Available online at www.sciencedirect.com
ScienceDirect
Journal of Computational Physics 227 (2008) 4425–4440
JOURNAL OF COMPUTATIONAL PHYSICS
www.elsevier.com/locate/jcp

Preconditioning for dual-time-stepping simulations of the shallow water equations including Coriolis and bed friction effects

B.T. Helenbrook [a,*], G.W. Cowles [b]

[a] *Mechanical and Aeronautical Engineering, Clarkson University, Potsdam, NY 13699-5725, USA*
[b] *Department of Fisheries Oceanography, School for Marine Sciences and Technology, University of Massachusetts-Dartmouth, Dartmouth, MA 02747, USA*

Received 5 September 2007; received in revised form 18 December 2007; accepted 4 January 2008
Available online 16 January 2008

Abstract

Diagonal preconditioners for implicit-unsteady and steady discretizations of the shallow water equations at low- and high-Rossby and drag-number limits are analyzed. For each case, sub and supercritical flow conditions are also considered. Based on the analysis, a preconditioner is derived for use with a multigrid cycle that performs as well as possible under all conditions. In addition, a streamwise-upwind Petrov–Galerkin discretization of the system is presented that is derived from the preconditioned system. Using this discretization, it is demonstrated that for most conditions, the preconditioner gives rapid convergence that is independent of the grid resolution and the flow parameters. Practical tests including an equatorial Rossby soliton and tide propagation over variable bathymetry are simulated to demonstrate the performance of this approach.

Keywords: Shallow water equations; Preconditioning; Dual-time stepping; Multigrid

1. Introduction

The shallow water equations are the mathematical model for a wide range of fluid flow problems. For hydraulic studies, they can be used to assess the extent of flooding from failed dams or swollen rivers. For coastal ocean applications, the shallow water equations can be employed to model seiches, tidal flows, storm surge, pollutant dispersal, and accurate residence times in estuarine pollution problems. In geophysical scale problems where the Earth's rotation becomes important, they can be used to model the barotropic motion of Kelvin and Rossby waves in the ocean and atmosphere.

* Corresponding author. Tel.: +1 315 268 2204; fax: +1 315 268 6695.
E-mail address: helenbrk@clarkson.edu (B.T. Helenbrook).

0021-9991/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jcp.2008.01.004

Difficulties in integrating the shallow water equations primarily stem from the disparity between convective velocities which typically are of the order 0.1–1 m/s and the gravity wave speed which can exceed 200 m/s off the continental shelf. The ratio of these values forms the relevant Froude number ($Fr = u/\sqrt{gh}$), where u is the convective speed, g the gravitational acceleration, and h the depth of the water column. Values for this parameter can be greater than unity in dam break problems but are typically much smaller, reaching 10^{-5} in deep ocean flows where the associated problem stiffness hinders explicit time-marching techniques.

To filter the gravity waves and thus relax the time step constraint, early work in large-scale geophysical applications used a rigid lid approximation [1]. This is a reasonable approach for meso- and basin-scale oceanic problems and atmospheric-scale problems where relatively little energy is contained in the gravity waves and tidal motions may be ignored. The rigid lid approach was later shown to have a detrimental effect on the energy transport by altering the group velocity of Rossby waves [2] and has subsequently seen reduced usage.

Presently, a common approach is to use a semi-implicit scheme where the linear terms responsible for the gravity wave's existence are treated implicitly [3,4]. These schemes allow a larger time step to be employed as the stability no longer depends on the celerity of the gravity wave. They are commonly formulated using Crank–Nicholson, backwards Euler, or θ-method approaches and many implementations neglect an important correction term, reducing them to first-order accuracy [5]. When Eulerian methods are used, these semi-implicit schemes are limited by a Courant–Friedrichs–Lewy (CFL) condition based on the local advection speed. Traditionally, this has not been a severe restriction, but recent advances in computer power in combination with unstructured grid methods in coastal ocean modeling have made it possible to use a fine resolution ($\sim$10 m) near the coast. These small cells necessitate small time steps in semi-implicit based schemes and can require O(10,000) steps to integrate a single period of the semi-diurnal M_2 tide. Recent efforts have focused on relaxing this CFL requirement using predictor–corrector algorithms [6] as well as semi-Lagrangian schemes [3,7] that discretize substantive time derivatives along characteristics.

In applications of the shallow water equations to hydraulic flow problems such as dam failures, the relatively shallow depths and large fluid velocities can lead to supercritical ($Fr > 1$) flows. Most of the scientific effort associated with these problems focuses on resolving moving flow discontinuities and simulating wetting/drying processes associated with flooding events. For these applications, explicit schemes are typically used because the higher Froude numbers and shorter problem time scales render them feasible (e.g. [8,9]). However in very shallow water, the stress exerted by the bed on the water column can lead to a significant reduction in allowable time step. This has led to the implementation of point-implicit approaches for the bottom friction [10].

By employing fully-implicit schemes, the CFL requirement can be completely relaxed. The system is unconditionally stable and the time step selection is based solely on desired solution accuracy. For coupled nonlinear hyperbolic systems of equations, fully implicit schemes require iterative solutions to advance the flow solutions each step. This problem is analogous to implicit numerical simulations of compressible flow and techniques derived for such simulations can often be applied to the shallow water equations as well. In this work, we develop a "dual time-stepping" approach [11–14] for solving the shallow water equations. This method is fully implicit, matrix free, and allows any order of temporal accuracy, but it requires the iterative solution of a set of nonlinear equations at each time step. To be cost effective, convergence must be rapid. For this work, we use an explicit iteration and local time stepping within a multigrid cycle [15–17]. We formulate a preconditioner for this iteration that allows fast solutions to be obtained over a wide range of flow regimes. These include large-scale geophysical flows in near geostrophic balance, tidally-driven coastal flows at low Froude, and friction-dominated flows in the shallows of estuarine intertidal zones.

