Pergamon

Computers Math. Applic. Vol. 27, No. 4, pp. 99–112, 1994
Printed in Great Britain.
0898-1221/94 $6.00 + 0.00

0898-1221(93)E0014-W

Stability, Accuracy and Efficiency of a Semi-Implicit Method for Three-Dimensional Shallow Water Flow

V. CASULLI AND E. CATTANI
Dipartimento di Matematica, Università di Trento
38050 Povo (TN), Italy

(Received and accepted November 1992)

Abstract—The stability analysis, the accuracy and the efficiency of a semi-implicit finite difference scheme for the numerical solution of a three-dimensional shallow water model are presented and discussed. The governing equations are the three-dimensional Reynolds equations in which pressure is assumed to be hydrostatic. The pressure gradient in the momentum equations and the velocities in the vertically integrated continuity equation are discretized with the θ-method, with θ being an implicitness parameter. It is shown that the method is stable for $\frac{1}{2} \le \theta \le 1$, unstable for $\theta < \frac{1}{2}$ and highest accuracy and efficiency is achieved when $\theta = \frac{1}{2}$. The resulting algorithm is mass conservative and naturally allows for the simulation of flooding and drying of tidal flats.

1. INTRODUCTION

A characteristic analysis of the two-dimensional, vertically integrated shallow water equations has shown that the celerity term $\sqrt{gH}$ in the equation for the characteristic cone arises from the barotropic pressure gradient in the momentum equations and from the velocity derivatives in the free surface equation [1]. Results of this analysis have led to a practical semi-implicit method which has been proven to be unconditionally stable, and which has proved to be very useful in several applications [2,3].

Recently, the semi-implicit finite difference method for the two-dimensional shallow water equations has been extended to the three-dimensional shallow water equations [4]. The Courant-Friedrich-Lewy (CFL) stability condition is not required by this method, because the barotropic pressure gradient in the momentum equations and the velocities in the vertically integrated continuity equation are finite-differenced implicitly.

Numerical experiments of the three-dimensional shallow water equations have shown that this algorithm is stable and is highly efficient. Moreover, when only one vertical layer is specified, this method reduces, as a special case, to the semi-implicit method for the two-dimensional vertically integrated shallow water equations as described by Casulli [1]. The resulting two- and three-dimensional methods, however, are only first-order accurate in time, and introduce some artificial damping.

Wave damping is a well-known and undesired effect of implicit methods when large time steps are used. It can be shown that the damping error can be reduced to a minimum by the use of Crank-Nicolson type of time averaging.

The characteristic analysis of the governing equations was first carried out by Casulli and Greenspan [5] on the inviscid, two-dimensional compressible flow equations in order to derive

The authors gratefully acknowledge the discussions with Enrico Bertolazzi, Maria Letizia Bertotti and Bruno Firmani who offered many valuable comments and suggestions.

an efficient semi-implicit finite difference method whose stability is independent of the speed of sound. Patnaik *et al.* [6] have improved the accuracy of this method by introducing an implicitness parameter θ. For $\theta = 1$, this method degenerates into the semi-implicit method of Casulli and Greenspan, but for $\theta = \frac{1}{2}$, the method remains stable independently of the speed of sound, and results in an improved time accuracy.

In order to obtain a more accurate method for the shallow water equations, in the present paper the scheme of Casulli and Cheng [4] is reconsidered with the inclusion of an implicitness parameter θ. When $\theta = 1$, this method reverts to the original semi-implicit scheme proposed in [4]. When $\theta = \frac{1}{2}$, the pressure gradient in the momentum equations and the velocities in the free surface equation are evaluated as an average of their values at time levels n and $n+1$, so that this discretization is second-order accurate in time. The present algorithm is shown to be stable for $\frac{1}{2} \le \theta \le 1$, and highest accuracy and efficiency is achieved for $\theta = \frac{1}{2}$. When $\theta < \frac{1}{2}$, the method is unstable.

The convective and viscous terms are conveniently discretized, as in [4], by using an Eulerian-Lagrangian approach. The computational efficiency of the resulting algorithm is competitive with other numerical schemes commonly used in three-dimensional models (see, e.g., [7–11]).

2. GOVERNING EQUATIONS

The governing three-dimensional, primitive variable equations describing constant density, free surface flows in estuarine embayments and coastal seas can be derived from the Navier-Stokes equations after turbulent averaging and under the simplifying assumption that the pressure is hydrostatic [4]. Such equations have the following form

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + w\frac{\partial u}{\partial z} = -g\frac{\partial \eta}{\partial x} + \mu\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) + \frac{\partial}{\partial z}\left(\nu\frac{\partial u}{\partial z}\right), \tag{1}$$

$$\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + w\frac{\partial v}{\partial z} = -g\frac{\partial \eta}{\partial y} + \mu\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2}\right) + \frac{\partial}{\partial z}\left(\nu\frac{\partial v}{\partial z}\right), \tag{2}$$

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0, \tag{3}$$

where $u(x,y,z,t)$, $v(x,y,z,t)$ and $w(x,y,z,t)$ are the velocity components in the horizontal x, y and vertical z directions; t is the time; $\eta(x,y,t)$ is the water surface elevation measured from the undisturbed water surface; g is the gravitational acceleration and μ and ν are the coefficients of horizontal and vertical eddy viscosity, respectively.

Integrating the continuity equation over the depth and using a kinematic condition at the free surface leads to the following free surface equation

$$\frac{\partial \eta}{\partial t} + \frac{\partial}{\partial x}\left[\int_{-h}^{\eta} u\,dz\right] + \frac{\partial}{\partial y}\left[\int_{-h}^{\eta} v\,dz\right] = 0, \tag{4}$$

where $h(x,y)$ is the water depth measured from the undisturbed water surface and $H(x,y,t)$ is the total water depth, given by $H(x,y,t) = h(x,y) + \eta(x,y,t)$.

The boundary conditions at the free surface are specified by the prescribed wind stresses, (τ_x^w, τ_y^w),

$$\nu\frac{\partial u}{\partial z} = \tau_x^w, \qquad \nu\frac{\partial v}{\partial z} = \tau_y^w, \tag{5}$$

and the boundary conditions at the sediment-water interface are given by specifying the bottom stress in a form of the Manning-Chezy formula

$$\nu\frac{\partial u}{\partial z} = \gamma\, u, \qquad \nu\frac{\partial v}{\partial z} = \gamma\, v, \tag{6}$$

where $\gamma = \dfrac{g\sqrt{u^2+v^2}}{C_z^2}$ and C_z is the Chezy friction coefficient. With properly specified initial and boundary conditions, equations (1)–(4) form a well-posed initial-boundary value problem for three-dimensional shallow water flow.

3. FREE SURFACE WAVE DAMPING

In order to analyze rigorously the stability and the accuracy of this method consider, first, the following linearized one-dimensional shallow water equations

$$\frac{\partial u}{\partial t} + g\,\frac{\partial \eta}{\partial x} = 0, \tag{7}$$

$$\frac{\partial \eta}{\partial t} + H\,\frac{\partial u}{\partial x} = 0, \tag{8}$$

where g and H, for the time being, are assumed to be constant. The celerity associated with equations (7), (8) is given by $\sqrt{gH}$. A staggered grid of size Δx is introduced and the discrete variables u and η are defined at alternate locations. The θ-method for equations (7), (8) is then defined as follows

$$u_{i+\frac{1}{2}}^{n+1} = u_{i+\frac{1}{2}}^{n} - g\frac{\Delta t}{\Delta x}\left[\theta\left(\eta_{i+1}^{n+1} - \eta_i^{n+1}\right) + (1-\theta)\left(\eta_{i+1}^{n} - \eta_i^{n}\right)\right], \tag{9}$$

$$\eta_i^{n+1} = \eta_i^n - H\frac{\Delta t}{\Delta x}\left[\theta\left(u_{i+\frac{1}{2}}^{n+1} - u_{i-\frac{1}{2}}^{n+1}\right) + (1-\theta)\left(u_{i+\frac{1}{2}}^{n} - u_{i-\frac{1}{2}}^{n}\right)\right]. \tag{10}$$

The stability analysis of the scheme (9), (10) is carried out using the von Neumann method under the assumption that the differential equations (7), (8) are defined on an infinite spatial domain, or with periodic boundary conditions on a finite domain. By changing the variable η to $z = \eta\sqrt{\dfrac{g}{H}}$, equations (9), (10) become

$$\begin{aligned} u_{i+\frac{1}{2}}^{n+1} + \theta\sqrt{gH}\,\frac{\Delta t}{\Delta x}\left[z_{i+1}^{n+1} - z_i^{n+1}\right] &= u_{i+\frac{1}{2}}^{n} - (1-\theta)\sqrt{gH}\,\frac{\Delta t}{\Delta x}\left[z_{i+1}^{n} - z_i^{n}\right], \\ z_i^{n+1} + \theta\sqrt{gH}\,\frac{\Delta t}{\Delta x}\left[u_{i+\frac{1}{2}}^{n+1} - u_{i-\frac{1}{2}}^{n+1}\right] &= z_i^{n} - (1-\theta)\sqrt{gH}\,\frac{\Delta t}{\Delta x}\left[u_{i+\frac{1}{2}}^{n} - u_{i-\frac{1}{2}}^{n}\right]. \end{aligned} \tag{11}$$

