–5].

The three-dimensional primitive equations of motion used by nearly all operational forecasting models (e.g., [3, 6]) include a system of nonlinear hyperbolic partial differential equations as well as different terms representing physical processes such as radiation, water–vapour, etc. This means that assessing a new numerical technique for this system is quite difficult, and it has become customary in developing new numerical methods for numerical weather prediction to study the simpler nonlinear shallow-water equations system instead (see Cullen [7]). One of the key issues in numerically solving these equations is how to treat the nonlinear advective terms (e.g., [1]).

Cullen [8] proposed a two-stage Galerkin method which was shown to have an

313

asymptotic error six times smaller than the usual single-stage Galerkin method in the limit by performing a truncation error estimate for the nonlinear term.

Cullen and Morton [1] also showed that the two-stage Galerkin approximation to the nonlinear advective operator gives a better description of nonlinear processes and has the advantage of using a better representation of the function at all stages of the calculation. It was suggested in [1] that the best technique could be to combine the Galerkin product with high-order difference approximations to derivatives.

This suggestion motivated the part of the approach in this paper where we propose a two-stage Numerov–Galerkin method applied to the nonlinear advective terms in the shallow-water equations on a limited-area domain.

In this approach the two-stage Galerkin product is combined with a high-order compact (hence the name Numerov) difference approximation to the first derivative. The compact Numerov finite-difference approximation to the first derivative has a truncation error of $O(h^{4l})$ and employs only a $2l + 1$ star of grid points (see Schwartz and Wendroff [9]). Here we used $l = 2$, i.e., and $O(h^8)$ approximation, because for $l = 1$ we recover the usual two-stage Galerkin method (e.g., Cullen and Morton [1]).

In order to recover the higher accuracy of the method it was found necessary to apply a Schuman [10] filter periodically (every 12 time-steps) to the v-component of velocity only.

In this paper it was decided not to employ the truncation error estimate for the nonlinear term employed in [1, 8, 11] as we are dealing with a multidimensional nonlinear problem and any conclusion drawn from a linear analysis could at best be only indicative. Moreover the problem has no analytic solution.

In the next section we first present the salient features of the single-stage Galerkin finite-element model using piecewise linear triangular elements on a regular mesh. The finite-element construction used here is carried out only for regular grids, if necessary achieved by a coordinate transformation.

In Section 3 we detail the two-stage Numerov-Galerkin (N.G.) method for the nonlinear advective terms, the required numerical boundary conditions and the structural simplifications introduced in the finite-element procedure by the N.G. technique.

As the two-stage Galerkin method does not conserve integral invariants (see Cullen [12]) we apply a new a posteriori technique using an augmented Lagrangian nonlinearly constrained optimization approach for enforcing the conservation of integral invariants of the shallow-water equations (see Navon and de Villiers [13]). This method is briefly exposed in Section 4.

A selective lumping used in conjunction with the N.G. method is presented in Section 5. While total lumping is known to produce less accurate results [11, 14] than the consistent mass approach, selective lumping, that is, a convex combination of the lumped and consistent mass matrices, can either provide a controlled dissipation [15, 16] or in some instances attain a better accuracy than the consistent mass matrix approach [4, 17].

In Section 6 numerical results relating to both medium-term and long-term integrations of the nonlinear shallow-water equations on a fairly standard test problem in a channel on a rotating β-plane [18, 19] are presented.

The computational efficiency of the different finite-element versions is compared and accuracy results for both medium-term (10 days) and long-term (20 days) numerical integrations are presented.

Finally in Section 7, the numerical results are discussed and a hypothesis is offered to explain the higher accuracy obtained with selective lumping N.G. schemes in the long-term integrations.

2. THE GALERKIN FINITE-ELEMENT MODEL
OF THE SHALLOW-WATER EQUATIONS

The shallow-water equations in a channel on a rotating earth can be written as follows:

$$u_t + uu_x + vu_y + \phi_x - fv = 0,$$

$$v_t + uv_x + vv_y + \phi_y + fu = 0, \qquad (1)$$

$$\phi_t + (\phi u)_x + (\phi v)_y = 0,$$

$$0 \leqslant x \leqslant L, \qquad 0 \leqslant y \leqslant D, \qquad t \geqslant 0,$$

where u and v are the velocity components in the x and y directions, respectively; f is the Coriolis parameter given by the β-plane approximation

$$f = \hat{f} + \beta \left(y - \frac{D}{2} \right) \qquad (2)$$

with $\hat{f}$ and β constants. $\phi = gh$ is the geopotential, h is the depth of the fluid and g is the acceleration of gravity.

Periodic boundary conditions are assumed in the x direction while rigid boundary conditions

$$v(x, 0, t) = v(x, D, t) = 0 \qquad (3)$$

are imposed in the y direction.

Using linear piecewise polynomials on triangular elements where over a given triangular element each variable is represented as a linear sum of interpolation function, e.g.,

$$u_{el} = \sum_{j=1}^{3} u_j(t) V_j, \qquad (4*)$$

where $u_j(t)$ represents the scalar nodal value of the variable u at the node j of the triangular element and V_j is the basis function (interpolation function) which can be defined by the coordinates of the nodes (see [2]). Here we use Galerkin formulation with the Einsteinian notation, e.g., that a repeated index implies summation with respect to that index.

The notation

$$\langle f(x, y), V_i \rangle = \sum_{\text{elements}}^{m} \iint f(x, y) V_i \, dx \, dy = \iint_{\text{global}} f(x, y) V_i \, dx \, dy$$

defines the inner product when a function is multiplied by the trial function.

Using linear piecewise polynomials on triangular elements the resulting Galerkin finite-element equations can be written as (see [4, 20])

$$M(\phi_j^{n+1} - \phi_j^n) - \frac{\Delta t}{2} K_1(\phi_j^{n+1} + \phi_j^n) = 0 \tag{4}$$

for the continuity equation, where n is the time level $(t_n = n \, \Delta t)$ and M is the mass matrix given by

$$M = \iint_A V_j V_i \, dA \tag{5}$$

and

$$K_1 = \iint_A V_i V_K u_K^* \frac{\partial V_j}{\partial x} \, dA + \iint_A V_i V_K v_K^* \frac{\partial V_j}{\partial y} \, dA, \tag{6}$$

where the element (3×3) mass matrix is

$$M = \frac{A}{12} \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}, \tag{5*}$$

where A is the area of the triangular element. K_1 gives also rise to a (3×3) element matrix and only after the assembly process we obtain the global $(N \times N)$ matrices. In our notation we already refer to the element matrices and u^* and v^* are given by

$$\begin{aligned} u^* = u^{n+1/2} = \tfrac{3}{2} u^n - \tfrac{1}{2} u^{n-1} + O(\Delta t^2), \\ v^* = v^{n+1/2} = \tfrac{3}{2} v^n - \tfrac{1}{2} v^{n-1} + O(\Delta t^2) \end{aligned} \tag{7}$$

and result from a time extrapolated Crank–Nicolson method (see [4, 21, 22]). This method is used to quasi-linearize the nonlinear advective terms.

After some algebra the u and v momentum equations obtained are

$$M(u_j^{n+1} - u_j^n) + \frac{\Delta t}{2} K_2^*(u_j^{n+1} + u_j^n) + \frac{\Delta t}{2}(K_{21}^{n+1} + K_{21}^n) + \Delta t P_2 = 0, \tag{8}$$

$$M(v_j^{n+1} - v_j^n) = \frac{\Delta t}{2} K_3^*(v_j^{n+1} + v_j^n) + \frac{\Delta t}{2}(K_{31}^{n+1} + K_{31}^n) + \Delta t P_3 = 0, \tag{9}$$

where the matrix definition used are

$$K_2^* = \iint_A u_K^* V_K V_i \frac{\partial V_j}{\partial x} \, dA + \iint_A v_K^* V_K V_i \frac{\partial V_j}{\partial y} \, dA, \tag{10}$$

$$K_{21}^{n+1} = \iint \phi_K^{n+1} \frac{\partial V_K}{\partial x} V_i \, dA, \tag{11}$$

$$P_2 = \iint_A f v_K^* V_k V_i \, dA, \tag{12}$$

$$K_3^* = \iint_A u_K^{n+1} V_K \frac{\partial V_j}{\partial x} \, dA + \iint v_K^* V_K \frac{\partial V_j}{\partial x} \, dA, \tag{13}$$

$$K_{31}^{n+1} = \iint_A \phi_K^{n+1} \frac{\partial V_K}{\partial y} V_i \, dA, \tag{14}$$

$$P_3 = \iint_A f u_K^{n+1} V_K V_i \, dA \tag{15}$$

and similar definitions for K_{31}^n, K_{21}^n, respectively.

For implementation of boundary conditions see Navon [4].

The resulting linear systems of equations were solved iteratively using either a Gauss–Seidel or an S.O.R. method.

3. The Two-Stage Numerov–Galerkin Scheme for Advective Terms

The two-stage Galerkin method (see [1]) is applied to the nonlinear advective terms of form $v \nabla v$. This involved calculating an intermediate approximation Z to the first derivative $\partial_x v$ (i.e., the closest piecewise linear approximation to $\partial_x v$) before incorporating it into the final Galerkin approximation to $v \, \partial v / \partial x$.

As shown by Cullen and Morton [1] if Z denotes the intermedite approximation to $\partial_x v$ we have

$$\tfrac{1}{6} Z_{j-1} + \tfrac{2}{3} Z_j + \tfrac{1}{6} Z_{j+1} = \tfrac{1}{2} h^{-1} (V_{j+1} - V_{j-1}) \tag{16}$$

and the second and final stage where $W = u \, \partial v / \partial x$ is

$$\tfrac{1}{6} W_{j-1} + \tfrac{2}{3} W_j + \tfrac{1}{6} W_{j-1}$$
$$= \tfrac{1}{12} (U_{j-1} Z_{j-1} + U_{j-1} Z_j + U_j Z_{j-1} + U_j Z_{j+1} + U_{j+1} Z_j + U_{j+1} Z_{j+1}) + \tfrac{1}{2} U_j Z_j. \tag{17}$$

By assuming Fourier modes, then in the asymptotic limit, Cullen and Morton [1] found that the truncation error associated with the two-stage Galerkin method is almost six times smaller than that of the single-stage Galerkin method applied to the advection operator.