2. Governing equations

Our starting point is the shallow water equations in conservative form

$$\frac{\partial w}{\partial t} + \frac{\partial f_x}{\partial x} + \frac{\partial f_y}{\partial y} = s \tag{1}$$

with the flow vector

$$w = \begin{bmatrix} h \\ hu \\ hv \end{bmatrix} \tag{2}$$

where h is the fluid depth, and u and v are the x- and y-direction velocities respectively. The x- and y-direction flux vectors, f_x and f_y, are given by

$$f_x = \begin{bmatrix} hu \\ hu^2 + g(h^2 - b^2)/2 \\ hvu \end{bmatrix} \tag{3}$$

$$f_y = \begin{bmatrix} hv \\ huv \\ hv^2 + g(h^2 - b^2)/2 \end{bmatrix} \tag{4}$$

where b is a variable bathymetry. s is the source term

$$s = \begin{bmatrix} 0 \\ fhv + \tau'_{bx} + g(h-b)b_x \\ -fhu + \tau'_{by} + g(h-b)b_y \end{bmatrix} \tag{5}$$

which includes the effect of the Earth's rotation through the Coriolis parameter, f, the stress exerted by bottom friction through τ'_{bx} and τ'_{by}, and forcing from the bathymetry gradient (b_x, b_y). The gravity terms are written in the above form such that the solution $h = b$ is guaranteed not to drive any spurious currents. The Coriolis parameter, f, is given by $f = 2\Omega\sin(\phi)$ where Ω is the angular velocity of the Earth and ϕ is the latitude. The bottom friction (τ'_{bx} and τ'_{by}) can be included using a variety of formulations. Here we use a quadratic stress representation based on a constant drag coefficient, C_d.

$$(\tau'_{bx}, \tau'_{by}) = -C_d\sqrt{u^2 + v^2}(u, v) \tag{6}$$

The bottom friction is denoted with primes to indicate that we have used specific quantities with the fluid density divided out.

3. Dual-time-stepping approach

The time derivative term is discretized with a one-step backwards difference scheme. Higher-order implicit schemes can also be used, but the one-step scheme simplifies the analysis. This results in an implicit set of equations that must be inverted at each time step. A dual-time-stepping approach is used to perform the inversion. In this method, an explicit advancement in pseudo-time is performed in order to solve the implicit equations. The form of the iteration is as follows:

$$P^{-1}\frac{\partial w}{\partial t^*} + \frac{w - w_{n-1}}{\Delta t} + \frac{\partial f_x}{\partial x} + \frac{\partial f_y}{\partial y} - s = 0 \tag{7}$$

where t^* is the pseudo-time variable and Δt is the physical time step. Because the one-step implicit scheme is A-stable [18], Δt is not constrained by stability criteria and can be set solely based on the desired accuracy of the simulation. The variable w_{n-1} is the solution from the previous time step. All other variables and fluxes are evaluated using $w(t^*)$. When this iteration reaches a steady state in pseudo time, the implicit problem is solved and the solution has been advanced from w_{n-1} to w_n. The matrix P is a preconditioning matrix designed to accelerate convergence in pseudo-time. We assume a simple form for P in which the u and v momentum equations are treated uniformly. P can thus be written as

$$P^{-1} = \begin{bmatrix} p^{-1} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{8}$$

The goal of this work is to find an expression for p that enables rapid convergence for a wide range of flow conditions.

4. Spatially discrete system

To analyze the above problem and determine an optimal preconditioning matrix, we first linearize the equations about a uniform flow state and bathymetry given by $u = u_0$, $v = 0$, $h = h_0$. In this case, the equations become

$$P^{-1}\frac{\partial w}{\partial t^*} + A\frac{\partial w}{\partial x} + B\frac{\partial w}{\partial y} = Sw \tag{9}$$

where A and B are the Jacobian matrices given by

$$A = \begin{bmatrix} 0 & 1 & 0 \\ -u_0^2 + gh_0 & 2u_0 & 0 \\ 0 & 0 & u_0 \end{bmatrix} \tag{10}$$

$$B = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & u_0 \\ gh_0 & 0 & 0 \end{bmatrix} \tag{11}$$

and Sw is the linearization of the source terms with S given by

$$S = \begin{bmatrix} -\Delta t^{-1} & 0 & 0 \\ 2C_d u_0^2/h_0 & -\Delta t^{-1} - 2C_d u_0/h_0 & f \\ 0 & -f & -\Delta t^{-1} - C_d u_0/h_0 \end{bmatrix} \tag{12}$$

where the backwards time discretization has been included in the source terms. Constant terms in the linearization and the backwards difference time advancement scheme are not included because these do not affect the convergence rate.

System analysis is ameliorated by multiplying by P, effectively multiplying the continuity equation by the preconditioning parameter p. We will denote the matrices PA and PB as $\widetilde{A}$ and $\widetilde{B}$. The above equations are discretized in space using a characteristic upwind scheme based on the preconditioned system. In this case, the x-discretization is as follows

$$\widetilde{A}\frac{\partial w}{\partial x} = \widetilde{A}\frac{w_{j+1,k} - w_{j-1,k}}{2\Delta x} - |\widetilde{A}|\frac{w_{j+1,k} - 2w_{j,k} + w_{j-1,k}}{2\Delta x} \tag{13}$$

where j and k are the horizontal and vertical mesh indices respectively, and $|\widetilde{A}|$ is the absolute value of the matrix defined as $V|\Lambda|V^{-1}$ where V and Λ are the eigenvector and eigenvalue matrix of $\widetilde{A}$. The y-direction discretization is analogous.

The eigenvalues of the matrix $\widetilde{A}$ are given by

$$\begin{bmatrix} u_0 \\ u_0 + \sqrt{u_0^2(1-p) + pgh_0} \\ u_0 - \sqrt{u_0^2(1-p) + pgh_0} \end{bmatrix} \tag{14}$$

If we let p be unity, we obtain the classical wave propagation speeds for the one-dimensional shallow water equations: u_0, $u_0 + \sqrt{gh_0}$, and $u_0 - \sqrt{gh_0}$. If the quantity under the square root becomes negative, the eigenvalues change from real to imaginary, indicating a change in the system of equations from hyperbolic to elliptic. This causes a modification of the characteristics and required boundary conditions and is thus undesirable. Using the constraint that the eigenvalues remain real gives an upper bound on p as

$$p < \frac{Fr^2}{Fr^2 - 1} \tag{15}$$

where the Froude number is defined as $u_0/\sqrt{gh_0}$. This upper bound must be satisfied only when Fr is greater than one (supercritical flow). In the limit that Fr goes to infinity, p must be less than or equal to one. The lower bound on p is zero, a condition derived from requiring the eigenvalues of the matrix $\widetilde{B}$, given by

$$\begin{bmatrix} 0 \\ \sqrt{pgh_0} \\ -\sqrt{pgh_0} \end{bmatrix} \tag{16}$$

to remain real.