Next, a Fourier mode is introduced for each field variable u and z. Specifically, $u_{i+\frac{1}{2}}^n$ and z_i^n are replaced in (11) by $\hat{u}^n e^{I(i+\frac{1}{2})\alpha}$ and $\hat{z}^n e^{Ii\alpha}$, respectively, where $\hat{u}^n$ and $\hat{z}^n$ are the amplitude functions of u and z at the time level t^n, $I = \sqrt{-1}$, and α is the phase angle. Thus, after some simplifications, equations (11) become

$$\begin{aligned} \hat{u}^{n+1} + \theta\,\Phi\left[\hat{z}^{n+1}\left(e^{I(\alpha/2)} - e^{-I(\alpha/2)}\right)\right] &= \hat{u}^n - (1-\theta)\,\Phi\left[\hat{Z}^n\left(e^{I(\alpha/2)} - e^{-I(\alpha/2)}\right)\right], \\ \hat{z}^{n+1} + \theta\,\Phi\left[\hat{u}^{n+1}\left(e^{I(\alpha/2)} - e^{-I(\alpha/2)}\right)\right] &= \hat{z}^n - (1-\theta)\,\Phi\left[\hat{u}^n\left(e^{I(\alpha/2)} - e^{-I(\alpha/2)}\right)\right], \end{aligned} \tag{12}$$

where $\Phi = \sqrt{gH}\,(\Delta t/\Delta x)$. Since $e^{I(\alpha/2)} - e^{-I(\alpha/2)} = 2I\sin(\alpha/2)$, by setting $p = 2\Phi\sin(\alpha/2)$, equations (12) in matrix notation can be written as

$$\mathbf{P}\hat{\mathbf{W}}^{n+1} = \mathbf{Q}\hat{\mathbf{W}}^n,$$

where

$$\hat{\mathbf{W}}^n = \begin{pmatrix} \hat{u}^n \\ \hat{z}^n \end{pmatrix} \quad \mathbf{P} = \begin{pmatrix} 1 & Ip\,\theta \\ Ip\,\theta & 1 \end{pmatrix} \quad \mathbf{Q} = \begin{pmatrix} 1 & -Ip\,(1-\theta) \\ -Ip\,(1-\theta) & 1 \end{pmatrix}.$$

The amplification matrix of the method is $\mathbf{G} = \mathbf{P}^{-1}\mathbf{Q}$, and a necessary and sufficient condition for stability is $\|\mathbf{G}\|_2 \le 1$ identically for every α. Since $\mathbf{G}$ is a normal matrix the norm of $\mathbf{G}$ is equal to its spectral radius. The eigenvalues of $\mathbf{G}$ are

$$\lambda_{1,2} = \frac{1 - p^2\,\theta\,(1-\theta) \pm I\,p}{1 + p^2\,\theta^2}.$$

Thus, the condition for the spectral radius of $\mathbf{G}$ to be not greater than unity is $1 - 2\theta \leq 0$, or equivalently,

$$\theta \geq \tfrac{1}{2}.$$

The accuracy of finite difference equations (9), (10) is examined by considering the following Taylor series expansions

$$\begin{aligned}
\eta_i^{n+1} &= \eta_i^n + \Delta t \left\{\eta_t\right\}_i^n + \frac{\Delta t^2}{2} \left\{\eta_{tt}\right\}_i^n + O(\Delta t^3), \\
u_{i\pm\frac{1}{2}}^n &= u_i^n \pm \frac{\Delta x}{2} \left\{u_x\right\}_i^n + \frac{\Delta x^2}{8} \left\{u_{xx}\right\}_i^n + O(\Delta x^3), \\
u_{i\pm\frac{1}{2}}^{n+1} &= u_i^n \pm \frac{\Delta x}{2} \left\{u_x\right\}_i^n + \Delta t \left\{u_t\right\}_i^n + \frac{\Delta x^2}{8} \left\{u_{xx}\right\}_i^n \\
&\quad \pm \frac{\Delta t \Delta x}{2} \left\{u_{xt}\right\}_i^n + \frac{\Delta t^2}{2} \left\{u_{tt}\right\}_i^n + O(\Delta x^3, \Delta t^3).
\end{aligned}$$

Substitution of these expressions into the finite difference equation (10), by using the differential equations (7), (8), and after simplifications, yields

$$\left\{\eta_t\right\}_i^n + H\left\{u_x\right\}_i^n = gH\Delta t \left(\theta - \frac{1}{2}\right) \left\{\eta_{xx}\right\}_i^n + O\left(\Delta x^2, \Delta t^2\right), \tag{13}$$

which shows that the finite difference equation (10) is consistent with the differential equation (8) with a second-order accuracy obtained only when $\theta = \frac{1}{2}$. For $\theta \neq \frac{1}{2}$, the first term on the right-hand side of equation (13) represents the damping error and has the form of a diffusion. The stability of the method ($\theta \geq \frac{1}{2}$) corresponds to nonnegative diffusion coefficient $gH\Delta t(\theta - \frac{1}{2})$. An accuracy analysis can be carried out for the finite difference equation (9) in an entirely similar fashion.

To compare the results obtained by the θ-method, consider an initial-boundary problem for equations (7), (8) with $g = H = 1$ and with the following initial and boundary conditions

$$\begin{aligned}
&u(x,0) = 0, \qquad &&\eta(x,0) = \cos(x), \qquad &&-\frac{\pi}{2} \leq x \leq \frac{\pi}{2}, \\
&u\left(-\frac{\pi}{2}, t\right) = 0, \qquad &&u\left(\frac{\pi}{2}, t\right) = 0, \qquad &&t \geq 0.
\end{aligned}$$

The analytical solution of this initial-boundary problem is

$$\begin{aligned}
u(x,t) &= \sin(x)\,\sin(t), \\
\eta(x,t) &= \cos(x)\,\cos(t).
\end{aligned}$$

The numerical results have confirmed that the accuracy is higher when $\theta = \frac{1}{2}$ and the method is stable for every $\theta \geq \frac{1}{2}$. Figure 1 shows the numerical solution obtained for $\eta(0,t)$ when the finite difference scheme (9), (10) is applied with $\Delta t = 0.1$, $\Delta x = \pi/100$ and with $\theta = 1$ and $\theta = \frac{1}{2}$.

4. A THREE-DIMENSIONAL SEMI-IMPLICIT NUMERICAL METHOD

In order to improve the efficiency and the time accuracy of the method proposed by Casulli and Cheng [4], the gradient of surface elevation in the momentum equations (1), (2), and the velocity in the free surface equation (4) will be discretized by the θ-method, and the vertical mixing terms will be discretized implicitly.

As shown in Figure 2, the spatial mesh consists of rectangular boxes of length Δx, width Δy and height Δz_k. Each box is numbered at its center with indices i, j and k. The discrete u velocity is then defined at half integer i and integers j and k; v is defined at integers i, k and half

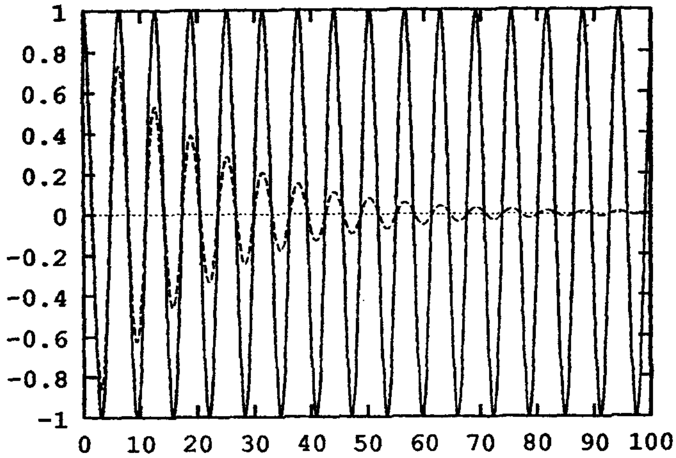

Figure 1. Effect of θ on the damping of a free surface wave: $\theta = 1$ dashed line; $\theta = \frac{1}{2}$ full line.