In our approach we combine the two-stage Galerkin product concept with a high-order compact implicit (hence the name Numerov) difference approximation to the first derivative $\partial_x v$ in the advection operator. The compact implicit finite-difference approximation to the first derivative has a truncation error of $O(h^{4l})$ and employs a compact star of only $2l + 1$ grid points at the price of solving a $2l + 1$ banded matrix system (see Schwartz and Wendroff [9], Navon and Riphagen [23]).

As our main aim is to extract a higher accuracy from the advective operator in the two-stage Numerov–Galerkin approach, we found out that it was necessary to use an intermediate approximation of $\partial_x v$ of order $O(h^8)$ or $l = 2$ for the Schwartz–Wendroff [9] symbol.

Using $l = 1$ would have retrieved the usual two-stage Galerkin method of [1].

The concise expression of the intermediate compact Numerov finite difference approximation to $\partial v / \partial x$ of order $O(h^8)$ is given by

$$\frac{1}{70} \left[\left(\frac{\partial u}{\partial x} \right)_{i+2} + 16 \left(\frac{\partial u}{\partial x} \right)_{i+1} + 36 \left(\frac{\partial u}{\partial x} \right)_{i} + 16 \left(\frac{\partial u}{\partial x} \right)_{i-1} + \left(\frac{\partial u}{\partial x} \right)_{i-2} \right]$$

$$= \frac{1}{84h} [-5u_{i-2} - 32u_{i-1} + 32u_{i+1} + 5u_{i+2}], \qquad h = \Delta x = \Delta y. \tag{18}$$

This necessitates the solution of a pentadiagonal system the matrix of which has the following entries for noncyclic boundary conditions

$$\frac{1}{70} \begin{bmatrix} 36 & 16 & 1 & & & & 0 \\ 16 & 36 & 16 & 1 & & & \\ 1 & 16 & 36 & 16 & 1 & & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \ddots & \ddots & 16 \\ 0 & & & 1 & 16 & 36 \end{bmatrix} \begin{bmatrix} \\ \\ \dfrac{\partial v}{\partial x} \\ \\ \end{bmatrix}$$

$$= \frac{1}{84h} \begin{bmatrix} 0 \\ -5v_0 - 32v_1 + 32v_3 + 5v_4 \\ -5v_1 - 32v_2 - 32v_4 + 5v_5 \\ \vdots \\ -5v_{N_y-4} - 32v_{N_y-2} + 32v_{N_y-1} + 5v_{N_y} \\ -5v_{N_y-3} - 32v_{N_y-2} + 32v_{N_y} + 5v_{N_y+1} \\ 0 \end{bmatrix}, \qquad j = 1, 2, ..., N_y \tag{19}$$

Here we interpolate v_0 and v_{N_y+1} using

$$v_0 = 4v_1 - 4v_2 + 4v_3 - v_4, \tag{20}$$

$$v_{N_y+1} = 4v_{N_y} - 6v_{N_y-1} + 4v_{N_y-2} - v_{N_y-3}, \tag{21}$$

while for the intermediate expression Z we have

$$Z_1 = \left(\frac{\partial v}{\partial y}\right)_1 = \frac{-25v_1 + 48v_2 - 36v_3 - 16v_4 - 3v_5}{12h} + O(h^4), \tag{22}$$

$$Z_{Ny} = \left(\frac{\partial v}{\partial y}\right)_{N_y} = \frac{3v_{N_y-4} - 16v_{N_y-3} + 36V_{N_y-2} - 48V_{N_y-1} + 25v_{N_y}}{12h} + O(h^4). \tag{23}$$

In the second stage of the Numerov–Galerkin method we have to solve a tridiagonal system of the form

$$\frac{1}{6}\begin{bmatrix} 4 & 1 & & & & \\ 1 & 4 & 1 & & & \\ & \ddots & \ddots & \ddots & & 0 \\ & & & & & 1 \\ 0 & & & & 1 & 4 \end{bmatrix}\begin{bmatrix} \\ w_j \\ \\ \end{bmatrix} = \frac{1}{12}\begin{bmatrix} v_{j-1}Z_{j-1} + v_jZ_{j-1} + v_{j-1}Z_j + \\ v_{j+1}Z_j + v_jZ_{j+1} + v_{j+1}Z_{j+1} + 6v_jZ_j \end{bmatrix}.$$

For this second stage we need the values of Z_0 and Z_{N_y+1} which we interpolate as

$$Z_0 = (-25v_0 + 48v_1 - 36v_2 + 16v_3 - 3v_4)/12h, \tag{25}$$

$$Z_{N_y+1} = (3v_{N_y-3} - 16v_{N_y-2} + 36v_{N_y-1} - 48v_{N_y} + 25v_{N_y+1})/12h. \tag{26}$$

For the influence of these boundary approximations on the overall accuracy, see Gustafsson [24].

A pentadiagonal solver was used for solving the system given by (19) (see Von-Rosenberg [25]).

In the case of cyclic boundary conditions a cyclic pentadiagonal matrix system has to be solved and here we generalized an algorithm due to Ahlberg, Nielsen and Walsh [26]. The general shape of the cyclic pentadiagonal matrix is

$$Ax = \begin{bmatrix} c & b & a & & & & a & b \\ b & c & b & a & & & & a \\ a & b & c & b & a & & & 0 \\ 0 & & \ddots & \ddots & \ddots & & & \\ \hline a & & & & \ddots & b & c & b \\ b & a & & & & a & b & c \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ 0 \\ x_{n-2} \\ \hline x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \\ d_{n-2} \\ \hline d_{n-1} \\ d_n \end{bmatrix} \tag{27}$$

Using this type of matrix splitting let

$$
E = \begin{bmatrix}
c & b & a & & & & 0 \\
b & c & b & a & & & \\
a & b & c & b & a & & \\
& \ddots & \ddots & \ddots & \ddots & \ddots & \\
& & \ddots & \ddots & \ddots & \ddots & \ddots \\
0 & & & a & b & c & b \\
& & & & a & b & c
\end{bmatrix}
$$

be an $(n-2)$ by $(n-2)$ matrix and

$$
\hat{X} = \begin{bmatrix} x_1 \\ \vdots \\ \vdots \\ x_{n-2} \end{bmatrix}, \qquad
g_n f_n = \begin{bmatrix}
a & & b \\
& & a \\
& 0 & \\
a & & \\
b & & a
\end{bmatrix}_{(n-2)\times 2}, \tag{29}
$$

$$
\hat{d} = \begin{bmatrix} d_1 \\ \vdots \\ \vdots \\ d_{n-2} \end{bmatrix} \tag{30}
$$

Then we have

$$
E\hat{X} + f_n \begin{bmatrix} X_{n-1} \\ X_n \end{bmatrix} = \hat{d}, \tag{31}
$$

$$
g_n^T \hat{X} + \begin{bmatrix} c & b \\ b & c \end{bmatrix} \begin{bmatrix} x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} d_{n-1} \\ d_n \end{bmatrix}. \tag{32}
$$

If we substitute (31), i.e.,

$$
\hat{X} = E^{-1}\hat{d} - E^{-1}f_n \begin{bmatrix} X_{n-1} \\ X_n \end{bmatrix}, \tag{33}
$$

in (32), we obtain

$$
g_n^T E^{-1}\hat{d} - g_n^T E^{-1}f_n \begin{bmatrix} X_{n-1} \\ X_n \end{bmatrix} + \begin{bmatrix} c & b \\ b & c \end{bmatrix} \begin{bmatrix} X_{n-1} \\ X_n \end{bmatrix} = \begin{bmatrix} d_{n-1} \\ d_n \end{bmatrix} \tag{34}
$$

or

$$
\left\{ \begin{bmatrix} c & b \\ b & c \end{bmatrix} - g_n^T E^{-1}f_n \right\} \begin{bmatrix} X_{n-1} \\ X_n \end{bmatrix} = \begin{bmatrix} d_{n-1} \\ d_n \end{bmatrix} - g_n^T E^{-1}\hat{d} \tag{35}
$$

which finally results in

$$\begin{bmatrix} X_{n-1} \\ X_n \end{bmatrix} = \left\{ \begin{bmatrix} c & b \\ b & c \end{bmatrix} - g_n^T E^{-1} f_n \right\}^{-1} \left\{ \begin{bmatrix} d_{n-1} \\ d_n \end{bmatrix} - g_n^T E^{-1} \hat{d} \right\}. \tag{36}$$

3.1. Changes Introduced by the Numerov–Galerkin Technique to the Finite-Element Matrices

Due to the Numerov–Galerkin technique the momentum equations for the u and v-components of velocity undergo structural changes. We will denote by Z_{xu} the intermediate Numerov approximation representing the first-stage derivative calculation ∂_{xu} and by Z_{xv} the corresponding intermediate approximation to $\partial_x v$ and use similar notation for the y derivatives.