5. Analysis

To analyze the above system, we Fourier transform in both the x- and y-directions. In this case the x-direction derivative becomes

$$\widetilde{A}\frac{\partial w}{\partial x} = \frac{\widetilde{A}}{\Delta x} i \sin(k_x \Delta x) + \frac{|\widetilde{A}|}{\Delta x}(1 - \cos(k_x \Delta x)) \tag{17}$$

where k_x is the wavenumber in the x-direction. The y-direction becomes

$$\widetilde{B}\frac{\partial w}{\partial y} = \frac{\widetilde{B}}{\Delta y} i \sin(k_y \Delta y) + \frac{|\widetilde{B}|}{\Delta y}(1 - \cos(k_y \Delta y)) \tag{18}$$

and the problem is reduced to a set of ODE's for each k_x, k_y wavenumber combination. The basic idea of the analysis is to find p to equilibrate the time scales of high wavenumber combinations. This ensures that the preconditioner will provide good damping for high wavenumber errors and thus make it useful for a multigrid algorithm.

Three high wavenumber cases must be analyzed to ensure good damping. These are $(k_x\Delta x, k_y\Delta y) = (\pi, 0), (0, \pi)$, and (π, π). (high wavenumber in x, high wavenumber in y, and high wavenumber in both). It is intractable to symbolically find the eigenvalues of these systems in the general case, so we approach the problem by taking limits. These are: large and small physical time step and Coriolis and friction dominated conditions, each for sub and supercritical flow.

5.1. Steady solutions

For the steady flow case, Δt^{-1}, f and C_d are set to zero. The wavenumber case of $(k_x\Delta x, k_y\Delta y) = (\pi, 0)$ then simplifies to

$$\frac{\partial w}{\partial t^*} + \frac{|A|}{\Delta x} w = 0 \tag{19}$$

The wavenumber $(k_x\Delta x, k_y\Delta y) = (0, \pi)$ simplifies to

$$\frac{\partial w}{\partial t^*} + \frac{|B|}{\Delta y} w = 0 \tag{20}$$

and for the bidirectional wavenumber case,

$$\frac{\partial w}{\partial t^*} + \left(\frac{|A|}{\Delta y} + \frac{|B|}{\Delta y}\right) w = 0 \tag{21}$$

We analyze the above three wavenumber cases separately for subcritical ($u_0 < \sqrt{gh_0}$) and supercritical ($u_0 > \sqrt{gh_0}$) conditions.

In the subcritical case, the eigenvalues of the above three equations are given by

$$\begin{bmatrix} Fr \\ Fr + \sqrt{Fr^2(1-p) + p} \\ -Fr + \sqrt{Fr^2(1-p) + p} \end{bmatrix} \quad \begin{bmatrix} 0 \\ \sqrt{p} \\ \sqrt{p} \end{bmatrix} \quad \begin{bmatrix} Fr + \sqrt{p} \\ \sqrt{Fr^2(1-p) + p} + \frac{\sqrt{p} + \sqrt{4Fr^2 + p}}{2} \\ \sqrt{Fr^2(1-p) + p} - \frac{\sqrt{p} + \sqrt{4Fr^2 + p}}{2} \end{bmatrix} \tag{22}$$

where we have non-dimensionalized by the wave speed $\sqrt{gh_0}$ and the grid size Δx. We also have made the assumption that the mesh is isotropic, $\Delta x = \Delta y$.

Examining the first set of eigenvalues, we see that if Fr goes to zero while p remains finite, the systems will be infinitely stiff with eigenvalues $(0, \sqrt{p}, -\sqrt{p})$. If the second two eigenvalues are to be of the same magnitude as the first (Fr), we must have $p \sim Fr^2$. Substituting the form $p = \eta^2 Fr^2$ and taking the limit as Fr goes to zero, we obtain

$$Fr\left(\begin{bmatrix}1\\ 1+\sqrt{1+\eta^2}\\ -1+\sqrt{1+\eta^2}\end{bmatrix}\quad\begin{bmatrix}0\\ \eta\\ \eta\end{bmatrix}\quad\begin{bmatrix}1+\eta\\ \sqrt{1+\eta^2}+1/2\eta+1/2\sqrt{4+\eta^2}\\ \sqrt{1+\eta^2}+1/2\eta-1/2\sqrt{4+\eta^2}\end{bmatrix}\right) \tag{23}$$

There is one eigenvalue that is zero, thus by definition the condition number of this system is infinite. The zero eigenvalue arises because there is no coupling in the y-direction for an eigenmode that is high-wavenumber only in y. This mode has physical significance and corresponds to a slip phenomena. If we equilibrate the remaining eigenvalues, we find that the minimum condition number occurs when $\eta^2 = 3$. Thus in the limit of low Fr, p should be $3Fr^2$.

In the supercritical case, the eigenvalues are similar, given by

$$\begin{bmatrix}Fr\\ Fr+\sqrt{Fr^2(1-p)+p}\\ Fr-\sqrt{Fr^2(1-p)+p}\end{bmatrix}\quad\begin{bmatrix}0\\ \sqrt{p}\\ \sqrt{p}\end{bmatrix}\quad\begin{bmatrix}Fr+\sqrt{p}\\ Fr+1/2\sqrt{p}+1/2\sqrt{4(1-p)Fr^2+5p}\\ Fr+1/2\sqrt{p}-1/2\sqrt{4(1-p)Fr^2+5p}\end{bmatrix} \tag{24}$$

In the supercritical limit, p cannot exceed one. If we let $p = 1$ and let Fr be large, the first and third set of eigenvalues are all proportional to Fr. The second set of eigenvalues are all small compared to Fr. This occurs because the flow variables are all being swept in the x-direction. The coupling in the y-direction is small because the wave speed is small relative to the convective velocity. There is no need to precondition in this limit because the problem becomes essentially one-dimensional with all variables propagating at the speed u_0.

A general form for the preconditioner for both of these cases is then

$$p = \frac{3Fr^2}{3Fr^2+1} \tag{25}$$

To determine a time step for the pseudo-time iteration, we also need to know the maximum eigenvalue. If we substitute the above form for p into the eigenvalues, we find that in the subcritical case the maximum dimensionless eigenvalue approaches 4.1 Fr and in the supercritical case it approaches Fr. These are reasonably approximated by

$$\lambda_{\max} = Fr\left(1+\frac{2+\sqrt{7}-\sqrt{3}}{2}(1-\sqrt{p})\right)+\sqrt{p} \tag{26}$$

5.2. Unsteady simulations

The next problem is the limit of small CFL, $u_0\Delta t/\Delta x$. There is no need to study the large CFL limit because it is essentially the steady problem studied above. To analyze the small CFL limit, u_0, f, and C_d are set to zero. Using our previous non-dimensionalization, there is only one remaining parameter, $\sqrt{gh_0}\Delta t/\Delta x$. A small value for this ratio implies that the time step is small relative to the time required for the gravity wave to propagate one grid cell. In this case, the equations are dominated by the Δt^{-1} term and the system is basically diagonal. Thus, $p = 1$ is appropriate.