integer j; w is defined at integers i, j and half integer k. Finally, η is defined at integers i, j. The water depth $h(x, y)$ is specified at the u and v horizontal grid points. Then, parametrized semi-implicit discretization for the momentum equations (1), (2) and for the free surface equation (4) takes the following form

$$u^{n+1}_{i+\frac{1}{2},j,k} = F\, u^{n}_{i+\frac{1}{2},j,k} - g\,\frac{\Delta t}{\Delta x}\left[\theta\left(\eta^{n+1}_{i+1,j} - \eta^{n+1}_{i,j}\right) + (1-\theta)\left(\eta^{n}_{i+1,j} - \eta^{n}_{i,j}\right)\right] + \Delta t\,\frac{\nu_{k+\frac{1}{2}}\,\dfrac{u^{n+1}_{i+\frac{1}{2},j,k+1} - u^{n+1}_{i+\frac{1}{2},j,k}}{\Delta z_{i+\frac{1}{2},j,k+\frac{1}{2}}} - \nu_{k-\frac{1}{2}}\,\dfrac{u^{n+1}_{i+\frac{1}{2},j,k} - u^{n+1}_{i+\frac{1}{2},j,k-1}}{\Delta z_{i+\frac{1}{2},j,k-\frac{1}{2}}}}{\Delta z_{i+\frac{1}{2},j,k}}, \tag{14}$$

$$v^{n+1}_{i,j+\frac{1}{2},k} = F\, v^{n}_{i,j+\frac{1}{2},k} - g\,\frac{\Delta t}{\Delta y}\left[\theta\left(\eta^{n+1}_{i,j+1} - \eta^{n+1}_{i,j}\right) + (1-\theta)\left(\eta^{n}_{i,j+1} - \eta^{n}_{i,j}\right)\right] + \Delta t\,\frac{\nu_{k+\frac{1}{2}}\,\dfrac{v^{n+1}_{i,j+\frac{1}{2},k+1} - v^{n+1}_{i,j+\frac{1}{2},k}}{\Delta z_{i,j+\frac{1}{2},k+\frac{1}{2}}} - \nu_{k-\frac{1}{2}}\,\dfrac{v^{n+1}_{i,j+\frac{1}{2},k} - v^{n+1}_{i,j+\frac{1}{2},k-1}}{\Delta z_{i,j+\frac{1}{2},k-\frac{1}{2}}}}{\Delta z_{i,j+\frac{1}{2},k}}, \tag{15}$$

$$\begin{aligned}
\eta^{n+1}_{i,j} = \eta^{n}_{i,j} &- \frac{\Delta t}{\Delta x}\,\theta\left[\sum_{k=m}^{M} \Delta z_{i+\frac{1}{2},j,k}\, u^{n+1}_{i+\frac{1}{2},j,k} - \sum_{k=m}^{M} \Delta z_{i-\frac{1}{2},j,k}\, u^{n+1}_{i-\frac{1}{2},j,k}\right] \\
&- \frac{\Delta t}{\Delta y}\,\theta\left[\sum_{k=m}^{M} \Delta z_{i,j+\frac{1}{2},k}\, v^{n+1}_{i,j+\frac{1}{2},k} - \sum_{k=m}^{M} \Delta z_{i,j-\frac{1}{2},k}\, v^{n+1}_{i,j-\frac{1}{2},k}\right] \\
&- \frac{\Delta t}{\Delta x}(1-\theta)\left[\sum_{k=m}^{M} \Delta z_{i+\frac{1}{2},j,k}\, u^{n}_{i+\frac{1}{2},j,k} - \sum_{k=m}^{M} \Delta z_{i-\frac{1}{2},j,k}\, u^{n}_{i-\frac{1}{2},j,k}\right] \\
&- \frac{\Delta t}{\Delta y}(1-\theta)\left[\sum_{k=m}^{M} \Delta z_{i,j+\frac{1}{2},k}\, v^{n}_{i,j+\frac{1}{2},k} - \sum_{k=m}^{M} \Delta z_{i,j-\frac{1}{2},k}\, v^{n}_{i,j-\frac{1}{2},k}\right],
\end{aligned} \tag{16}$$

where m and M denote the limit of k-index representing the bottom and the top finite difference stencil, respectively. Moreover, $\Delta z_{i+\frac{1}{2},j,k}$ and $\Delta z_{i,j+\frac{1}{2},k}$ are, in general, the thickness of the k^{th} water layer. If, however, a vertical face of the finite difference box is not fully filled (because either the bottom or the free surface crosses a vertical face of the box), then $\Delta z_{i+\frac{1}{2},j,k}$ and/or $\Delta z_{i,j+\frac{1}{2},k}$ are defined to be the wetted height of the corresponding face. Of course, some of the Δz can be allowed to be zero. The height of the surface layer depends on the position of the free

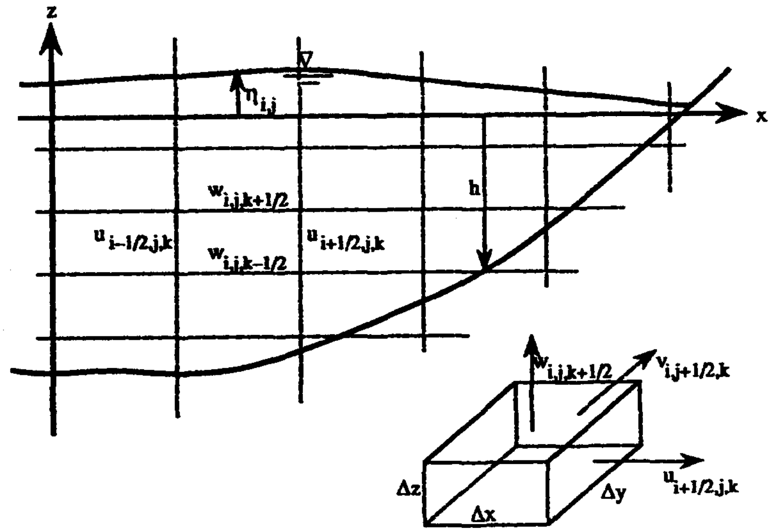

Figure 2. Spatial mesh.

surface, and since the free surface changes with time, thus the surface Δz also depends on the time level n.

In (14) and (15), F is an explicit, nonlinear finite difference operator, which includes the contributions arising from the Eulerian-Lagrangian discretization of the convective and the horizontal eddy viscosity terms (see [4] for details).

The values of u and v above free surface and below bottom in (14) and (15) are eliminated by means of the boundary conditions (5), (6) which are written in difference form as

$$\begin{aligned} \nu_{i+\frac{1}{2},j,M+\frac{1}{2}} \frac{u^{n+1}_{i+\frac{1}{2},j,M+1} - u^{n+1}_{i+\frac{1}{2},j,M}}{\Delta z_{i+\frac{1}{2},j,M+\frac{1}{2}}} &= \tau^w_x, \\ \nu_{i,j+\frac{1}{2},M+\frac{1}{2}} \frac{v^{n+1}_{i,j+\frac{1}{2},M+1} - v^{n+1}_{i,j+\frac{1}{2},M}}{\Delta z_{i,j+\frac{1}{2},M+\frac{1}{2}}} &= \tau^w_y, \end{aligned} \tag{17}$$

and

$$\begin{aligned} \nu_{i+\frac{1}{2},j,m-\frac{1}{2}} \frac{u^{n+1}_{i+\frac{1}{2},j,m} - u^{n+1}_{i+\frac{1}{2},j,m-1}}{\Delta z_{i+\frac{1}{2},j,m-\frac{1}{2}}} &= \gamma^{n+\frac{1}{2}}_{i+\frac{1}{2},j,m}\, u^{n+1}_{i+\frac{1}{2},j,m}, \\ \nu_{i,j+\frac{1}{2},m-\frac{1}{2}} \frac{v^{n+1}_{i,j+\frac{1}{2},m} - v^{n+1}_{i,j+\frac{1}{2},m-1}}{\Delta z_{i,j+\frac{1}{2},m-\frac{1}{2}}} &= \gamma^{n+\frac{1}{2}}_{i,j+\frac{1}{2},m}\, v^{n+1}_{i,j+\frac{1}{2},m}, \end{aligned} \tag{18}$$

where $\gamma^{n+\frac{1}{2}} = \frac{g}{C_z^2}\sqrt{(u^{n+\frac{1}{2}})^2 + (v^{n+\frac{1}{2}})^2}$ and $u^{n+\frac{1}{2}}$ and $v^{n+\frac{1}{2}}$ are taken to be

$$\begin{aligned} u^{n+\frac{1}{2}}_{i+\frac{1}{2},j,m} &= u^n_{i+\frac{1}{2},j,m} - g\frac{\Delta t}{2\Delta x}\left(\eta^n_{i+1,j} - \eta^n_{i,j}\right), \\ v^{n+\frac{1}{2}}_{i,j+\frac{1}{2},m} &= v^n_{i,j+\frac{1}{2},m} - g\frac{\Delta t}{2\Delta y}\left(\eta^n_{i,j+1} - \eta^n_{i,j}\right). \end{aligned}$$

If the computational domain is subdivided into $N_x \times N_y \times N_z$ computational boxes, then equations (14)–(16) constitute a linear system of $N_x N_y (2N_z + 1)$ equations. This system has to be solved at each time step in order to calculate the new field variables $u^{n+1}_{i+\frac{1}{2},j,k}$, $v^{n+1}_{i,j+\frac{1}{2},k}$ and $\eta^{n+1}_{i,j}$ throughout the flow domain.