The u-momentum equation takes the form

$$\langle (u_j^{n+1} - u_j^n) V_i, V_j \rangle + \Delta t [\langle (u Z_{xu})_j^* V_j, V_i \rangle$$
$$+ \langle (v Z_{yu})_j^* V_j, V_i \rangle] - \Delta t [\langle f_j V_j^* V_j, V_i \rangle]$$
$$+ \frac{\Delta t}{2} \left[\left\langle \phi_K^{n+1} \frac{\partial V_K}{\partial x}, V_i \right\rangle + \left\langle \phi_K^n \frac{\partial V_K}{\partial x}, V_j \right\rangle \right] = 0. \tag{37}$$

Using the matrix notation of Section 2, i.e.,

$$M = \iint_A V_j V_i \, dA \tag{38}$$

for the mass matrix, we obtain the following matrix equation:

$$M\{(u_j^{n+1} - u_j^n) + \Delta t (u Z_{xu})_j^* + (V Z_{yu})_j^* - f_j V_j^*]\} = \Delta t \overline{K_{21}} \tag{39}$$

In a similar manner we obtain the following modified V-momentum equation:

$$M(v_j^{n+1} - v_j^n) + \Delta t (v Z_{yv})_j + (u^{n+1} Z_{xv})_j + f_j u_j^{u+1} = \Delta t \overline{K_{31}}, \tag{40}$$

where we denote

$$\overline{K_{21}} = \tfrac{1}{2}(K_{21}^{n+1} + K_{21}^n) \overline{K_{31}} = \tfrac{1}{2}(K_{31}^{n+1} + K_{31}^n). \tag{41}$$

Compared to the single-stage Galerkin finite-element method we observe that Eqs. (39) and (40) result in a simplification for the momentum equations as the mass matrix M is calculated only once, and the solution process is simplified compared to the single-stage Galerkin where we have to solve

$$\left(M + \frac{\Delta t}{2} K_2^* \right) (u_j^{n+1} - u_j^n) = \Delta t (\overline{K_{21}} + P_2 + K_2^* u_j^n). \tag{42}$$

4. AN AUGMENTED LAGRANGIAN COMBINED PENALTY–MULTIPLIER METHOD FOR ENFORCING DISCRETE CONSERVATION OF INTEGRAL INVARIANTS

It is well known that the Galerkin finite-element method has the property of satisfying certain conservation laws of the shallow-water equations.

However, the two-stage Galerkin method does not conserve quadratic invariants (Cullen [12]).

Here we briefly present a new a posteriori method for enforcing conservation of integral invariants of the shallow-water equations, namely, potential enstrophy (Z), total energy (E) and total mass (H), using augmented Lagrangian methods due to Bertsekas [27], resulting from viewing the problem of enforcing conservation laws as a nonlinearly constrained optimization problem with nonlinear equality constraints (see Navon and de Villiers [13]).

It should be mentioned that strictly speaking neither Z nor E are *quadratic* invariants so that the procedure to be outlined would be required for the Galerkin method also.

We start by defining the following functional:

$$f = \sum_{j=1}^{N_x} \sum_{k=1}^{N_y} \left[\tilde{\alpha}(u - \tilde{u})^2 + \tilde{\alpha}(v - \tilde{v})^2 + \tilde{\beta}(h - \tilde{h})^2 \right]_{jk}^n, \qquad (43)$$

where

$$L = N_x \, \Delta x, \qquad D = N_y \, \Delta y, \qquad \Delta x = \Delta y = h, \qquad (44)$$

where $(\tilde{u}, \tilde{v}, \tilde{h})_{jk}^n$ are the grid variables predicted by the N.G. finite-element method shallow-water equations solver at the nth time level and $(u, v, h)_{jk}^n$ are the grid variables adjusted by the Augmented Lagrangian method so as to satisfy the conservation of integral invariants up to a given accuracy.

$\tilde{\alpha}$ and $\tilde{\beta}$ are weights which can be chosen as in Sasaki [28] and for the functional f we adopt the same three basic principles as in [28].

The augmented Lagrangian takes the form

$$L(\mathbf{x}, \mathbf{U}, \mathbf{r}) = f(\mathbf{x}) + \mathbf{U}^T \mathbf{e}(\mathbf{x}) + \frac{1}{2r} |\mathbf{e}(\mathbf{x})|^2, \qquad (45)$$

while considering the problem

$$\begin{cases} \text{minimize } f(\mathbf{x}) \\ \text{subject to } m \text{ nonlinear equality constraints } \mathbf{e}(\mathbf{x}) = 0, \end{cases}$$

where

$$\mathbf{x} = (\tilde{u}_{11} \cdots \tilde{u}_{N_x,N_y}, \tilde{v}_{11} \cdots \tilde{v}_{N_x,N_y}, \tilde{h}_{11} \cdots \tilde{h}_{N_x,N_y})^n. \qquad (46)$$

In our case

$$e(x) = E^n - E^0,$$
$$= Z^n - Z^0, \tag{47}$$
$$= H^n - H^0,$$

where

$$E^n = \frac{1}{2} \sum_{j=1}^{N_x} \sum_{k=1}^{N_y} [\tilde{h}(\tilde{u}^2 + \tilde{v}^2) + g\tilde{h}^2]_{jk}^n \, \varDelta x \, \varDelta y, \tag{48}$$

$$Z^n = \frac{1}{2} \sum_{j=1}^{N_x} \sum_{k=1}^{N_y} \left\{ \left[\frac{\partial \tilde{v}/\partial x - \partial \tilde{u}/\partial y + f}{\tilde{h}} \right]_{jk}^n \right\}^2 \, \varDelta x \, \varDelta y, \tag{49}$$

$$H^n = \sum_{j=1}^{N_x} \sum_{k=1}^{N_y} \tilde{h}_{jk} \, \varDelta x \, \varDelta y, \tag{50}$$

and E^0, Z^0, H^0 are the initial values of the integral constraints respectively. In Eq. (45) U is the m-complement multiplier vector and r is the penalty parameter (different penalty parameters can be used for the different equality constraints).

The basic idea of the penalty multiplier method is to solve the nonlinearly constrained minimization problem by performing a sequence of unconstrained minimizations of the following problem:

$$\min_{x \in \mathbb{R}^n} L_{r_K}(x, U_K) = f(x) + \sum_{i=1}^{n} U_K^i e_i(x) + \frac{1}{2r_K} |e(x)|^2. \tag{51}$$

The stopping criteria

$$\|\nabla_x L_{r_K}(x_K, U_K)\| \leqslant \{\eta_K\} \|e(x_K)\| \tag{52}$$

with $\{\eta_K\}$ a decreasing sequence tending to zero was used.

The updating of multipliers and penalties is done as in [27]. We used the method of inexact minimization of the augmented Lagrangian and a scaling of both the constraints and the variables. The method was only activated once one of the following relations were violated:

$$\left| \frac{E^n - E^0}{E^0} \right| \leqslant \delta_E, \qquad \left| \frac{H^n - H^0}{H^0} \right| \leqslant \delta_H, \qquad \text{and} \qquad \left| \frac{Z^n - Z^0}{Z^0} \right| \leqslant \delta_Z. \tag{53}$$

As an illustration we give the penalty–multiplier algorithm minimizing the augmented Lagrangian in Eq. (51).

Algorithm steps

Preparatory step. Select an initial vector of multipliers U_0, based either on prior knowledge or else start with the zero vector. Select penalty multipliers $r_0^i > 0$ as in [27] and a sequence $\{\eta_K\}$ with $\{\eta_K\}_{K \to \infty} \to 0$.

Step 1.   Given a multiplier vector $\mathbf{U}_K$, penalty parameters r_k^i and y_K find a vector $\mathbf{x}_K$ satisfying

$$\|\nabla_x L_{r_k}(\mathbf{x}_K, \mathbf{U}_K)\| \leqslant \eta_K \|e(\mathbf{x}_K)\| \tag{54}$$

by solving an inexact unconstrained optimization problem.

Step 2.   If

$$|e_i(\mathbf{x}_K)| \leqslant \varepsilon_i, \qquad i = 1,\ldots, m \qquad \text{then} \tag{55}$$

Stop. Use the new corrected fields obtained from the vector $\mathbf{x}_K$ as starting values for a new time-step prediction using the N.G. shallow-water equations solver. Otherwise go to 3.

Step 3.   Update the multiplier vector using

$$U_{k+1} = U_k + r_K^{-1} e(\mathbf{x}_K). \tag{56}$$

Update penalty parameters as in [27], i.e., $r_{k+1}^i \in (0, r_K^i)$. Select $\eta_{k+1} > 0$ s.t. $\eta_{k+1} < \eta_k$ using a formula of the type

$$\eta_K = l^k, \tag{57}$$

where $0 < l < 1$, e.g., $l = (0.8)^k$.

Return to Step 1.

5. The Selective Lumping N.G. Method

Full lumping of the mass matrix in a finite-element scheme is a way of obtaining a less expensive scheme due to the diagonalization of the mass matrix.

In applications (many engineering problems, [29, 30]), where a time-dependent problem is solved as a way of reaching a steady state, lumping does not affect the accuracy and saves computer time. Accuracy of the Numerov derivatives depends on the correct form of the mass matrix. If the mass is lumped, the explicit finite-difference formula is needed. If the mass is lumped in the Galerkin product, it produces damping, but loses accuracy in medium term integrations (see also [11, 14]).

However, when employing a selective lumping method which consists of a convex combination of the lumped and the consistent mass matrices such that

$$M_{SL} = \alpha M_C + (1 - \alpha) M_L, \qquad 0 \leqslant \alpha \leqslant 1 \tag{58}$$

where M_C and M_L are the consistent and lumped mass matrices, respectively, it was found out that in some occurrences [4, 17, 31] selective lumping may result in a certain accuracy gain. In other schemes [15, 16] selective lumping is used as a selective damping and diffusion operator.

In this study it was decided to conduct numerical experiments in order to assess the influence of full lumping and selective lumping of the mass matrix on the medium-term and long-term accuracy of the nonlinear shallow-water equations using the Numerov–Galerkin technique.

As we are using a higher-order approximation in the nonlinear advection operator, we could also ascertain the conclusions in [11] where quadratic finite elements were used.