The more interesting problem is the subcritical case in which the CFL number is small, but $\sqrt{gh_0}\Delta t/\Delta x$ is large i.e. the gravity waves propagate through many grid cells in one time step. For this case, the eigenvalues of the three different wavenumber combinations are

$$\begin{bmatrix} \Delta t^{-1} \\ \sqrt{p} + p\Delta t^{-1} \\ \sqrt{p} + \Delta t^{-1} \end{bmatrix} \quad \begin{bmatrix} \sqrt{p} + \Delta t^{-1} \\ \sqrt{p} + p\Delta t^{-1} \\ \Delta t^{-1} \end{bmatrix} \quad \begin{bmatrix} 2\sqrt{p} + p\Delta t^{-1} \\ \sqrt{p} + \Delta t^{-1} \\ \sqrt{p} + \Delta t^{-1} \end{bmatrix} \tag{27}$$

where we have again non-dimensionalized using $\sqrt{gh_0}$ and Δx. In the limit of $\Delta t^{-1} \ll 1$ (large time step relative to gravity wave-based CFL), the problem will be badly conditioned unless $p \sim \Delta t^{-2}$. Substituting $p = \eta^2 \Delta t^{-2}$ and taking the limit that Δt^{-1}, we find that the minimum condition number occurs when $\eta = 1$.

A form for the preconditioner that is valid for the small CFL limit is

$$p = \frac{\Delta t^{-2}}{\Delta t^{-2} + 1} \tag{28}$$

Using this form, p is equal to Δt^{-2} when Δt^{-1} is small and 1 when it is large (relative to $\sqrt{gh_0}/\Delta x$). The maximum eigenvalue occurs when Δt^{-1} is small and has a value of $2\Delta t^{-1}$. $\lambda_{\max}$ asymptotes to Δt^{-1} as Δt^{-1} increases. $\lambda_{\max}$ in both cases is bounded by $\Delta t^{-1} + \sqrt{p}$.

5.3. Friction-dominated flow

The parameter governing the importance of friction is the nondimensional surface drag number $\alpha = C_d \Delta x / h_0$. This parameter is normally denoted r, however, for clarity in Eq. (29) through (31) we use α. When α is large, damping at the grid scale associated with friction dominates the advective terms. To see this, we compare the source matrix S (Eq. (12)) to the matrix A. Any entry proportional to u_0 in A has a corresponding entry proportional to $C_d u_0 / h_0$ in S. Assuming α is large, we can then neglect the terms proportional to u_0 in the A matrix.

Assuming α is large, for the subcritical case, the eigenvalues of the three wavenumber cases become

$$\begin{bmatrix} \sqrt{p} \\ \alpha Fr \\ \sqrt{p} + 2\alpha Fr \end{bmatrix} \quad \begin{bmatrix} 2\alpha Fr \\ \sqrt{p} \\ \sqrt{p} + \alpha Fr \end{bmatrix} \quad \begin{bmatrix} \sqrt{p} + 2\alpha Fr \\ \sqrt{p} + \alpha Fr \\ 2\sqrt{p} \end{bmatrix} \tag{29}$$

To equilibrate the damping rates of all eigenvalues as much as possible, p must take the form $\alpha^2 Fr^2$ in this limit.

In the supercritical limit, if we let α approach infinity while p and Fr remain bounded, the eigenvalues are

$$\begin{bmatrix} \alpha Fr \\ 2\alpha Fr \\ pFr \end{bmatrix} \quad \begin{bmatrix} 2\alpha Fr \\ \alpha Fr \\ \sqrt{p} \end{bmatrix} \quad \begin{bmatrix} \alpha Fr \\ \sqrt{p} + pFr \\ 2\alpha Fr \end{bmatrix} \tag{30}$$

In this case, there is no choice of p that will equilibrate the eigenvalues. In the limit of large Fr, p is bounded by 1. If α is large, then αFr will be much larger than Fr and the system will be poorly conditioned. A choice of p that gives the correct scaling in the subcritical limit and does not exceed the required bounds on p in the supercritical limit is

$$p = \frac{\alpha^2 Fr^2}{\alpha^2 Fr^2 \max(0, (Fr^2 - 1)/Fr^2) + 1} \tag{31}$$

With this choice of p, the maximum eigenvalue in both the sub- and supercritical friction-dominated cases is bounded by $\sqrt{p} + 2\alpha Fr$.

5.4. Coriolis-dominated conditions

In this case, the flow dynamics are dominated by the Coriolis terms and we set $u_0, \Delta t^{-1}$, and C_d to 0. This corresponds to the limit that the grid scale Rossby number, $u_0/(f\Delta x)$ is small. There is only one relevant parameter in this case which is $\sqrt{gh}/(f\Delta x)$, the ratio of the Rossby deformation radius to the local grid scale. In the subcritical case, the eigenvalues are lengthy and are not reproduced here. Upon examination, we find

that unless p scales as f^2, the time scales associated with the momentum equations will be much different than for the continuity equation. Following the same procedure, we introduce $p = \eta^2 f^2$, with the result that all eigenvalues scale linearly with f. The minimum condition number occurs when $\eta = 1$. $\lambda_{\max}$ becomes $2f$ and is independent of $\sqrt{gh}/(f\Delta x)$.

In the supercritical case, p cannot exceed $Fr^2/(Fr^2 - 1)$. If we take the limit as f approaches infinity with Fr and p bounded for each of the three wavenumber cases, two of the eigenvalues approach $\pm if$. For the high wavenumber in x case, the third eigenvalue scales as $1/f^2$ so this mode will be badly conditioned. For the other two cases, the third eigenvalue approaches a constant in the limit of large f. Thus, these two modes will also be badly conditioned. The poor conditioning for this limit may not be detrimental because for most practical supercritical flow problems, the associated length scales are considerably less than the Rossby radius of deformation.

The following definition of p for Coriolis-dominated conditions gives the correct behavior for the subcritical case and ensures that p remains properly bounded for the supercritical case:

$$p = \frac{f^2}{f^2 \max(0, (Fr^2 - 1)/Fr^2) + 1} \tag{32}$$

The maximum eigenvalue in both cases is bounded by $f + \sqrt{p}$.