When $\theta = 1$, the finite difference scheme (14)–(16) reduces to the one introduced by Casulli and Cheng [4].

5. SOLUTION ALGORITHM

Computationally, since a linear system of $N_xN_y(2N_z+1)$ equations can be quite large even for modest values of N_x, N_y and N_z, the system of equations (14)–(16) is first decomposed into a set of $2N_xN_y$ independent tridiagonal systems of N_z equations, and one five-diagonal system of N_xN_y equations. Specifically, equations (14)–(16) are first written in matrix form as

$$\mathbf{A}^n_{i+\frac{1}{2},j}\,\mathbf{U}^{n+1}_{i+\frac{1}{2},j} = \mathbf{G}^n_{i+\frac{1}{2},j} - g\,\frac{\Delta t}{\Delta x}\,\left[\theta\left(\eta^{n+1}_{i+1,j}-\eta^{n+1}_{i,j}\right)\right]\,\mathbf{\Delta Z}^n_{i+\frac{1}{2},j}, \tag{19}$$

$$\mathbf{A}^n_{i,j+\frac{1}{2}}\,\mathbf{V}^{n+1}_{i,j+\frac{1}{2}} = \mathbf{G}^n_{i,j+\frac{1}{2}} - g\,\frac{\Delta t}{\Delta y}\,\left[\theta\left(\eta^{n+1}_{i,j+1}-\eta^{n+1}_{i,j}\right)\right]\,\mathbf{\Delta Z}^n_{i,j+\frac{1}{2}}, \tag{20}$$

$$\begin{aligned}\eta^{n+1}_{i,j} &= \delta^n_{i,j} - \frac{\Delta t}{\Delta x}\,\theta\left[\left(\mathbf{\Delta Z}^n_{i+\frac{1}{2},j}\right)^{\top}\mathbf{U}^{n+1}_{i+\frac{1}{2},j} - \left(\mathbf{\Delta Z}^n_{i-\frac{1}{2},j}\right)^{\top}\mathbf{U}^{n+1}_{i-\frac{1}{2},j}\right]\\ &\quad - \frac{\Delta t}{\Delta y}\,\theta\left[\left(\mathbf{\Delta Z}^n_{i,j+\frac{1}{2}}\right)^{\top}\mathbf{V}^{n+1}_{i,j+\frac{1}{2}} - \left(\mathbf{\Delta Z}^n_{i,j-\frac{1}{2}}\right)^{\top}\mathbf{V}^{n+1}_{i,j-\frac{1}{2}}\right],\end{aligned} \tag{21}$$

where $\mathbf{U}$, $\mathbf{V}$, $\mathbf{\Delta Z}$, $\mathbf{G}$, δ and $\mathbf{A}$ are defined as follows:

$$\mathbf{U}^{n+1}_{i+\frac{1}{2},j} = \begin{pmatrix} u^{n+1}_{i+\frac{1}{2},j,M}\\ u^{n+1}_{i+\frac{1}{2},j,M-1}\\ u^{n+1}_{i+\frac{1}{2},j,M-2}\\ \vdots\\ u^{n+1}_{i+\frac{1}{2},j,m+1}\\ u^{n+1}_{i+\frac{1}{2},j,m}\end{pmatrix},\qquad \mathbf{V}^{n+1}_{i,j+\frac{1}{2}} = \begin{pmatrix} v^{n+1}_{i,j+\frac{1}{2},M}\\ v^{n+1}_{i,j+\frac{1}{2},M-1}\\ v^{n+1}_{i,j+\frac{1}{2},M-2}\\ \vdots\\ v^{n+1}_{i,j+\frac{1}{2},m+1}\\ v^{n+1}_{i,j+\frac{1}{2},m}\end{pmatrix},\qquad \mathbf{\Delta Z} = \begin{pmatrix}\Delta z_M\\ \Delta z_{M-1}\\ \Delta z_{M-2}\\ \vdots\\ \Delta z_{m+1}\\ \Delta z_m\end{pmatrix},$$

$$\mathbf{G}^n_{i+\frac{1}{2},j} = \begin{pmatrix} \Delta z_M\left[F\,u^n_{i+\frac{1}{2},j,M} - g\,\frac{\Delta t}{\Delta x}\,(1-\theta)\,\left(\eta^n_{i+1,j}-\eta^n_{i,j}\right)\right] + \Delta t\,\tau^w_x\\ \Delta z_{M-1}\left[F\,u^n_{i+\frac{1}{2},j,M-1} - g\,\frac{\Delta t}{\Delta x}\,(1-\theta)\,\left(\eta^n_{i+1,j}-\eta^n_{i,j}\right)\right]\\ \Delta z_{M-2}\left[F\,u^n_{i+\frac{1}{2},j,M-2} - g\,\frac{\Delta t}{\Delta x}\,(1-\theta)\,\left(\eta^n_{i+1,j}-\eta^n_{i,j}\right)\right]\\ \vdots\\ \Delta z_{m+1}\left[F\,u^n_{i+\frac{1}{2},j,m+1} - g\,\frac{\Delta t}{\Delta x}\,(1-\theta)\,\left(\eta^n_{i+1,j}-\eta^n_{i,j}\right)\right]\\ \Delta z_m\left[F\,u^n_{i+\frac{1}{2},j,m} - g\,\frac{\Delta t}{\Delta x}\,(1-\theta)\,\left(\eta^n_{i+1,j}-\eta^n_{i,j}v\right)\right]\end{pmatrix},$$

$$\mathbf{G}^n_{i,j+\frac{1}{2}} = \begin{pmatrix} \Delta z_M\left[F\,v^n_{i,j+\frac{1}{2},M} - g\,\frac{\Delta t}{\Delta y}\,(1-\theta)\,\left(\eta^n_{i,j+1}-\eta^n_{i,j}\right)\right] + \Delta t\,\tau^w_y\\ \Delta z_{M-1}\left[F\,v^n_{i,j+\frac{1}{2},M-1} - g\,\frac{\Delta t}{\Delta y}\,(1-\theta)\,\left(\eta^n_{i,j+1}-\eta^n_{i,j}\right)\right]\\ \Delta z_{M-2}\left[F\,v^n_{i,j+\frac{1}{2},M-2} - g\,\frac{\Delta t}{\Delta y}\,(1-\theta)\,\left(\eta^n_{i,j+1}-\eta^n_{i,j}\right)\right]\\ \vdots\\ \Delta z_{m+1}\left[F\,v^n_{i,j+\frac{1}{2},m+1} - g\,\frac{\Delta t}{\Delta y}\,(1-\theta)\,\left(\eta^n_{i,j+1}-\eta^n_{i,j}\right)\right]\\ \Delta z_m\left[F\,v^n_{i,j+\frac{1}{2},m} - g\,\frac{\Delta t}{\Delta y}\,(1-\theta)\,\left(\eta^n_{i,j+1}-\eta^n_{i,j}\right)\right]\end{pmatrix},$$

$$\delta_{i,j}^n = \eta_{i,j}^n - \frac{\Delta t}{\Delta x}(1-\theta)\left[\left(\mathbf{\Delta Z}_{i+\frac{1}{2},j}\right)^\top \mathbf{U}_{i+\frac{1}{2},j}^n - \left(\mathbf{\Delta Z}_{i-\frac{1}{2},j}\right)^\top \mathbf{U}_{i-\frac{1}{2},j}^n\right] - \frac{\Delta t}{\Delta y}(1-\theta)\left[\left(\mathbf{\Delta Z}_{i,j+\frac{1}{2}}\right)^\top \mathbf{V}_{i,j+\frac{1}{2}}^n - \left(\mathbf{\Delta Z}_{i,j-\frac{1}{2}}\right)^\top \mathbf{V}_{i,j-\frac{1}{2}}^n\right],$$

$$\mathbf{A} = \begin{pmatrix} \Delta z_M + a_{M-\frac{1}{2}} & -a_{M-\frac{1}{2}} & & 0 \\ -a_{M-\frac{1}{2}} & \Delta z_{M-1} + a_{M-\frac{1}{2}} + a_{M-\frac{3}{2}} & -a_{M-\frac{3}{2}} & \\ \vdots & \vdots & \vdots & \vdots \\ 0 & & -a_{m+\frac{1}{2}} & \Delta z_m + a_{m+\frac{1}{2}} + \gamma \end{pmatrix}$$

with $a_k = \dfrac{\nu_k \, \Delta t}{\Delta z_k}$.