In the numerical experiments to be described we varied the selective lumping coefficient α, from $\alpha = 1$ (consistent mass matrix) to $\alpha = 0$ (full lumping of the mass matrix) by using the intermediate values of $\alpha = 0.9, 0.7, 0.5$ and 0.3.

6. Numerical Experiments with the Nonlinear Shallow-Water Equations in Two Dimensions

(a)   To compare the computational efficiency and the accuracy of the proposed new Numerov–Galerkin technique with those of the single-stage Galerkin method (Navon [4], Navon and Muller [20]) the test problem used is the one for the nonlinear shallow-water equations in a channel on a rotating earth, i.e., the initial height field condition No. 1 used by Grammeltvedt [18], which has been tested by different researchers [1, 19], etc.). This initial field condition can be written as

$$gh(x, y) = \phi = g \left\{ H_0 + H_1 \tanh\left(\frac{9(D/2 - y)}{2D}\right) \right.$$

$$\left. + H_2 \operatorname{sech}^2\left(\frac{9(D/2 - y)}{D}\right) \frac{\sin 2\pi x}{L} \right\}. \tag{59}$$

The initial velocity fields were derived from the initial height field via the geostrophic relationship, i.e.,

$$fu = -\phi_y \quad \text{or} \quad u = \left(\frac{-g}{f}\right)\frac{\partial h}{\partial y}, \tag{60}$$

$$fv = \phi_x \quad \text{or} \quad v = \left(\frac{g}{f}\right)\frac{\partial h}{\partial x}. \tag{61}$$

The constants used were

$$\begin{aligned}
&L = 6000 \text{ km}, &\quad &g = 10 \text{ m s}^{-2}, \\
&D = 4400 \text{ km}, &\quad &H_0 = 2000 \text{ m}, \\
&\hat{f} = 10^{-4} \text{ s}^{-1}, &\quad &H_1 = -220 \text{ m}, \\
&\beta = 1.5 \times 10^{-11} \text{ s}^{-1} \text{ m}^{-1}, &\quad &H_2 = 133 \text{ m}.
\end{aligned} \tag{62}$$

The time and space increments used were

$$h = \Delta x = \Delta y = 400 \text{ km}, \qquad \Delta t = 1800 \text{ s}, \tag{63}$$

respectively.

All the calculations were carried out on a CYBER CDC 750. A series of numerical experiments including medium-range (10 days) and long-term (20 days) integrations of the nonlinear shallow-water were carried out. These include

(i) The single-stage Galerkin f.e.m. (also called FESW) for medium- and long-term integrations.

(ii) The Numerov–Galerkin (N.G.) method including a posteriori enforcing of integral invariants by the method of augmented Lagrangian nonlinear constrained optimization and periodic filtering of short-wave noise by a Shuman filter application [10, 34] to the v-component of velocity (every 24 time-steps).

(iii) The Schuman filter is the simplest filter designed to filter out short-wavelength components. In its simplest form it is a one-dimensional three-point operator given by

$$\bar{f}_j = (1 + S) f_j + \frac{S}{2} (f_{j+1} - f_{j-1}), \tag{64}$$

where $x = j\Delta x$ is the field to be smoothed and S is a constant. Applied to the harmonic form $f = A e^{ikx}$ where the wave number is $k = 2\pi/L$ the result is

$$\bar{f} = Rf, \tag{65}$$

where R is the response function given by

$$R = 1 - 2 \sin^2(\pi \Delta x/L). \tag{66}$$

If

$$S = \tfrac{1}{2}, \qquad R(\tfrac{1}{2}) = \cos^2(\pi \Delta x/L) \tag{67}$$

which means total elimination of two gridlength waves.

(iv) The N.G. method as above but using a varying selective lumping of the mass matrix using the following values of the selective lumping coefficient

$$\alpha = 1.0, \qquad \alpha = 0.5, \qquad \text{and} \qquad \alpha = 0.0.$$

This method will be denoted N.G.S.L. (i.e., selective lumping). The case $\alpha = 1.0$ corresponds to a consistent mass matrix.

(b) *Computational efficiency.* The methods FESW, N.G. and N.G.S.L. (with different $\alpha - s$) were compared for computational efficiency by finding the run times in seconds per full time-step. These results are displayed in the first entry of Table I.

TABLE I

Computational Efficiency of the Different Finite-Element Methods

Method	Run time per full time step(sec)	Total run time for 20 days numerical integr.(sec)
Single-stage Galerkin FESW	0.38	364.82
Numerov–Galerkin (NG) Consistent-mass matrix $\alpha = 1$	0.25	243.679
N.G. Method Selective lumping $\alpha = 0.9$	0.24	231.340
N.G. Method Selective lumping $\alpha = 0.7$	0.22	213.103
N.G. Method Selective lumping $\alpha = 0.5$	0.20	192.760
N.G. Method Selective lumping $\alpha = 0.3$	0.19	182.727
N.G. Lumped mass matrix $\alpha = 0$	0.15	146.280

The second entry of Table I displays the total run time for 20 days of numerical integration of the nonlinear shallow-water equations.

The Numerov–Galerkin technique turns out to be significantly more economical of CPU time than the single-stage Galerkin method due to the simplifications in the solution/assembly process.

When the N.G. method is combined with selective lumping of the mass matrix, further computational economy results as the selective lumping coefficient α decreases from 1 to 0 (full lumping). This is due to the increased diagonal dominance of the mass matrix which speeds up the convergence of the iterative methods (Gauss–Seidel, S.O.R.) used for solving the linear systems of equations of the N.G. method.

By surveying Table I a total economy of 40% per full-time-step is achieved by the Numerov–Galerkin method as compared with FESW if we take the CPU time as our measure.

(c) *Accuracy tests.* In order to provide a basis for comparison between the new Numerov–Galerkin method and its selective lumping variants on one hand and the single-stage Galerkin (FESW) on the other hand, and in the absence of an analytic solution to the nonlinear shallow-water equations, a very fine grid (50 km) finite-difference approximation was taken to give the definitive result.

The fine-mesh results are denoted by W_{EX} where

$$W = (u, v, \phi)^T. \tag{68}$$

TABLE II

Accuracy Results for the Different Finite-Element Methods
$\|E_{AP}\|/\|W_{EX}\|$, $\Delta x = \Delta y = 400$ km, $\Delta t = 1800$ sec
Medium-Term Runs

Method time (days)	F.E.S.W.	N.G. Consistent	N.G./S.L. $\alpha = 0.9$	N.G./S.L. $\alpha = 0.7$	N.G./S.L. $\alpha = 0.5$	N.G./S.L. $\alpha = 0.3$	N.G. $\alpha = 0$ Lumping
1	$1.06 \times$ E-3	$8.43 \times$ E-4	$8.42 \times$ E-4	$8.50 \times$ E-4	$9.01 \times$ E-4	$9.42 \times$ E-4	$9.63 \times$ E-4
2	$1.85 \times$ E-3	$1.33 \times$ E-3	$1.36 \times$ E-3	$1.44 \times$ E-3	$1.49 \times$ E-3	$1.58 \times$ E-3	$1.64 \times$ E-3
3	$2.43 \times$ E-3	$1.63 \times$ E-3	$1.72 \times$ E-3	$1.83 \times$ E-3	$1.94 \times$ E-3	$2.03 \times$ E-3	$2.15 \times$ E-3
4	$2.86 \times$ E-3	$1.82 \times$ E-3	$1.88 \times$ E-3	$2.00 \times$ E-3	$2.11 \times$ E-3	$2.25 \times$ E-3	$2.41 \times$ E-3
5	$3.20 \times$ E-3	$2.11 \times$ E-3	$2.15 \times$ E-3	$2.19 \times$ E-3	$2.24 \times$ E-3	$2.35 \times$ E-3	$2.50 \times$ E-3
6	$3.38 \times$ E-3	$2.30 \times$ E-3	$2.34 \times$ E-3	$2.33 \times$ E-3	$2.33 \times$ E-3	$2.35 \times$ E-3	$2.42 \times$ E-3
7	$3.34 \times$ E-3	$2.13 \times$ E-3	$2.15 \times$ E-3	$2.13 \times$ E-3	$2.13 \times$ E-3	$2.14 \times$ E-3	$2.17 \times$ E-3
8	$3.45 \times$ E-3	$2.00 \times$ E-3	$2.02 \times$ E-3	$2.01 \times$ E-3	$2.05 \times$ E-3	$2.09 \times$ E-3	$2.14 \times$ E-3
9	$3.33 \times$ E-3	$2.03 \times$ E-3	$2.06 \times$ E-3	$2.11 \times$ E-3	$2.19 \times$ E-3	$2.25 \times$ E-3	$2.33 \times$ E-3
10	$3.16 \times$ E-3	$2.10 \times$ E-3	$2.07 \times$ E-3	$2.09 \times$ E-3	$2.10 \times$ E-3	$2.17 \times$ E-3	$2.26 \times$ E-3

Representing the FESW/SSG and the Numerov–Galerkin finite-element methods by W_{FE} and W_{NG}, respectively, the error as in [19] is given by

$$\varepsilon_{FE} = W_{FE} - W_{EX} \tag{69}$$

and the relative error by (see Gustafsson [19], Fairweather and Navon [32])

$$\text{relative error} = \frac{\|\varepsilon_{FE}\|}{\|W_{EX}\|}, \tag{70}$$

where the norm $\|\ \|$ is defined as follows:

Define a Hilbert space H by considering all vector functions of the form (68).

$$W_{jk} = W_{j,N_x+k}, \qquad v_{j,0} = v_{j,N_y} = 0. \tag{71}$$

The inner produce of two vectors α and β is defined by

$$(\alpha, \beta) = \Delta x\, \Delta y \sum_{j=1}^{N_x} \left\{ \sum_{k=1}^{N_y-1} \alpha_{jk}^T \beta_{jk} + \tfrac{1}{2}(\alpha_{j0}^T \beta_{j0} + \alpha_{j,N_y}^T \beta_{j,N_y}) \right\} \tag{72}$$

and the norm by

$$\|\alpha\|^2 = (\alpha, \alpha). \tag{73}$$

The relative errors for both medium-range (10 days) and long-term (10–20 days) are displayed in Tables II and III, respectively, for the following finite-element methods: FESW, N.G. (consistent), N.G.S.L. ($\alpha = 0.9$), N.G.S.L. ($\alpha = 0.7$), N.G.S.L ($\alpha = 0.5$), N.G.S.L. ($\alpha = 0.3$) and N.G. ($\alpha = 0$) (full lumping).