5.5. Combined result

If we combine all of the previous results and generalize to a flow moving in any direction, we arrive at the following result for p and $\lambda_{\max}$ which can be written in dimensional form as

$$p = \frac{(3 + \alpha^2)||V||^2 + \Delta_{\max}^2(\Delta t^{-2} + f^2)}{(3 + \sigma\alpha^2)||V||^2 + \Delta_{\max}^2(\Delta t^{-2} + \sigma f^2) + gh} \tag{33}$$

with σ defined as

$$\sigma = \max\left(\frac{||V||^2 - gh}{||V||^2}, 0\right) \tag{34}$$

and $\Delta_{\max}^2$ defined as $\Delta x^2 + \Delta y^2$. Here, α has been generalized to $C_d \Delta_{\max}/(h)$ and $\|V\|$ is the magnitude of the flow velocity. With this choice of preconditioner, $\lambda_{\max}$ is given by

$$\lambda_{\max} = \left(1 + \frac{2 + \sqrt{7} - \sqrt{3}}{2} \max\left(\sqrt{1} - \sqrt{p}, 0\right)\right) \frac{||V||}{\Delta_{\min}} + \Delta t^{-1} + f + \frac{2\alpha||V||}{\Delta_{\min}} + \frac{\sqrt{pgh}}{\Delta_{\min}} \tag{35}$$

where $\Delta_{\min} = \Delta x \Delta y / \Delta_{\max}$. The use of $\Delta_{\min}$ gives a conservative estimate of $\lambda_{\max}$ on anisotropic meshes.

6. SUPG stabilization

The numerical validation of the above preconditioner will be performed using a streamwise-upwind Petrov–Galerkin (SUPG) finite element formulation [19] which uses the weak form

$$0 = \sum_{e=1}^{n_{el}} \left\{ \int_{\Omega_e} \left[\phi^T \left(\frac{\partial w}{\partial t} - s \right) - \frac{\partial \phi^T}{\partial x} f_x - \frac{\partial \phi^T}{\partial y} f_y \right] d\Omega_e + \int_{\Omega_e} \left[\frac{\partial \phi^T}{\partial x} A + \frac{\partial \phi^T}{\partial y} B \right] \tau \left[\frac{\partial w}{\partial t} - s + \frac{\partial f_x}{\partial x} + \frac{\partial f_y}{\partial y} \right] d\Omega_e \right\} + \int_{\Gamma} \phi^T (f_x, f_y) \cdot \vec{n}\, d\Gamma \tag{36}$$

where the sum is performed over all of the elements in the mesh, 1 to n_{el}, and ϕ is a vector test function. Ω_e is the element area. Γ and $\vec{n}$ are the boundary to the domain and the outward normal respectively. The matrix τ is a stabilization matrix that must be determined.

The classical definition of τ is

$$\tau = \frac{\Delta_{\max}}{2} V \widehat{\Lambda}^{-1} V^{-1} \tag{37}$$

where V is the matrix of eigenvectors of $|A| + |B|$ and $\widehat{\Lambda}$ is defined as

$$\widehat{\Lambda} = |\Lambda| + \Delta_{\max}\mathrm{diag}(V^{-1}SV) \tag{38}$$

where Λ is the eigenvalue matrix of $|A| + |B|$. In the following we use an alternative definition that is easier to analyze and more appropriate when the source terms strongly influence the dynamics:

$$\tilde{\tau} = \frac{\Delta_{\max}}{2}\left(|\widetilde{A}| + |\widetilde{B}| + \Delta_{\max} S\right)^{-1} \tag{39}$$

Using $\widetilde{A}$ and $\widetilde{B}$ in the definition of $\tilde{\tau}$ gives the stabilization for the preconditioned system. Using this definition, we can see that the stabilization matrix is closely tied to the preconditioning analysis of modes that are high-wavenumber in both x and y. Often the form of $\tilde{\tau}$ is simpler because the time scales of the preconditioned system have been normalized to be the same magnitude. If the time scales were exactly equal, $\tilde{\tau}$ would be diagonal because it could be written in the form $V\Lambda^{-1}V^{-1}$ where the eigenvalues and eigenvectors are derived from the matrix inside the parenthesis in Eq. (39). If all the λ's of this matrix are the same, this becomes $VV^{-1}/\lambda = I/\lambda$ where I is the identity matrix.

For the steady and unsteady super and subcritical cases, as well as the subcritical friction- and Coriolis-dominated cases, the preconditioner was able to reasonably equilibrate the eigenvalues of the system for the case that is high-wavenumber in both x and y. Therefore for these cases, the eigenvalues are all similar and as a reasonable estimate of these eigenvalues we can use $\lambda_{\max}$. $\tilde{\tau}$ for these cases is then

$$\tilde{\tau} = \lambda_{\max}^{-1} I \tag{40}$$

The more difficult cases to analyze are supercritical friction- and Coriolis-dominated flow. For these conditions, the eigenvalues of the bidirectional high-wave number system are found to scale differently. (For the friction-dominated case, this can be seen by examining the third column of eigenvalues in Eq. (30).) For these cases we use Eq. (39) directly. In the limit of large α Eq. (39) gives

$$\tilde{\tau} = \begin{bmatrix} \frac{1}{2Fr\alpha\left(1+\sqrt{p}Fr\right)} & 0 & \frac{1}{\sqrt{p}+pFr} \\ 0 & \frac{1}{Fr\alpha} & 0 \\ -\frac{\sqrt{p}}{2\alpha\left(1+\sqrt{p}Fr\right)} & 0 & \frac{1}{\sqrt{p}+pFr} \end{bmatrix} \tag{41}$$

The matrix is not diagonal. The $\tilde{\tau}$ matrix for the Coriolis-dominated case is also not diagonal. As supercritical friction- and Coriolis-dominated flows are less relevant to physical applications, the increased computational cost associated with employing a non-diagonal form for $\tilde{\tau}$ is not worthwhile. Instead we use the definition given by Eq. (40). Because this formulation uses $\lambda_{\max}$ for all three eigenvalues, the stabilization terms in the supercritical friction and Coriolis-dominated cases will be smaller than optimal.

The above stabilization matrix is for a system that is multiplied by P, but in actual implementations the inverse of the preconditioning matrix would multiply the pseudo-time derivative term. To revert back to the original form, we note that the stabilization term can be written (in 1D, but the same arguments hold in 2D) as

$$\widetilde{A}\tilde{\tau}\widetilde{A} = PA\tilde{\tau}PA. \tag{42}$$

To revert to the original system, we must multiply by P^{-1}. This gives

$$A\tilde{\tau}PA \tag{43}$$

and we can therefore define the stabilization matrix for the original system, τ, as $\tilde{\tau}P$. The resulting final form for τ is then

$$\tau = \lambda_{\max}^{-1}\begin{bmatrix} p & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{44}$$