Then, formal substitution of the expressions for $\mathbf{U}_{i\pm\frac{1}{2},j}^{n+1}$ and $\mathbf{V}_{i,j\pm\frac{1}{2}}^{n+1}$ from (19) and (20) into (21) yields

$$\begin{aligned} \eta_{i,j}^{n+1} &- g\,\frac{\Delta t^2}{\Delta x^2}\,\theta^2 \left[\left[(\mathbf{\Delta Z})^\top \mathbf{A}^{-1}\,\mathbf{\Delta Z}\right]_{i+\frac{1}{2},j}^n \left(\eta_{i+1,j}^{n+1} - \eta_{i,j}^{n+1}\right) - \left[(\mathbf{\Delta Z})^\top \mathbf{A}^{-1}\,\mathbf{\Delta Z}\right]_{i-\frac{1}{2},j}^n \left(\eta_{i,j}^{n+1} - \eta_{i-1,j}^{n+1}\right)\right] \\ &- g\,\frac{\Delta t^2}{\Delta y^2}\,\theta^2 \left[\left[(\mathbf{\Delta Z})^\top \mathbf{A}^{-1}\,\mathbf{\Delta Z}\right]_{i,j+\frac{1}{2}}^n \left(\eta_{i,j+1}^{n+1} - \eta_{i,j}^{n+1}\right) - \left[(\mathbf{\Delta Z})^\top \mathbf{A}^{-1}\,\mathbf{\Delta Z}\right]_{i,j-\frac{1}{2}}^n \left(\eta_{i,j}^{n+1} - \eta_{i,j-1}^{n+1}\right)\right] \\ = \delta_{i,j}^n &- \frac{\Delta t}{\Delta x}\,\theta \left[\left[(\mathbf{\Delta Z})^\top \mathbf{A}^{-1}\,\mathbf{G}\right]_{i+\frac{1}{2},j}^n - \left[(\mathbf{\Delta Z})^\top \mathbf{A}^{-1}\,\mathbf{G}\right]_{i-\frac{1}{2},j}^n\right] \\ &- \frac{\Delta t}{\Delta y}\,\theta \left[\left[(\mathbf{\Delta Z})^\top \mathbf{A}^{-1}\,\mathbf{G}\right]_{i,j+\frac{1}{2}}^n - \left[(\mathbf{\Delta Z})^\top \mathbf{A}^{-1}\,\mathbf{G}\right]_{i,j-\frac{1}{2}}^n\right]. \end{aligned} \tag{22}$$

Since $\mathbf{A}$ is positive definite, $\mathbf{A}^{-1}$ is also positive definite and therefore $(\mathbf{\Delta Z})^\top \mathbf{A}^{-1}\,\mathbf{\Delta Z}$ is a non-negative number. Hence equations (22) constitute a linear five-diagonal system of $N_x N_y$ equations for $\eta_{i,j}^{n+1}$. This system is symmetric and positive definite. Thus, it has a unique solution which can be determined very efficiently by a conjugate gradient method [4,12].

It is important to emphasize that when $\theta = \frac{1}{2}$ is used, the off-diagonal coefficients in system (22) are reduced by a factor $\theta^2 = \frac{1}{4}$. Clearly, this is not the case for the main diagonal term which is given by

$$1 + g\,\frac{\Delta t^2}{\Delta x^2}\,\theta^2 \left[\left[(\mathbf{\Delta Z})^\top \mathbf{A}^{-1}\,\mathbf{\Delta Z}\right]_{i+\frac{1}{2},j}^n + \left[(\mathbf{\Delta Z})^\top \mathbf{A}^{-1}\,\mathbf{\Delta Z}\right]_{i-\frac{1}{2},j}^n\right] + g\,\frac{\Delta t^2}{\Delta y^2}\,\theta^2 \left[\left[(\mathbf{\Delta Z})^\top \mathbf{A}^{-1}\,\mathbf{\Delta Z}\right]_{i,j+\frac{1}{2}}^n + \left[(\mathbf{\Delta Z})^\top \mathbf{A}^{-1}\,\mathbf{\Delta Z}\right]_{i,j-\frac{1}{2}}^n\right].$$

Thus, when $\theta = \frac{1}{2}$ system (22) is better conditioned, and, accordingly, a faster convergence of the conjugate gradient method is achieved.

Once the new free surface location has been determined, equations (19) and (20) constitute a set of simple $2N_x N_y$ linear, tridiagonal systems with N_z equations each. All these systems are independent from each other, symmetric and positive definite. Thus, they can be conveniently solved by a direct method.

Finally, by discretizing the continuity equation (3) the vertical component of the velocity at the new time level can be computed by setting $w^{n+1}_{i,j,m-\frac{1}{2}} = 0$, and

$$w^{n+1}_{i,j,k+\frac{1}{2}} = w^{n+1}_{i,j,k-\frac{1}{2}} - \frac{\Delta z^n_{i+\frac{1}{2},j,k} u^{n+1}_{i+\frac{1}{2},j,k} - \Delta z^n_{i-\frac{1}{2},j,k} u^{n+1}_{i-\frac{1}{2},j,k}}{\Delta x} - \frac{\Delta z^n_{i,j+\frac{1}{2},k} v^{n+1}_{i,j+\frac{1}{2},k} - \Delta z^n_{i,j-\frac{1}{2},j,k} v^{n+1}_{i,j-\frac{1}{2},k}}{\Delta y}, \quad k = m, m+1, \dots, M. \tag{23}$$

The numerical algorithm presented above includes the simulation of flooding and drying of tidal flats. To this purpose at each time step the new water depths $H^{n+1}_{i+\frac{1}{2},j}$ and $H^{n+1}_{i,j+\frac{1}{2}}$ are defined as

$$H^{n+1}_{i+\frac{1}{2},j} = \max\left(0,\, h^{n+1}_{i+\frac{1}{2},j} + \eta^{n+1}_{i,j},\, h^{n+1}_{i+\frac{1}{2},j} + \eta^{n+1}_{i+1,j}\right),$$
$$H^{n+1}_{i,j+\frac{1}{2}} = \max\left(0,\, h^{n+1}_{i,j+\frac{1}{2}} + \eta^{n+1}_{i,j},\, h^{n+1}_{i,j+\frac{1}{2}} + \eta^{n+1}_{i,j+1}\right),$$

with the understanding that an occurrence of the zero value for the total depth H simply means a dry point which may be flooded at a later time when H becomes positive. The vertical grid spacings $\mathbf{\Delta Z}^n_{i+\frac{1}{2},j}$ and $\mathbf{\Delta Z}^n_{i,j+\frac{1}{2}}$ are updated accordingly.

Note, finally, that when the vertical spacing Δz is taken to be large enough so that both the bottom and the free surface always fall within one vertical layer, this algorithm reduces to a two-dimensional semi-implicit numerical method, which is consistent with the two-dimensional, vertically integrated shallow water equations and which, for $\theta = 1$, yields the method described by Casulli [1].

6. NUMERICAL STABILITY

The stability analysis of the semi-implicit finite difference method (14)–(16) will be carried out by using the von Neumann method under the assumptions that the governing differential equations (1)–(4) are linear, with constant coefficients and defined on an infinite horizontal domain, or with periodic boundary conditions on a finite domain. Thus, if Δz denotes the constant layer thickness, by neglecting the wind stresses ($\tau^w_x = 0, \tau^w_y = 0$), assuming that $a_1 = a_2 = \cdots = a_{N_z}$ and γ in the matrix $\mathbf{A}$ are constants, the difference equations (19)–(21) reduce to

$$\mathbf{A}\,\mathbf{U}^{n+1}_{i+\frac{1}{2},j} + g\frac{\Delta t}{\Delta x}\theta\left(\eta^{n+1}_{i+1,j} - \eta^{n+1}_{i,j}\right)\mathbf{\Delta Z} = \Delta z\, F\,\mathbf{U}^n_{i+\frac{1}{2},j} - (1-\theta)\, g\frac{\Delta t}{\Delta x}\left(\eta^n_{i+1,j} - \eta^n_{i,j}\right)\mathbf{\Delta Z}, \tag{24}$$

$$\mathbf{A}\mathbf{V}^{n+1}_{i,j+\frac{1}{2}} + g\frac{\Delta t}{\Delta y}\theta\left(\eta^{n+1}_{i,j+1} - \eta^{n+1}_{i,j}\right)\mathbf{\Delta Z} = \Delta z\, F\,\mathbf{V}^n_{i,j+\frac{1}{2}} - (1-\theta)\, g\frac{\Delta t}{\Delta y}\left(\eta^n_{i,j+1} - \eta^n_{i,j}\right)\mathbf{\Delta Z}, \tag{25}$$