TABLE III

Accuracy Results for the Different Finite-Element Methods
$\|E_{\mathrm{AP}}\|/\|W_{\mathrm{EX}}\|$, $\Delta x = \Delta y = 400$ km, $\Delta t = 1800$ sec
Long-Term Runs

Method time (days)	F.E.S.W.	N.G. Consistent	N.G./S.L. $\alpha = 0.9$	N.G./S.L. $\alpha = 0.7$	N.G./S.L. $\alpha = 0.5$	N.G./S.L. $\alpha = 0.3$	N.G. $\alpha = 0$ Lumping
11	$3.20 \times$ E-3	$2.34 \times$ E-3	$2.26 \times$ E-3	$2.18 \times$ E-3	$2.13 \times$ E-3	$2.14 \times$ E-3	$2.21 \times$ E-3
12	$3.22 \times$ E-3	$2.41 \times$ E-3	$2.35 \times$ E-3	$2.28 \times$ E-3	$2.17 \times$ E-3	$2.12 \times$ E-3	$2.14 \times$ E-3
13	$3.06 \times$ E-3	$2.30 \times$ E-3	$2.28 \times$ E-3	$2.20 \times$ E-3	$2.10 \times$ E-3	$2.07 \times$ E-3	$2.06 \times$ E-3
14	$3.08 \times$ E-3	$2.40 \times$ E-3	$2.38 \times$ E-3	$2.31 \times$ E-3	$2.19 \times$ E-3	$2.11 \times$ E-3	$2.10 \times$ E-3
15	$3.17 \times$ E-3	$2.58 \times$ E-3	$2.57 \times$ E-3	$2.45 \times$ E-3	$2.27 \times$ E-3	$2.17 \times$ E-3	$2.13 \times$ E-3
16	$3.33 \times$ E-3	$2.79 \times$ E-3	$2.81 \times$ E-3	$2.62 \times$ E-3	$2.42 \times$ E-3	$2.29 \times$ E-3	$2.20 \times$ E-3
17	$3.59 \times$ E-3	$2.97 \times$ E-3	$3.01 \times$ E-3	$2.86 \times$ E-3	$2.67 \times$ E-3	$2.50 \times$ E-3	$2.32 \times$ E-3
18	$3.65 \times$ E-3	$3.00 \times$ E-3	$3.03 \times$ E-3	$2.91 \times$ E-3	$2.78 \times$ E-3	$2.62 \times$ E-3	$2.46 \times$ E-3
19	$3.56 \times$ E-3	$2.93 \times$ E-3	$2.99 \times$ E-3	$2.81 \times$ E-3	$2.79 \times$ E-3	$2.67 \times$ E-3	$2.55 \times$ E-3
20	$3.49 \times$ E-3	$2.87 \times$ E-3	$3.03 \times$ E-3	$2.82 \times$ E-3	$2.80 =$ E-3	$2.72 \times$ E-3	$2.62 \times$ E-3

(d) *Conservation of integral invariants using the method of augmented Lagrangian.* In Figs. 1–6 we display the time variation of the three integral invariants of total mass (H), total energy (E) and potential enstrophy (Z) of the shallow-water equations as functions of their initial values using either the augmented Lagrangian method with combined penalty and multipliers or the penalties only.

In Figs. 1 and 2 we display the time variation of the integral invariants for the N.G. (consistent) case with penalty and multipliers or penalties only, respectively, for medium-term integrations (10 days). Figures 3 and 4 display the same results for the case N.G.S.L. ($\alpha = 0.5$). Figures 5 and 6 for N.G. full lumping case ($\alpha = 0.0$). In Figs. 7–9 we display the long-term time variation (up to 20 days) of the three integral invariants of the shallow-water equations as functions of their initial values —using only the combined multiplier penalties augmented Lagrangian method. This is done for the cases N.G. ($\alpha = 1$), N.G.S.L. ($\alpha = 0.5$) and N.G. ($\alpha = 0.0$), respectively.

(e) *Results of the numerical integrations.* In Fig. 10 we display the initial distribution of the height field depicted by isopleths drawn every 50 m.

In Fig. 11 we display a 20-day forecast of the height-field using the single-stage Galerkin (FESW) method, while in Figs. 12–15 we display a 20-day forecast of the height using the consistent Numerov–Galerkin method, N.G.S.L. ($\alpha = 0.9$), N.G.S.L. ($\alpha = 0.5$) and the lumped version of the Numerov–Galerkin method, respectively.
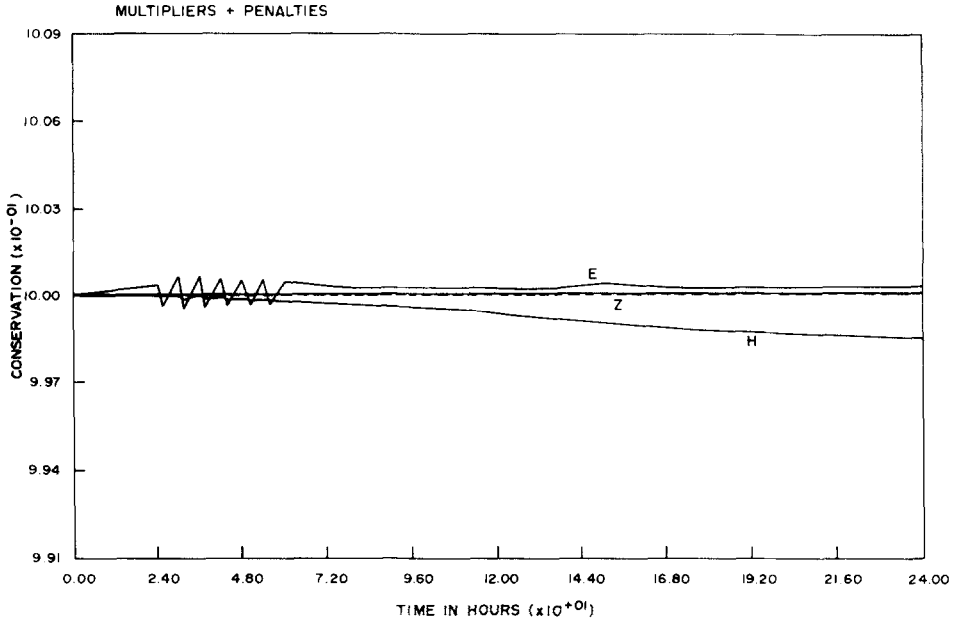
I. M. NAVON