7. Numerical tests

To verify the optimality of the preconditioner, we examine convergence rates from a SUPG discretization using the stabilization defined above. The details of the finite element code are described in Ref. [20]. It employs a finite element discretization of Eq. (1) on triangular elements with a continuous, piecewise polynomial basis. Although this scheme allows high-order polynomials, for simplicity we restrict the study to linear polynomials. When using linear polynomials, the discretization is analogous to an unstructured, vertex-based, finite-volume discretization. The discrete equations are solved using multigrid. The multigrid algorithm is a full approximation and storage (FAS) algorithm [21], which means that it obtains the solution to the non-linear equations. The prolongation and restriction operators are the standard linear interpolation operators used for unstructured, vertex-based, triangular meshes [21]. The relaxation scheme is given by Eq. (7) with a 5-stage explicit Runge–Kutta scheme [22] to advance the solution in t^*. Eqs. (33) and (35) for p and $\lambda_{\max}$ are generalized for use on an unstructured mesh by taking $\|V\|$ and h as the maximum value over the three vertices associated with a given element. $\Delta_{\max}$ is the maximum length of the three element sides, $\Delta_{\min}$ is twice the triangle area divided by $\Delta_{\max}$. The time step in pseudotime Δt^* is inversely proportional to $\lambda_{\max}$. A "local time-stepping" approach is used where both p and Δt^* are calculated for each element.

7.1. Free-stream flow

The test problems in this section correspond to the analyses performed in Section 5. They are uniform flows with speed u_∞ on a unit square. On the left and right side of the square, characteristic boundary conditions are employed. The characteristic decomposition is based on the $\widetilde{A}$ matrix whose eigenvalues are given in Eq. (10). Periodic boundary conditions are used on the upper and lower boundaries. The problems are non-dimensionalized such that the baseline depth and the gravitational acceleration are unity, therefore u_∞ corresponds to the Froude number (Fr). The mesh is composed of 32×32 squares that are then subdivided into triangles. Coarser meshes in the multigrid cycle have the same structure except the resolution is halved. Six levels of grids are used such that the coarsest grid consists of only two triangles.

We examine the eight cases shown in Table 1. The first two cases examine steady solutions with no source terms in the sub- and supercritical limit. The third and fourth cases correspond to unsteady simulations. In the third, the implicit time step gives a convective CFL number equal to 1 and a wave-speed-based CFL of 100. In the fourth case, the convective CFL is 0.1 and the wave-speed-based CFL is 10.0. It is not necessary to examine the supercritical case because this case corresponds to a steady supercritical flow when the time step is large or a flow regime where an explicit simulation could be performed when the time step is small. The last four cases test the friction- and Coriolis-dominated limit in sub- and supercritical conditions. For these cases, a constant source term is added to the governing equations so that solution converges to uniform flow. To determine whether the preconditioner provides optimal convergence rates, the optimal value of the preconditioner, p_{opt}, given by Eq. (33), is multiplied by five factors: 1/4, 1/2, 1, 2, and 4. For initial conditions we use a perturbation to free-stream that includes both high- and low-wavenumber components

Table 1
Cases studied

Case no.	$\frac{u_\infty}{\sqrt{gh}}$	$\frac{\Delta t\sqrt{gh}}{\Delta x}$	$\frac{C_d \Delta x}{h}$	$\frac{f\Delta x}{u_\infty}$	Comments
1	0.01	0	0	0	Subcritical, steady
2	100	0	0	0	Supercritical, steady
3	0.01	100	0	0	Subcritical, convective Δt
4	0.01	10	0	0	Subcritical, intermediate Δt
5	0.01	0	10	0	Subcritical, friction
6	100	0	10	0	Supercritical, friction
7	0.01	0	0	10	Subcritical, Coriolis
8	100	0	0	10	Supercritical, Coriolis

$$\begin{aligned}\frac{u}{u_\infty} &= 1 + Ax(1-x)(\sin(2\pi x) + \sin(16\pi x))(\sin(2\pi y) + \sin(16\pi y))\\ \frac{v}{u_\infty} &= Ax(1-x)(\sin(2\pi x) + \sin(16\pi x))(\sin(2\pi y) + \sin(16\pi y))\\ h &= 1.0 + A_h x(1-x)(\sin(2\pi x) + \sin(16\pi x))(\sin(2\pi y) + \sin(16\pi y))\end{aligned} \tag{45}$$

where A is the amplitude of the velocity perturbation and A_h is the amplitude of the depth perturbation. For all of the calculations, A was set to 0.01. A_h was set to 0.0001 in the subcritical case and 0.01 in the supercritical case. It was necessary for A_h to be smaller in the subcritical case because small changes in free-surface height result in large changes in flow velocity when the flow is subcritical.

The tests with multigrid are performed using a W-cycle with two five-stage Runge–Kutta relaxations performed after each restriction and prolongation phase of the multigrid. With this configuration, the scheme becomes essentially a two-level iteration; the equations on the coarse grid are solved precisely for each fine grid iteration. This set-up was chosen so that we could more accurately assess the performance of the relaxation scheme in damping the high wavenumbers. It would be more computationally efficient to use a V-cycle with less relaxations on each grid level, but in previous work [23] we have found that if the coarse grid residuals are not significantly reduced, a boundary-condition-driven instability can develop. Table 2 shows the convergence factors for each case studied. The convergence factor is the ratio of the magnitude of the residuals before and after one multigrid cycle. The W-cycle is terminated on first return to the finest mesh, thus there are four relaxations on the fine mesh per cycle (two before the restriction to the first coarser mesh, and two after the prolongation back to the fine mesh). We determine a mean convergence factor by iterating until machine convergence is reached then fitting an exponential curve to a plot of the residual as a function of iteration number. In fitting the curve, we ignore the initial region of the curve where faster decaying modes are eliminated and only fit the data after a nearly constant convergence factor is reached. This convergence factor is then due to the slowest decaying mode of the iteration. The number in parentheses is the number of iterations required to reduce the residual error one order of magnitude based on the mean convergence factor. A "u" indicates that the iteration failed to converge and a "–" indicates that the chosen value of p violated the limit established in Eq. (15).