$$\begin{aligned}\eta^{n+1}_{i,j} = \eta^n_{i,j} &- \frac{\Delta t}{\Delta x}\theta\left[(\mathbf{\Delta Z})^\top \mathbf{U}^{n+1}_{i+\frac{1}{2},j} - (\mathbf{\Delta Z})^\top \mathbf{U}^{n+1}_{i-\frac{1}{2},j}\right]\\ &- \frac{\Delta t}{\Delta y}\theta\left[(\mathbf{\Delta Z})^\top \mathbf{V}^{n+1}_{i,j+\frac{1}{2}} - (\mathbf{\Delta Z})^\top \mathbf{V}^{n+1}_{i,j-\frac{1}{2}}\right]\\ &- \frac{\Delta t}{\Delta x}(1-\theta)\left[(\mathbf{\Delta Z})^\top \mathbf{U}^{n}_{i+\frac{1}{2},j} - (\mathbf{\Delta Z})^\top \mathbf{U}^{n}_{i-\frac{1}{2},j}\right]\\ &- \frac{\Delta t}{\Delta y}(1-\theta)\left[(\mathbf{\Delta Z})^\top \mathbf{V}^{n}_{i,j+\frac{1}{2}} - (\mathbf{\Delta Z})^\top \mathbf{V}^{n}_{i,j-\frac{1}{2}}\right],\end{aligned} \tag{26}$$

where $F\,\mathbf{U}^n_{i+\frac{1}{2},j}$ and $F\,\mathbf{V}^n_{i+\frac{1}{2},j}$ are the explicit finite difference discretization of the horizontal eddy viscosity terms. Specifically,

$$\begin{aligned}F\,\mathbf{U}^n_{i+\frac{1}{2},j} = \mathbf{U}^n_{i+\frac{1}{2},j} &+ \mu\frac{\Delta t}{\Delta x^2}\left(\mathbf{U}^n_{i+\frac{3}{2},j} - 2\,\mathbf{U}^n_{i+\frac{1}{2},j} + \mathbf{U}^n_{i-\frac{1}{2},j}\right)\\ &+ \mu\frac{\Delta t}{\Delta y^2}\left(\mathbf{U}^n_{i+\frac{1}{2},j+1} - 2\,\mathbf{U}^n_{i+\frac{1}{2},j} + \mathbf{U}^n_{i+\frac{1}{2},j-1}\right),\end{aligned} \tag{27}$$

$$F\,\mathbf{V}^n_{i,j+\frac{1}{2}} = \mathbf{V}^n_{i,j+\frac{1}{2}} + \mu\,\frac{\Delta t}{\Delta x^2}\left(\mathbf{V}^n_{i+1,j+\frac{1}{2}} - 2\,\mathbf{V}^n_{i,j+\frac{1}{2}} + \mathbf{V}^n_{i-1,j+\frac{1}{2}}\right) + \mu\,\frac{\Delta t}{\Delta y^2}\left(\mathbf{V}^n_{i,j+\frac{3}{2}} - 2\,\mathbf{V}^n_{i,j+\frac{1}{2}} + \mathbf{V}^n_{i,j-\frac{1}{2}}\right), \tag{28}$$

while the nonlinear convective terms are not being considered.

THEOREM 1. *The semi-implicit finite difference scheme (24)–(26) is stable in the von Neumann sense if* $\frac{1}{2} \le \theta \le 1$ *and if the time step* Δt *satisfies the following inequality*

$$\Delta t \le \left[2\mu\left(\frac{1}{\Delta x^2} + \frac{1}{\Delta y^2}\right)\right]^{-1}. \tag{29}$$

PROOF. By replacing $\mathbf{U}^n_{i+\frac{1}{2},j}$, $\mathbf{V}^n_{i,j+\frac{1}{2}}$ and $\eta^n_{i,j}$ in (24)–(28) with the corresponding Fourier components $\hat{\mathbf{U}}^n e^{I\,[(i+\frac{1}{2})\alpha+j\beta]}$, $\hat{\mathbf{V}}^n e^{I\,[i\alpha+(j+\frac{1}{2})\beta]}$ and $\hat{\eta}^n\,e^{I\,(i\alpha+j\beta)}$, after some simplifications, (24)–(26) become

$$\mathbf{A}\,\hat{\mathbf{U}}^{n+1} + Ip\,\theta g\,\hat{\eta}^{n+1}\,\mathbf{\Delta Z} = \Delta z\,f\,\hat{\mathbf{U}}^n - Ip\,(1-\theta)g\,\hat{\eta}^n\,\mathbf{\Delta Z}, \tag{30}$$

$$\mathbf{A}\,\hat{\mathbf{V}}^{n+1} + Iq\,\theta g\,\hat{\eta}^{n+1}\,\mathbf{\Delta Z} = \Delta z\,f\,\hat{\mathbf{V}}^n - Iq\,(1-\theta)g\,\hat{\eta}^n\,\mathbf{\Delta Z}, \tag{31}$$

$$\hat{\eta}^{n+1} + Ip\,\theta\,(\mathbf{\Delta Z})^{\top}\,\hat{\mathbf{U}}^{n+1} + Iq\,\theta\,(\mathbf{\Delta Z})^{\top}\,\hat{\mathbf{V}}^{n+1} = \hat{\eta}^n - Ip(1-\theta)\,(\mathbf{\Delta Z})^{\top}\,\hat{\mathbf{U}}^n - Iq\,(1-\theta)\,(\mathbf{\Delta Z})^{\top}\,\hat{\mathbf{V}}^n, \tag{32}$$

where $\hat{\mathbf{U}}^n$, $\hat{\mathbf{V}}^n$ and $\hat{\eta}^n$ are the amplitude functions of $\mathbf{U}$, $\mathbf{V}$ and η at time level t^n; α and β are the x and the y phase angles; $p = 2(\Delta t/\Delta x)\sin(\alpha/2)$; $q = 2(\Delta t/\Delta y)\sin(\beta/2)$; and the amplification factor of the explicit difference operator F is given by

$$f = 1 - 2\mu\,\Delta t\left[\frac{1-\cos(\alpha)}{\Delta x^2} + \frac{1-\cos(\beta)}{\Delta y^2}\right]. \tag{33}$$

Equations (30)–(32) can be written in a more compact matrix form as

$$\mathbf{P}\,\hat{\mathbf{W}}^{n+1} = \mathbf{Q}\,\hat{\mathbf{W}}^n, \tag{34}$$

where

$$\hat{\mathbf{W}}^k = \begin{bmatrix} \hat{\mathbf{U}}^n \\ \hat{\mathbf{V}}^n \\ \hat{\eta}^n \end{bmatrix}, \qquad \mathbf{P} = \begin{bmatrix} \mathbf{A} & 0 & Ip\,\theta g\,\mathbf{\Delta Z} \\ 0 & \mathbf{A} & Iq\,\theta g\,\mathbf{\Delta Z} \\ Ip\,\theta(\mathbf{\Delta Z})^{\top} & Iq\,\theta(\mathbf{\Delta Z})^{\top} & 1 \end{bmatrix},$$

and

$$\mathbf{Q} = \begin{bmatrix} \Delta z\,f\,\mathbf{I}_d & 0 & -Ip\,(1-\theta)g\,\mathbf{\Delta Z} \\ 0 & \Delta z\,f\mathbf{I}_d & -Iq\,(1-\theta)g\,\mathbf{\Delta Z} \\ -Ip\,(1-\theta)\,(\mathbf{\Delta Z})^{\top} & -Iq\,(1-\theta)(\mathbf{\Delta Z})^{\top} & 1 \end{bmatrix},$$

with $\mathbf{I}_d$ being the identity matrix of order N_z. Thus, the amplification matrix of the method is $\mathbf{G} = \mathbf{P}^{-1}\,\mathbf{Q}$ and a condition for stability is that the spectral radius of $\mathbf{G}$ does not exceed 1 identically for every α and β. Equivalently, the modulus of each eigenvalue of $\mathbf{G}^{-1}$ must be no less than 1. The characteristic polynomial of matrix $\mathbf{G}^{-1}$ is $\det(\mathbf{P} - \lambda\,\mathbf{Q}) = 0$, that is,

$$\det\begin{bmatrix} \mathbf{A} - \lambda\,\Delta z\,f\,\mathbf{I}_d & 0 & Ip\,gs\,\mathbf{\Delta Z} \\ 0 & \mathbf{A} - \lambda\,\Delta z\,f\,\mathbf{I}_d & Iq\,gs\,\mathbf{\Delta Z} \\ Ip\,s(\mathbf{\Delta Z})^{\top} & Iq\,s(\mathbf{\Delta Z})^{\top} & 1-\lambda \end{bmatrix} = 0, \tag{35}$$

where $s = \theta + \lambda\,(1-\theta)$. Next to be shown is that equation (35) cannot be satisfied by any complex number λ when $|\lambda| < 1$. Assume that $|\lambda| < 1$ and consider the matrix $\mathbf{A}$ which is real, symmetric and strictly diagonally dominant. Note that when inequality (29) is satisfied, equation (33) implies $|f| \le 1$. Therefore, $|\lambda\,\Delta z\, f| < \Delta z$, and hence, the matrix $\mathbf{A} - \lambda\,\Delta z\, f\,\mathbf{I}_d$ remains strictly diagonally dominant and invertible. Consider then,

$$\mathbf{T} = \begin{bmatrix} \mathbf{I}_d & 0 & 0 \\ 0 & \mathbf{I}_d & 0 \\ -Ip\, s(\mathbf{\Delta Z})^\top[\mathbf{A} - \lambda\,\Delta z\, f\,\mathbf{I}_d]^{-1} & -Iq\, s(\mathbf{\Delta Z})^\top[\mathbf{A} - \lambda\,\Delta z\, f\,\mathbf{I}_d]^{-1} & 1 \end{bmatrix},$$

so that

$$\det(\mathbf{P} - \lambda\,\mathbf{Q}) = \det(\mathbf{T}) \times \det(\mathbf{P} - \lambda\,\mathbf{Q}) = \det[\mathbf{T}(\mathbf{P} - \lambda\,\mathbf{Q})].$$