MULTIPLIERS + PENALTIES

FIG. 1.   Time variation of total mass (H), total energy (E) and potential enstrophy (Z) as ratios of their initial values with combined penalty–multiplier nonlinearly constrained optimization using the Numerov–Galerkin finite-element model. $(\alpha = 1)$ Inflection points correspond to adjustment times.
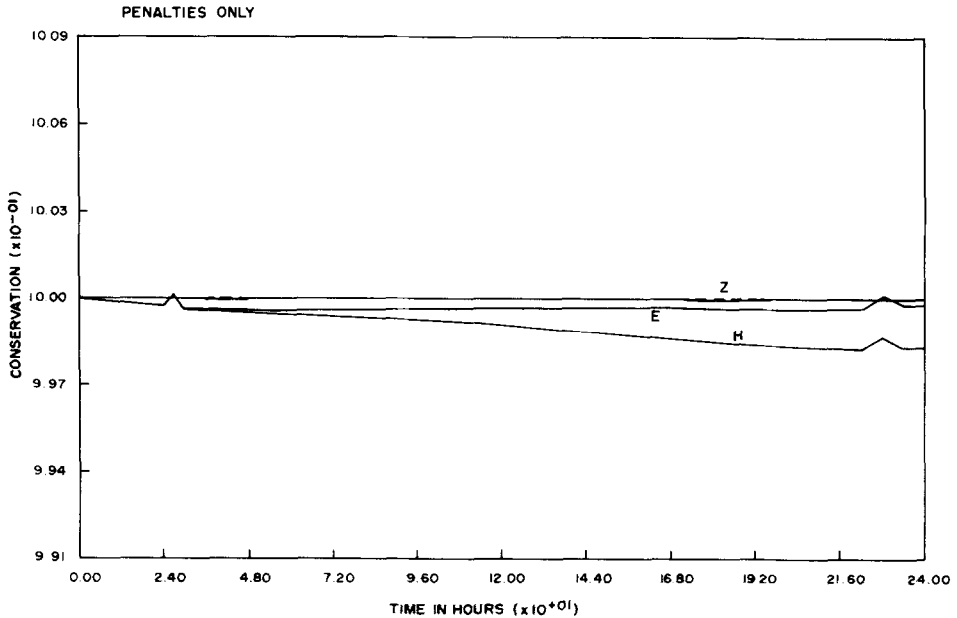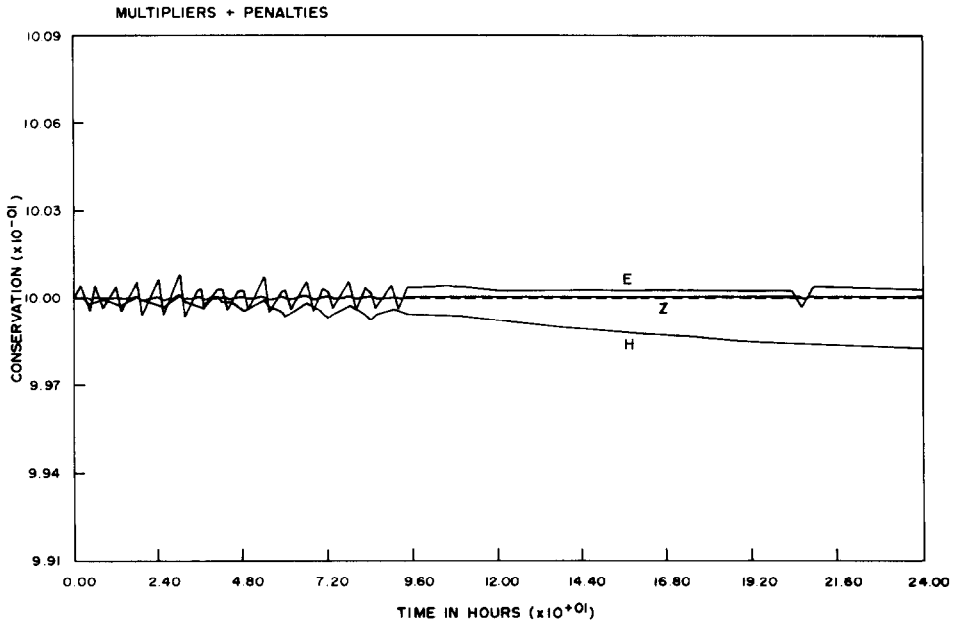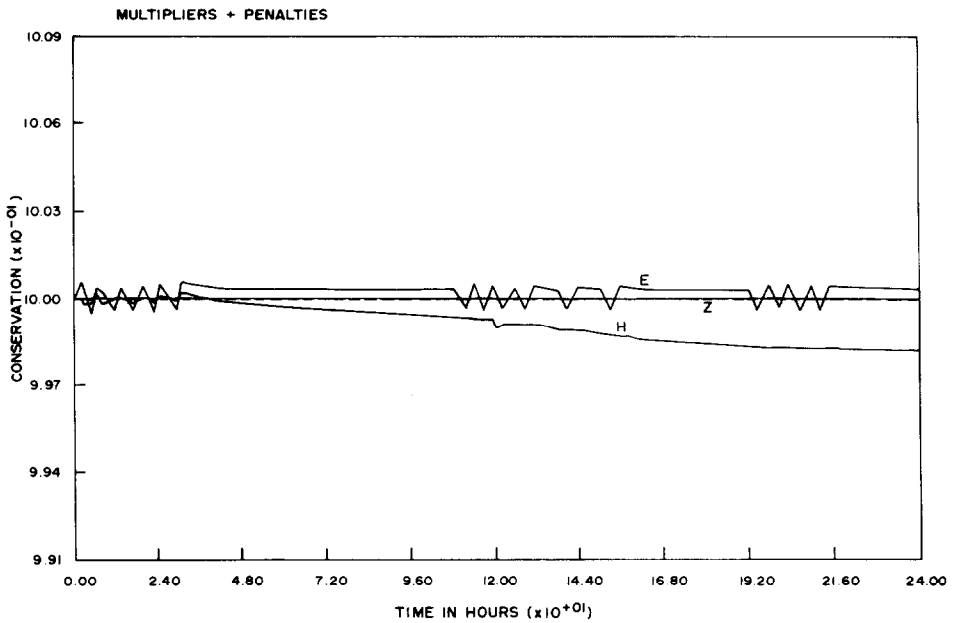
PENALTIES ONLY

FIG. 2.   Same as Fig. 1 but using a quadratic penalty method.

MULTIPLIERS + PENALTIES

FIG. 3. Time variation of total mass (H), total energy (E) and potential enstrophy (Z) as ratios of their values using the combined penalty-multiplier method (Numerov–Galerkin selective lumping/N.G./S.L. with $\alpha = 0.5$).

PENALTIES ONLY
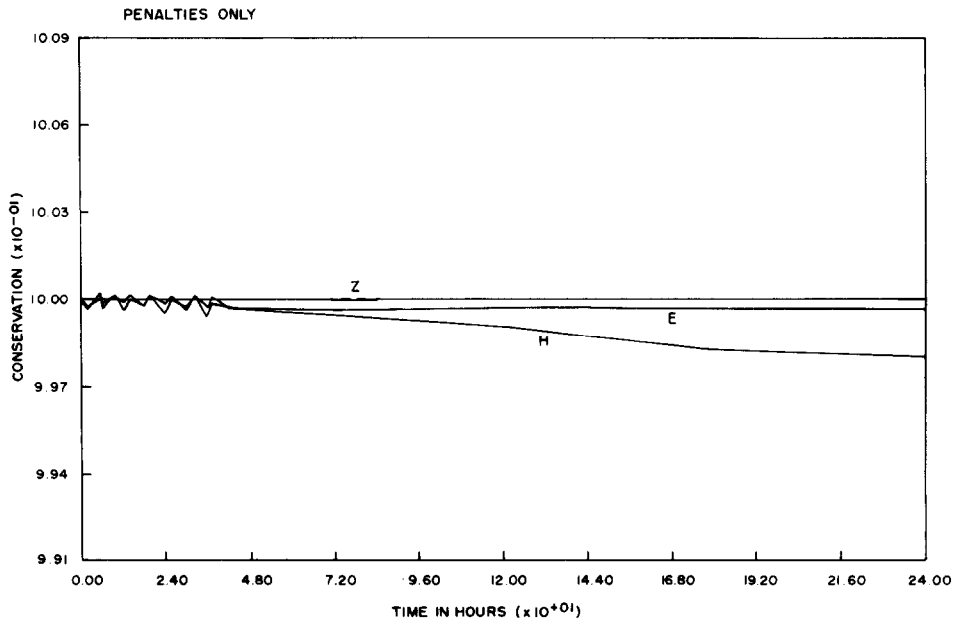
FIG. 4. Same as Fig. 3 but using a quadratic penalty method.