There are several significant trends in this table. Cases 1 (steady, subcritical), 3 (convective CFL), and 4 (intermediate CFL), all converge faster with $1/4\ p_{\text{opt}}$. We have rerun these cases with $p/p_{\text{opt}} = 1/8$ and found that the optimal value is at $p/p_{\text{opt}} = 1/4$. There are several possible explanations for why the optimum was not exactly predicted by the analysis. First, there are implementation differences because the tests are run on an unstructured mesh with different definitions of Δx and Δy than in the analysis. Second, the analysis was based on a first-order upwind scheme rather than the SUPG FEM stabilization. Lastly, the analysis was based on optimal damping of only the highest wavenumber on the mesh. Optimal multigrid convergence requires that the entire spectrum of nondimensional wavenumbers from $\pi/2$ to π be damped. In spite of this, the analysis did an excellent job predicting the scaling of p for optimum convergence and a reasonably accurate job of predicting the exact value necessary. For all of the remaining cases except case 8, the analysis predicted the value

Table 2
Convergence factors for various dynamical limits in free stream flow

p/p_{opt}	1/4	1/2	1	2	4
Case 1	0.91 (25.8)	0.93 (30.2)	0.94 (39.2)	0.96 (54.5)	0.95 (47.4)
Case 2	u	u	0.75 (8.1)	–	–
Case 3	0.35 (2.2)	0.36 (2.3)	0.42 (2.6)	0.46 (3.0)	0.52 (3.5)
Case 4	0.12 (1.1)	0.16 (1.3)	0.24 (1.6)	0.34 (2.2)	0.40 (2.6)
Case 5	0.65 (5.3)	0.47 (3.0)	0.35 (2.2)	0.37 (2.3)	0.42 (2.6)
Case 6	0.97 (88.4)	0.97 (73.1)	0.94 (40.5)	–	–
Case 7	u	u	0.94 (35.4)	0.95 (43.1)	0.97 (74.7)
Case 8	u	u	u	–	–

A range of p/p_{opt} is tested in a w-cycle multigrid iteration. A "u" indicates divergence and a "–" indicates that the chosen value of p violates the limit established in Eq. (15). The numbers in parentheses are the number of iterations required to reduce the error one order of magnitude.

of p_{opt} perfectly; other values of p either give larger convergence factors, diverged, or violated the maximum constraint.

As expected from the analysis, the supercritical friction-dominated case (6) and supercritical Coriolis-dominated case (8) do not perform well. For the friction-dominated case, the convergence factor is 0.94 and in the Coriolis-dominated case the iteration does not converge at all. As mentioned previously, these cases are not typically physically significant, but if there was a problem that encountered these conditions, one would either have to decrease the physical time step until it stabilized the iteration or use another iterative approach. The Coriolis-dominated case is especially difficult because the Coriolis terms cause oscillation at the highest wavenumbers rather than damping.

Several of the iterations, although convergent, required a large number of iterations to reduce the error an order of magnitude. These include cases 1 and 7 and to a lesser extent case 2. Cases 1 and 2 are the subcritical and supercritical steady cases. In the analysis of these cases, the conditions that were high wavenumber in y and 0 wavenumber in x resulted in a zero eigenvalue (middle set of eigenvalues in Eqs. (22) and (24)). These are high-wavenumber error modes that lack any damping because the local discretization allows slip between adjacent fluid layers. The only way the iteration can eliminate slip is by propagating the specified uniform inlet velocity across the entire domain. Because the error mode is high wavenumber in the y-direction, multigrid does not help this process, and the convergence rate should be dependent on the number of grid cells in the mesh. To verify this, we repeated cases 1 and 2 on a 64×64 grid. The convergence factor for case 1 on the finer grid is 0.95 (49.1) versus 0.94 (39.2) on the 32×32 grid. Thus, for this case the convergence rate depends on the mesh resolution. In case 2, the supercritical case, the convergence factor is 0.75 (7.9) versus 0.75 (8.1). Thus, in this case the convergence is grid independent. This is rather unexpected because the zero eigenvalue should affect the convergence rate. For case 7 (Coriolis-dominated, subcritical), there is no zero eigenvalue and the results should be grid independent. For that case, the convergence factor on the 64×64 grid is 0.93 (30.9) and 0.94 (35.4) on the 32×32 grid, indicating grid-independence. The reason that convergence is slow is again because the Coriolis terms cause oscillation rather than damping.

Based on the results for these tests, we introduce a modified version of the preconditioner, defined as

$$p = \frac{(1+\alpha^2)||V||^2 + \Delta_{max}^2(\Delta t^{-2}/4 + f^2)}{(1+\sigma\alpha^2)||V||^2 + \Delta_{max}^2(\Delta t^{-2}/4 + \sigma f^2) + gh} \tag{46}$$

This modified formulation for the preconditioner is employed in the remaining tests.

7.2. Equatorial Rossby soliton

To examine the performance for a more realistic application, we consider a problem from geophysical fluid dynamics with a known analytical solution. The test problem considers the propagation of a Rossby soliton on an equatorial β-plane [24]. The soliton propagates westward at a fixed celerity and shape. While this is a common test case for consistency and accuracy of shallow water equation solvers, we will employ this case to test the effects of the preconditioner on iterative convergence in a flow with rotation. The β-plane approximation used for the Coriolis parameter is $f = f_o + \beta y$ with $f_o = 0$ as the problem is symmetric about the equator. The spatial and temporal coordinates of the problem are non-dimensionalized using combinations of the variables β and $\sqrt{gb}$ where b is the constant bathymetric depth. β can thus be set to one in the preceding expression. The computational domain is a rectangle with limits $-24 \leqslant x \leqslant 24$ and $-8.0 \leqslant y \leqslant 8.0$ (see Fig. 1). Bottom friction is neglected ($C_d = 0$).

Characteristic boundary conditions are used on the meridional sides of the domain, and for the zonal boundaries, a slip condition is employed. The initial conditions for the water depth and velocity nondimensionalized using b and $\sqrt{gb}$ are given by

$$\begin{aligned} h &= 1 + \phi(x)\left(\tfrac{3+6y^2}{4}\right)e^{-y^2/2} \\ u &= \phi(x)\left(\tfrac{-9+6y^2}{4}\right)e^{-y^2/2} \\ v &= \tfrac{\partial \phi}{\partial x}(2y)e^{-y^2/2} \end{aligned} \tag{47}$$

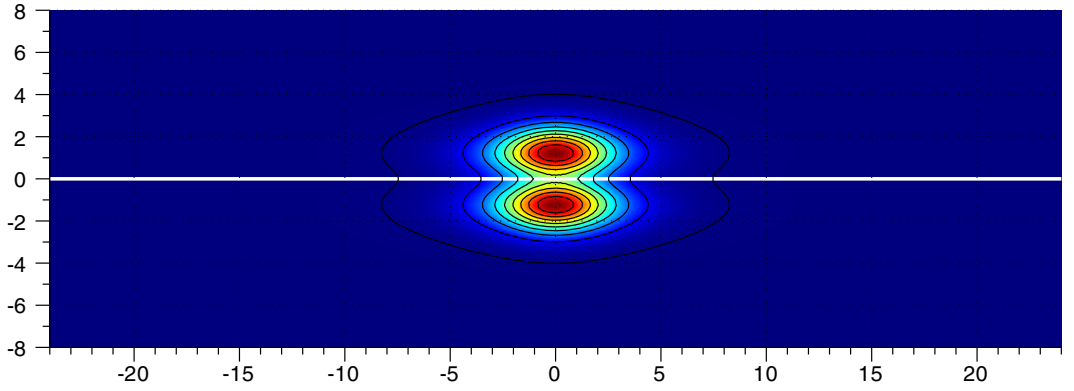

Fig. 1. Domain and initial surface elevation for Rossby soliton test case (*white line*: equator).