Thus, equation (35) can also be written as

$$\det \begin{bmatrix} \mathbf{A} - \lambda\,\Delta z\, f\,\mathbf{I}_d & 0 & Ip\,gs\,\mathbf{\Delta Z} \\ 0 & \mathbf{A} - \lambda\,\Delta z\, f\,\mathbf{I}_d & Iq\,gs\,\mathbf{\Delta Z} \\ 0 & 0 & b \end{bmatrix} = 0,$$

where

$$b = g\,(p^2 + q^2)\, s^2(\mathbf{\Delta Z})^\top[\mathbf{A} - \lambda\,\Delta z\, f\,\mathbf{I}_d]^{-1}\mathbf{\Delta Z} + 1 - \lambda.$$

Since $\det(\mathbf{A} - \lambda\,\Delta z\, f\,\mathbf{I}_d) \neq 0$ it is only necessary to show that for $\frac{1}{2} \le \theta \le 1$ one has $b \neq 0$. On the other hand, since $\mathbf{A}$ is real and symmetric, it possesses N_z real eigenvalues $\lambda_k \ge \Delta z$ and a complete system of eigenvectors $\mathbf{x}_k$, which form an orthonormal basis. Consequently, the vector $[\mathbf{A} - \lambda\,\Delta z\, f\,\mathbf{I}_d]^{-1}\mathbf{\Delta Z}$ can be expressed as

$$[\mathbf{A} - \lambda\,\Delta z\, f\,\mathbf{I}_d]^{-1}\mathbf{\Delta Z} = \sum_{k=1}^{N_z} \frac{(\mathbf{\Delta Z})^\top\mathbf{x}_k}{\lambda_k - \lambda\,\Delta z\, f}\,\mathbf{x}_k.$$

Thus, the inequality $b \neq 0$ can be written as

$$g\,(p^2 + q^2)\, s^2 \sum_{k=1}^{N_z} \frac{[(\mathbf{\Delta Z})^\top\mathbf{x}_k]^2}{|\lambda_k - \lambda\,\Delta z\, f|^2}\,(\lambda_k - \bar\lambda\,\Delta z\, f) + 1 - \lambda \neq 0. \tag{36}$$

Since $s \neq 0$, inequality (36) can be divided by s, to obtain

$$g\,(p^2 + q^2)\, s \sum_{k=1}^{N_z} \frac{[(\mathbf{\Delta Z})^\top\mathbf{x}_k]^2}{|\lambda_k - \lambda\,\Delta z\, f|^2}\,(\lambda_k - \bar\lambda\,\Delta z\, f) + \frac{(1-\lambda)\,\bar s}{|s|^2} \neq 0. \tag{37}$$

In this latter inequality one has $|s|^2 > 0$ and

$$g\,(p^2 + q^2)\,\frac{[(\mathbf{\Delta Z})^\top\mathbf{x}_k]^2}{|\lambda_k - \lambda\,\Delta z\, f|^2} \ge 0.$$

Moreover,

$$\mathrm{Re}\left[s(\lambda_k - \bar\lambda\,\Delta z\, f)\right] > 0 \tag{38}$$

and

$$\mathrm{Re}\left[(1-\lambda)\,\bar s\,\right] > 0. \tag{39}$$

In fact,

$$\begin{aligned}
&\mathrm{Re}[s\,(\lambda_k - \bar\lambda\,\Delta z\, f)] \\
&= \mathrm{Re}\left\{[\theta + \lambda\,(1-\theta)]\,(\lambda_k - \bar\lambda\,\Delta z\, f)\right\} \\
&= \theta\,[\lambda_k - f\,\Delta z\ \mathrm{Re}(\lambda)] + (1-\theta)[\lambda_k\ \mathrm{Re}(\lambda) - |\lambda|^2\,\Delta z\, f] \\
&= (1-\theta)[\lambda_k + \mathrm{Re}(\lambda)(\lambda_k - \Delta z\, f) - |\lambda|^2 f\,\Delta z] + (2\theta - 1)[\lambda_k - f\,\Delta z\ \mathrm{Re}(\lambda)] \\
&\ge (1-\theta)[\lambda_k - |\lambda|\,(\lambda_k - \Delta z\, f) - |\lambda|^2 f\,\Delta z] + (2\theta - 1)[\lambda_k - |f\,\Delta z\,\lambda|\,] \\
&= (1-\theta)[(\lambda_k + f\,\Delta z\,|\lambda|\,)\,(1 - |\lambda|\,)] + (2\theta - 1)[\lambda_k - |f\,\Delta z\,\lambda|\,] \\
&\ge (\lambda_k - |f\,\Delta z\,\lambda|\,)[(1-\theta)\,(1 - |\lambda|\,) + (2\theta - 1)] > 0.
\end{aligned}$$

The validity of (39) can be shown in a similar fashion with $\bar{\lambda}$ replacing λ and $\lambda_k = f = \Delta z = 1$. This proves that the left-hand side of (37) has strictly positive real part. Thus, when $\frac{1}{2} \leq \theta \leq 1$ under the stability restriction (29) the matrix $\mathbf{G}^{-1}$ does not have eigenvalues λ such that $|\lambda| < 1$. Therefore, the spectral radius of $\mathbf{G}$ is no greater than 1 and the scheme (24)–(26) is stable in the von Neumann sense.

Note that the stability of the semi-implicit finite difference scheme (24)–(26) is independent of the celerity, bottom friction and vertical viscosity. It does depend on the horizontal viscosity through the mild stability condition (29). This method becomes unconditionally stable when the horizontal viscosity terms are neglected. The presence of non-linear convective terms may affect the stability of the method when they are discretized explicitly by standard schemes which use, for example, central, upwind differences, or Eulerian-Lagrangian methods (ELMs). Use of ELMs as described in [4] is always recommended because of their higher accuracy and because additional conditions for the stability are not required.

7. COMPUTER APPLICATIONS

The present model, in its original formulation with $\theta = 1$, has been applied extensively at several sites including San Francisco Bay, California, and the Lagoon of Venice, Italy [2,4]. This section is aimed at emphasizing the superior efficiency and accuracy obtained by this method in the simulation of complex three-dimensional flows when $\theta = \frac{1}{2}$ is used.

The Lagoon of Venice is a very complex sea water basin whose area is about $50\,\text{km}^2$ and which consists of several interconnected narrow channels with a maximum width of 1 km and depth of 50 m encircling large and flat shallow areas. Additionally, several tidal marshes with a bathymetry of only 20–40 cm above sea level require proper treatment of flooding and drying. The Lagoon is connected to the Adriatic Sea through three narrow inlets, namely Lido, Malamocco, and Chioggia (see Figure 3). The city of Venice is situated upon the largest island near Lido inlet. Tides propagate from the Adriatic Sea into the Lagoon through the three inlets. In the numerical model the Lagoon has been covered with a $N_x = 384$ by $N_y = 426$ finite difference mesh of equal $\Delta x = \Delta y = 100\,\text{m}$. At the three inlets, an M_2 tide of 0.5 m amplitude and 12 lunar hour period has been specified.

The integration time step is chosen to be $\Delta t = 15\,\text{min}$ and the computations have been carried out by solving, at each time step, a corresponding linear, five-diagonal system of $N_x N_y = 163,584$ equations.