FIG. 5. Time variation of total mass (H), total energy (E) and potential enstrophy (Z) as ratios of their initial values using the combined penalty–multiplier method (N.G./S.L. with $\alpha = 0$, i.e., total lumping case).
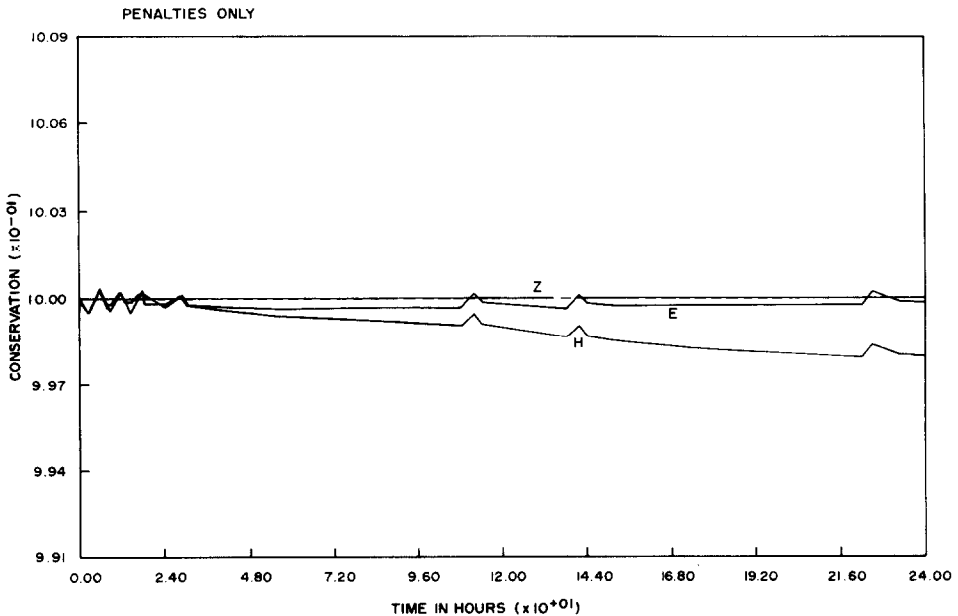
FIG. 6. Same as Fig. 11 but using a quadratic penalty method.
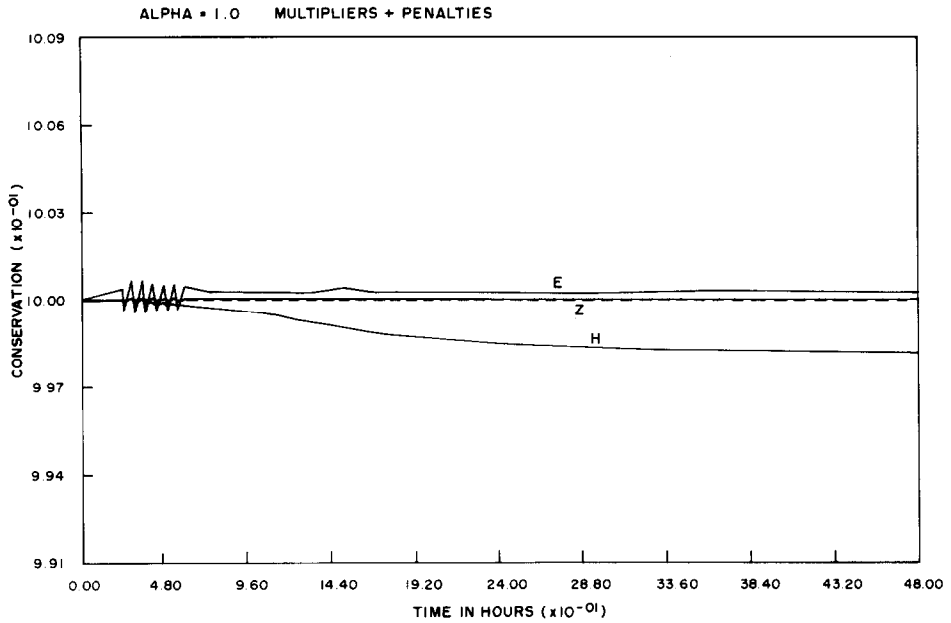
FIG. 7. Long-term (20 days) time variation of total mass (H), total energy (E) and potential enstrophy (Z) as ratios of their initial values using the combined penalty multiplier method and the consistent N.G. f.e.m. $(\alpha = 1)$.
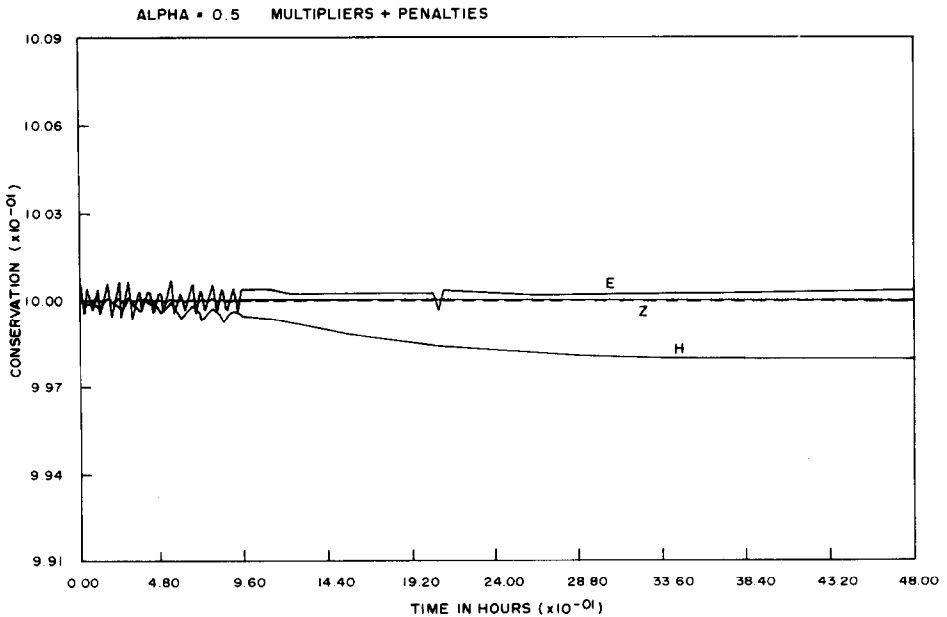
FIG. 8. Same as Fig. 13, but with the N.G./S.L. method using $\alpha = 0.5$.
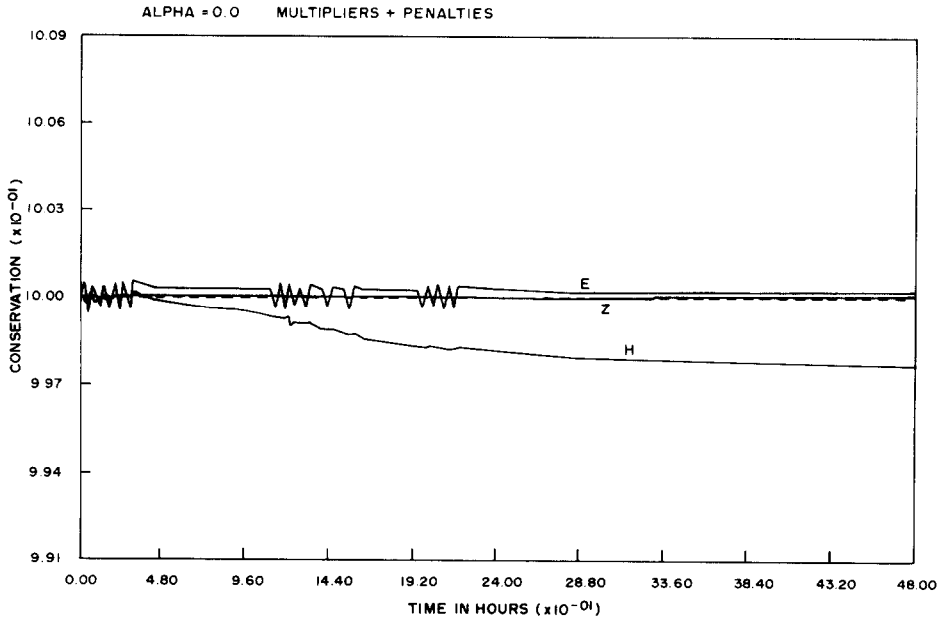
FIG. 9.   Same as Fig. 13, but with the N.G./S.L. method using $\alpha = 0.0$, i.e., the case of full lumping of the mass matrix.
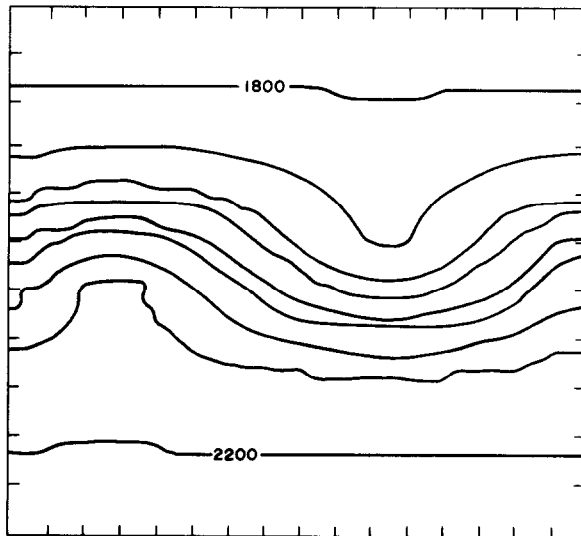
FIG. 10.   The initial distribution of the height field depicted by isopleths drans at 50-m intervals. $\Delta x = \Delta y = 400$ km, $\Delta t = 1800$ s.
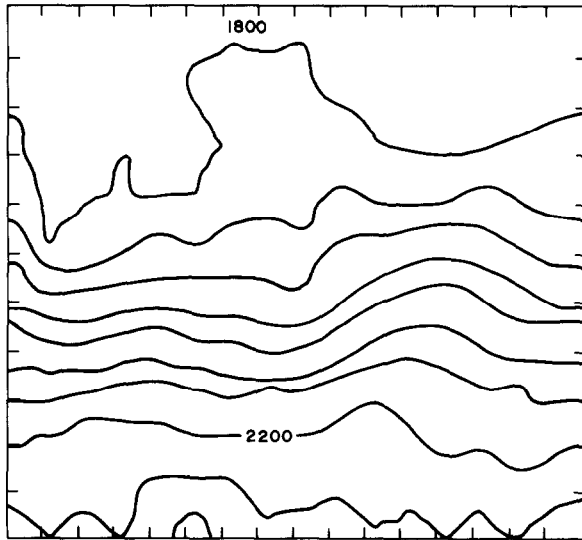
FIG. 11. A 20-day forecast of the height field using the single-stage Galerkin finite-element method. $\Delta x = \Delta y = 400$ km, $\Delta t = 1800$ s.
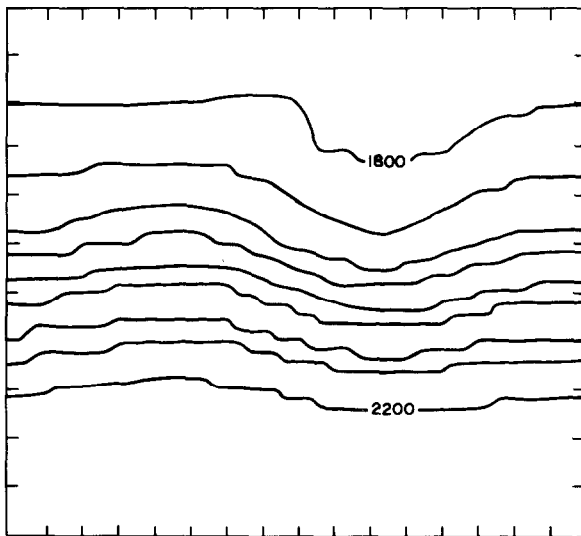
FIG. 12. A 20-day forecast of the height field using the two-stage Numerov–Galerkin f.e.m. with a consistent mass matrix ($\alpha = 1$).
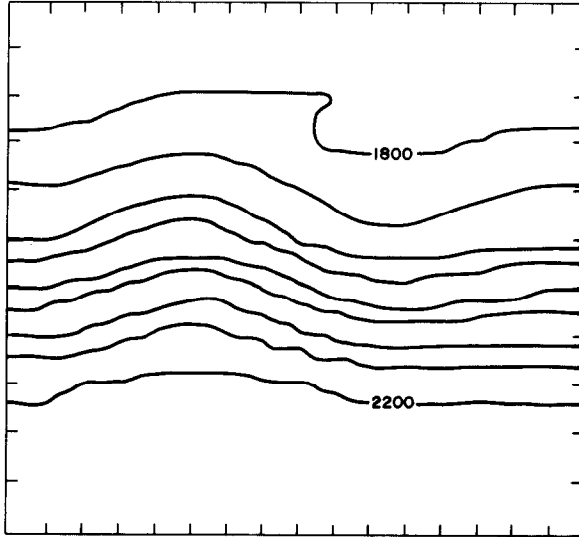
FIG. 13.   A 20 day forecast of the height field using the two-stage N.G./S.L. f.e.m. with selective lumping coefficient $\alpha = 0.9$.
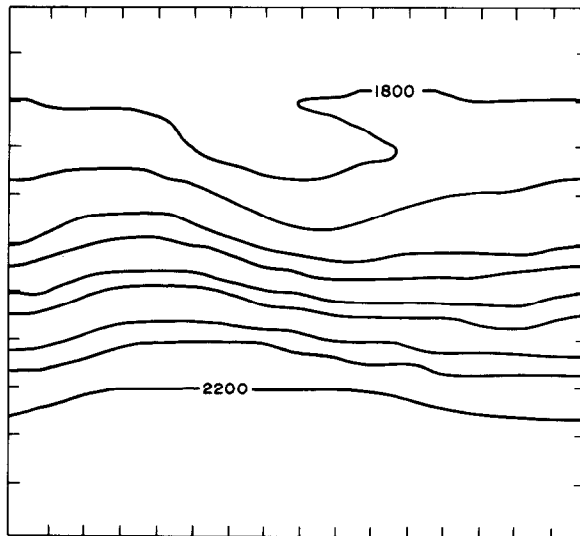

FIG. 14.   A 20-day forecast of the height field using the two-stage N.G./S.L. f.e.m. with selective lumping coefficient $\alpha = 0.5$.
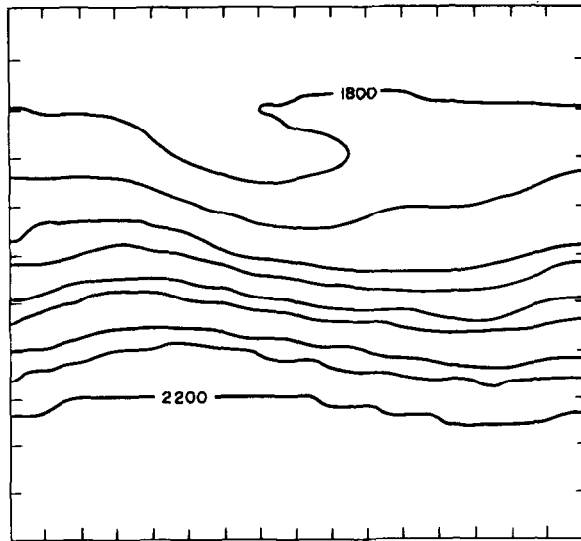
FIG. 15.   A 20-day forecast of the height field using the two-stage Numerov–Galerkin f.e.m. method with a full lumping of the mass matrix ($\alpha = 0$).

7. DISCUSSION OF NUMERICAL RESULTS AND CONCLUSIONS

The Numerov–Galerkin scheme is computationally more efficient than the single-stage Galerkin (FESW). Its selective lumping variants result in a considerable saving of CPU time.

The numerical experiments confirm that there is a considerable and consistent improvement in accuracy when using the Numerov–Galerkin method for solving the nonlinear shallow-water equations. The improvement is quite considerable both for medium-range (up to 10 days) and for long-range experiments (10 days to 20 days numerical integration) and is about 30% in the average in the relative-error norm.

As far as the selective lumping and full-lumping versions of the Numerov–Galerkin method are concerned, their accuracy is consistently lower than that of the consistent-mass Numerov–Galerkin method for the medium-range numerical integrations (up to 10 days). This confirms results due to [11, 14].

However, for the long-range runs (10–20 days) there is a marqued tendency for different selective lumping versions of the Numerov–Galerkin method including the full lumping method to have a better accuracy than the consistent Numerov–Galerkin method. This improvement varies between 10% and 20%. After day 13 the best accuracy is attained by the fully lumped ($\alpha = 0$) version of the Numerov–Galerkin method. It is the author's conjecture that this improvement in accuracy is due to the lumping operator acting as a dissipative operator filtering out short-wave noise for long-term runs.

In this sense the Numerov–Galerkin method, due to the high-order compact approximation of the first derivative, behaves like the first-order wave equation using either quadratic or Hermite cubic elements, i.e., it contains an accurate solution with an additional spurious wave [11, 33]. The growth of the noise caused by this spurious wave is well controlled for medium-range integrations of the nonlinear shallow-water equations by the periodic application of a Schuman filter to the v-component of velocity.

The numerical results for long-term integrations point out that the periodic application of the Schuman filter is not sufficient but that the use of the lumping operator takes care of the additional short-wave noise accumulated.

The augmented Lagrangian combined penalty–multiplier method copes efficiently with the task of enforcing the integral constraints. The method has to be applied every 12–15 time-steps on the first days of the integration and it is practically not required at all in long-term integrations (10–20 days). As such it is very economical in CPU time.

It appears that by using the Numerov–Galerkin technique with linear elements on a regular grid one is extracting a higher-order accuracy, when using it in conjunction with an augmented Lagrangian method for enforcing conservation of integral invariants of the nonlinear shallow-water equations and with a periodic application of a Shuman filter for the v-component of velocity. As such this method is recommended for treating the nonlinear advective terms when using the finite-element Galerkin algorithm for solving the nonlinear shallow-water equations in meteorological applications.

REFERENCES

1. M. J. P. CULLEN AND K. W. MORTON, *J. Comput. Phys.* **34** (1980), 245–267.
2. M. J. P. CULLEN AND C. D. HALL, *Quart. J. Roy. Meteor. Soc.* **102** (1979), 571–592.
3. A. STANIFORTH AND R. W. DALEY, *Monthly Weather Rev.* **107** (1979), 107–121.
4. I. M. NAVON, *Appl. Math. Modelling* **3** (1979), 337–348.
5. A. STANIFORTH, A review of the application of the finite-element method to meteorological flows, *in* "Finite-Element Flow Analysis," Proc. of the Fourth International Symposium on Finite-Element Methods in Flow Problems, Tokyo 1982, pp. 835–842, University of Tokyo Press, Tokyo, 1982.
6. G. J. HALTINER AND R. T. WILLIAMS, "Numerical Prediction and Dynamic Meteorology," Wiley, New York, 1980.
7. M. J. P. CULLEN, Numerical methods in weather forecasting, *in* "Numerical Methods in Applied Fluid Dynamics" (B. Hunt, Ed.), pp. 313–333, Academic Press, London, 1980.
8. M. J. P. CULLEN, *J. Inst. Math. Appl.* **13** (1974), 233–247.
9. B. SWARTZ AND B. WENDROFF, *SIAM J. Numer. Anal.* **11** (1974), 979–993.
10. F. G. SCHUMAN, Joint Numerical Weather Prediction Unit. Tech. Memo No. 7, 1955.

11. M. J. P. CULLEN, *J. Comput. Phys.* **45** (1982), 221–245.

12. M. J. P. CULLEN, The finite-element method, *in* "Numerical Methods Used in Atmosphere Models, Vol. 2," ICSU/WMO GARP Publication Series No. 17, World Meteorological Organization, Geneva, Switzerland, 1979.

13. I. M. NAVON AND R. DE VILLIERS, Combined penalty–multiplier nonlinearly constrained optimization methods for enforcing conservation of integral invariants in shallow-water equations models. TWISK 268, Technical Report, NRIMS, CSIR, Pretoria, South Africa, 1982.

14. P. M. GRESHO, R. L. LEE, AND R. L. SANI, *in* "Finite-Elements in Fluids," Vol. 3, Wiley, New York, 1977.

15. M. KAWAHARA, H. HIRANA, K. TSUBOTA, AND K. INAGAKI, *Int. J. Numer. Methods Fluid* **2** (1982), 89–112.

16. H. HIRANO, H. HARA, AND M. KAWAHARA, *in* "Finite Element Flow Analysis," Proc. Fourth International Symposium on Finite-Element Methods in Fluid Flow Problems, Tokyo 1982, pp. 121–128, University of Tokyo Press, Tokyo, 1982.

17. K. ISHIHARA, *Mem. Numer. Math.* **4** (1977), 1–23.

18. A. GRAMMELTVEDT, *Monthly Weather Rev.* **97** (1969), 384–404.

19. B. GUSTAFSSON, *J. Comput. Phys.* **7** (1971), 239–254.

20. I. M. NAVON AND U. MULLER, *Advan. Engrg. Software* **1** (1979), 77–86.

21. J. DOUGLAS AND T. DUPONT, *SIAM J. Numer. Anal.* **7** (1970), 575–626.

22. H. H. WANG, H. HALPERN, J. DOUGLAS, AND T. DUPONT. *Monthly Weather Rev.* **100** (1972), 738–746.

23. I. M. NAVON AND H. A. RIPHAGEN, *Monthly Weather Rev.* **107** (1979), 1107–1127.

24. B. GUSTAFSSON, *Math. Comput.* **29** (1975), 396–406.

25. DALE U. VON ROSENBERG, "Methods for the Numerical Solution of Partial Differential Equations, American Elsevier, New York, 1969.

26. H. H. AHLBERG, E. N. NILSON AND J. L. WALSH, The theory of splines and their application, *in* "Mathematics in Science and Engineering," Vol. 38, Academic Press, New York, 1967.

27. D. P. BERTSEKAS, Penalty and multiplier methods, *in* "Nonlinear Optimization, Theory and Algorithms" (L. C. W. Dixon, E. Spedicato, and G. P. Szego, Eds.), Chap. 11, pp. 253–278, Birkhäuser, Boston, 1980.

28. Y. SASAKI, *J. Comput. Phys.* **21** (1976), 270–278.

29. J. FRIED AND D. S. MALKUS, *Int. J. Solid Structures* **11** (1975), 461.

30. S. W. KEY AND Z. E. BEISINGER, *in* "Proc. Third Conf. Matrix Methods in Structural Engineering," Air Force Inst. of Technology, Wright Patterson AFB, Ohio, 1971.

31. K. W. MORTON, "Initial Value Problems by Finite Difference and Other Methods in State of the Art in Numerical Analysis" (D. Jacobs, Ed.), Academic Press, New York, 1977.

32. G. FAIRWEATHER AND I. M. NAVON, *J. Comput. Phys.* **37** (1980), 1–18.

33. G. W. HEDSTROM, *SIAM J. Numer. Anal.* **16** (1979), 385–393.

34. F. G. SHUMAN, *Monthly Weather Rev.* **85** (1957), 357–361.