The numerical simulations have been performed with both $\theta = 1$ and $\theta = \frac{1}{2}$. With $\theta = \frac{1}{2}$ the matrix of coefficients of the five-diagonal system is better conditioned because the off-diagonal terms have been reduced by a factor $\frac{1}{4}$. Thus, in general, the preconditioned conjugate gradient method requires a lower number of iterations to obtain the desired solution with a given accuracy. For example, by fixing an accuracy of 10^{-3}, a simulation of one tidal cycle using the model with one vertical layer and $\theta = 1$, requires an average of 100 iterations at each time step. When $\theta = \frac{1}{2}$, the averaged number of iterations drops down to 65. Accordingly, the total CPU time required by a CRAY Y-MP8/432 to simulate one tidal cycle reduces from 60 s, when $\theta = 1$, to 49 s when $\theta = \frac{1}{2}$. When the model is considered for three-dimensional applications with more vertical layers, the number of iterations and the corresponding computing time required to solve the large five-diagonal system remains essentially the same as for the one layer case. The total computing time, however, is increased by the solution of the corresponding tridiagonal systems. For example, using four vertical layers with $\theta = 1$ and with $\theta = \frac{1}{2}$, a total computing time of 92 s and 82 s, respectively, is required for every tidal cycle. Of course, when higher accuracy is required, the number of iterations increases and the corresponding saving in computing time becomes more evident.

The computed results indicate that the amplitude of the free surface waves in various points of the Lagoon is generally 5% higher when $\theta = \frac{1}{2}$ is used. Figure 4 shows the values of η calculated

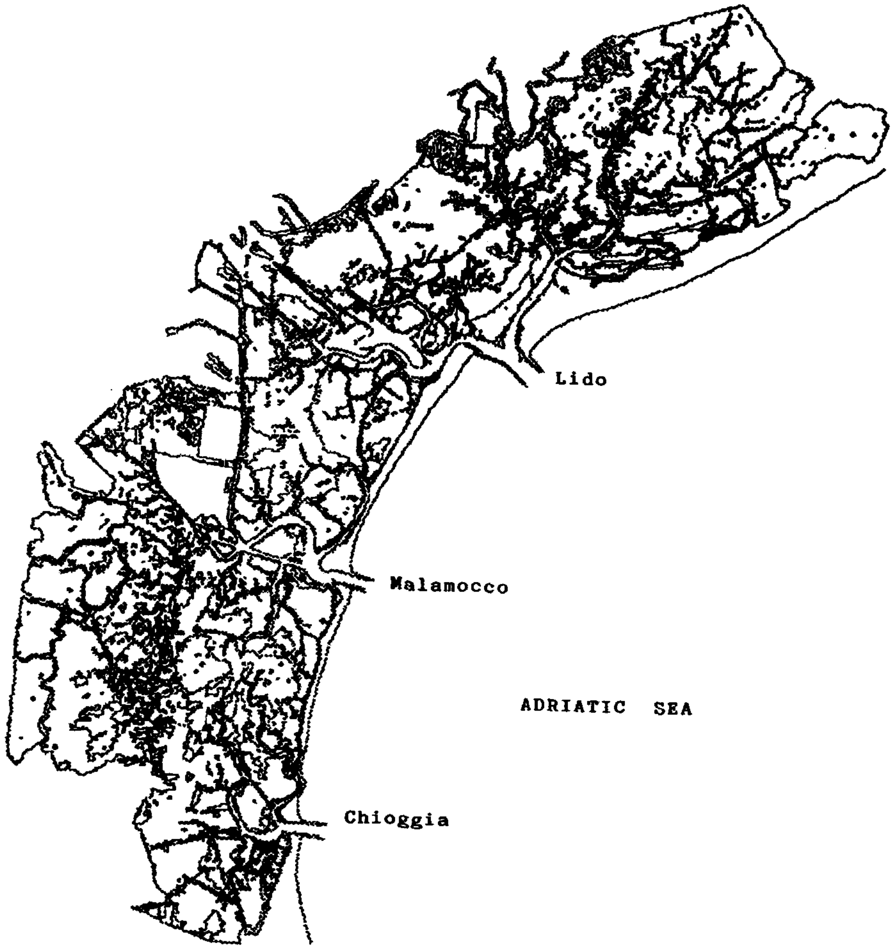

Figure 3. The Lagoon of Venice.

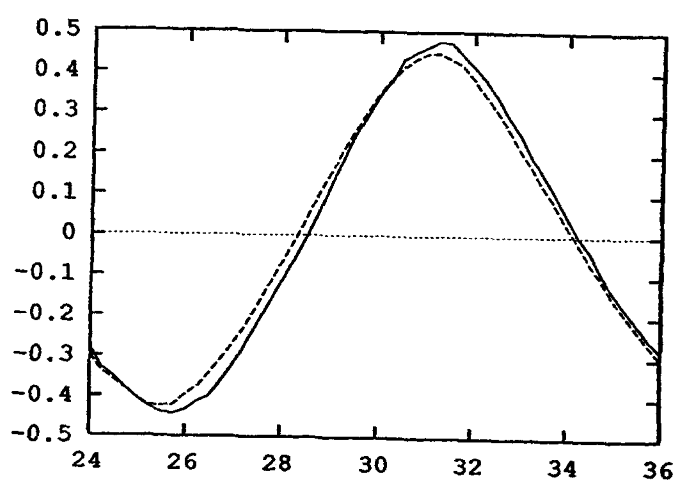

Figure 4. Effect of θ on the free surface in the Lagoon of Venice: $\theta = 1$ dashed line; $\theta = \frac{1}{2}$ full line.

at a typical point located west of the city of Venice. When $\theta = \frac{1}{2}$, the values of η range between $-44\,\text{cm}$ and $48\,\text{cm}$; while, for $\theta = 1$, the values of η range between $-42\,\text{cm}$ and $45\,\text{cm}$. As expected, when $\theta = \frac{1}{2}$, the reduced damping error results in a higher wave amplitude.

8. CONCLUSIONS

A parametrized semi-implicit finite difference model for the three-dimensional shallow water flow has been analyzed from the point of view of its numerical stability, accuracy and efficiency. A rigorous stability and accuracy analysis has been presented. It is shown that when the implicitness parameter θ is set to be $\frac{1}{2}$ the method is stable and achieves highest accuracy and efficiency. In the particular case that only one vertical layer is specified, this scheme remains consistent with the vertically integrated two-dimensional shallow water equations. Computationally, the resulting algorithm is suitable for the simulations of complex three-dimensional flows using fine spatial resolution and relatively large time steps. The present formulation is fully vectorizable and naturally allows for the simulation of flooding and drying of tidal flats.

REFERENCES

1. V. Casulli, Semi-implicit finite difference methods for the two-dimensional shallow-water equation, *Jour. of Computational Physics* **86**, 56–74 (1990).
2. R.T. Cheng, V. Casulli and J.W. Gartner, Tidal, Residual, Inter-tidal Mud-flat (TRIM) Model with applications to San Francisco Bay, *Estuarine, Coastal Shelf Science* **36**, 235–280 (1993).
3. R.P. Signell and B. Butman, Modeling tidal exchange and dispersion in Boston Harbor, *Jour. of Geophysical Research* **97** (C10), 15591–15606 (1992).
4. V. Casulli and R.T. Cheng, Semi-implicit finite difference methods for three-dimensional shallow-water flow, *Int. Jour. for Numerical Methods in Fluids* **15**, 629–648 (1992).
5. V. Casulli and D. Greenspan, Pressure method for the numerical solution of transient, compressible fluid flow, *Int. Jour. for Numerical Methods in Fluids* **4**, 1001–1012 (1984).
6. G. Patnaik, R.H. Guirguis, J.P. Boris and E.S. Oran, A barely implicit correction for flux-corrected transport, *Jour. of Computational Physics* **71**, 1–20 (1987).
7. A.F. Blumberg and G.L. Mellor, A description of a three-dimensional coastal ocean circulation model, In *Three-Dimensional Coastal Ocean Circulation Models, Coastal and Estuarine Sciences*, (Edited by N.S. Heaps), Vol. 4, 1–16, AGU, Washington, DC, (1987).
8. A.M. Davies and J.N. Aldridge, A stable algorithm for bed friction in three-dimensional shallow sea modal models, *Int. Jour. for Numerical Methods in Fluids* **14**, 477–493 (1992).
9. E.D. de Goede, Numerical methods for the three-dimensional shallow water equations on Supercomputers, Thesis, Centrum voor Wiskunde en Informatica, Amsterdam (1992).
10. K.C. Duwe, R.R. Hewer and J.O. Backhaus, Results of a semi-implicit two-step method for the simulation of markedly nonlinear flow in coastal seas, *Continental Shelf Research* **2** (4), 255–274 (1983).
11. J.J. Leendertse, A new approach to three-dimensional free-surface flow modelling, Rep. R-3712-NETH/RC, Rand Corporation, Santa Monica, California (1989).
12. E. Bertolazzi, Metodo PCG ed applicazione ad un modello di acque basse, Tesi, Dipartimento di Matematica, Università di Trento (1990).