with

$$\begin{aligned}\phi(x) &= 0.771B^2\operatorname{sech}^2 Bx \\ \tfrac{\partial\phi}{\partial x} &= -2B\tanh(Bx)\phi \\ B &= 0.395\end{aligned} \tag{48}$$

The equatorial Rossby soliton test case is run using three time steps ($\Delta t = 0.5, 1.0, 2.0$) on a sequence of meshes with resolution ($\Delta x = 1.0, 0.5, 0.25, 0.125$). Because of the spatially varying flow velocity, most of the non-dimensional parameters can only be defined locally. The local Froude number is independent of grid resolution and time step but varies with the flow velocity. The minimum Fr is zero because the flow velocity approaches zero near the zonal boundaries. The maximum value is 0.25. The CFL number for the coarsest mesh has maximum value of 0.125 and for the finest mesh the maximum was 4. The CFL number based on the wave celerity ranges from 0.5 (small Δt coarse mesh) to 16 (large Δt fine mesh). The ratio of the grid scale to the local Rossby deformation radius, $\frac{f\Delta x}{\sqrt{gb}}$ varies linearly with y because of the β-plane assumption. For the coarse grid the maximum value was 8 and on the fine grid the maximum was 1. Overall, on the coarse mesh, the maximum explicit time step allowable is roughly 1/8 due to the Coriolis terms. On the fine mesh, the maximum allowable time step due to Coriolis terms and the wave celerity is both 1/8. If these effects are additive, then the maximum explicit time step is probably around 1/16.

For each mesh and time step, five preconditioning multiplicative factors are examined: $p/p_{\text{opt}} = 0.25, 0.5, 1.0, 2.0, 4$. For specification of p_{opt}, the modified preconditioner (Eq. (46)) is utilized. In the iterative scheme, local flow values are used to calculate a local p and iterative time step at each grid point. All cases with $p/p_{\text{opt}} = 4$ diverged. Results for cases with the remaining factors are shown in Table 3. In general, conditions with more dominant Coriolis effects (coarse grid with large time steps) converged slower. At the finest grid resolution using $\Delta t = 2.0$, it takes 7.5 cycles to reduce the error an order of magnitude. This corresponds to 30 relaxations because there are 4 fine grid relaxations per cycle. Because the allowable explicit time step for these conditions is roughly 1/16, an explicit scheme would take 32 time steps to simulate the same length of time. Thus, it seems that the two methods are comparable, although to make a fair comparison we should determine how much the iterative error must be reduced each time step to maintain a specified target accuracy. Factoring in the overhead of multigrid, the explicit scheme is probably more efficient. This is not unexpected because this problem does not have characteristics that make it stiff (Fr is O(1), uniform mesh resolution).

Looking at the variation of the damping factors with p/p_{opt} it seems that most cases converged best in the 1/2 to 2 range, which indicates that the scaling of p is primarily correct even for this case in which the conditions vary widely across the mesh. Coriolis-dominated conditions ($\Delta x = 1.0, \Delta t = 2.0$) again cause difficulties. It seems that one must consider a non-diagonal preconditioner to eliminate this problem.

Table 3
Convergence factors for the Rossby equatorial soliton test case

Δx	Δt	p/p_{opt}			
		1/4	1/2	1	2
1.0	2.0	0.97 (104.1)	0.96 (62.71)	u	u
0.5	2.0	0.80 (130.0)	0.95 (46.7)	0.90 (21.5)	u
0.25	2.0	0.95 (48.7)	0.90 (22.4)	0.877 (17.5)	0.81 (11.1)
0.125	2.0	0.95 (43.4)	0.87 (17.1)	0.74 (7.5)	0.79 (9.6)
1.0	1.0	0.58 (4.2)	0.83 (12.3)	u	u
0.5	1.0	0.40 (2.5)	0.50 (3.3)	0.74 (7.6)	u
0.25	1.0	0.69 (6.2)	0.55 (3.9)	0.64 (5.1)	0.77 (9.0)
0.125	1.0	0.84 (13.7)	0.72 (6.9)	0.60 (4.53)	0.66 (5.6)
1.0	0.5	0.53 (3.6)	0.48 (3.1)	u	u
0.5	0.5	0.51 (3.5)	0.31 (2.0)	0.61 (4.8)	u
0.25	0.5	0.60 (4.4)	0.40 (2.5)	0.49 (3.2)	0.69 (6.2)
0.125	0.5	0.76 (8.5)	0.58 (4.5)	0.49 (3.3)	0.55 (3.8)

A range of p/p_{opt} is tested in a preconditioned W-cycle multigrid iteration. A "u" indicates divergence. The numbers in parentheses are the number of iterations required to reduce the error one order of magnitude.

7.3. Tide propagation in a semi-enclosed channel

In this test, we examine the efficacy of the preconditioner for a tidally-driven flow at length and depth scales relevant to coastal oceanography. The test case is a long wave run-up on a linearly sloping bottom. Bottom friction and Coriolis terms are neglected ($C_d = 0, f = 0$). The domain is a semi-enclosed channel of length ($L = 580$ km) (Fig. 2). The problem is one-dimensional and thus the width is not a critical parameter. The bathymetry ranges from $H_o = 20$ m at the open boundary to 0 m at the closed end. As the code does not currently support flooding/drying dynamics, the domain is truncated at ($L = 575$ km) to maintain positive depth. The flow is forced at the open boundary ($x = 0$) using a prescribed free surface elevation of amplitude 1 cm and period $T_p = 12.42h$, corresponding to the dominant tidal component for most coastal regions. The problem has an analytical solution, based on linearized equations, which could in principle be used to evaluate the accuracy of the scheme. For our interests, however, this test case represents a tidal-flow problem with typical depth, length, and time scales for coastal ocean applications and thus is a good test of the preconditioner for time-dependent, low *Fr* flows.

The tidally-forced semi-enclosed channel test case was run using a range of time steps per tidal period (5, 10, 20, 40), and mesh resolution ($\Delta x = 1.25, 2.5, 5.0, 10$ km). For this problem, the mean Froude number over the wave period at the midpoint of the domain ($x = 290$ km) is approximately $Fr = 0.01$, thus the disparity between gravity wave speed and local convective speed is approximately two orders of magnitude. The

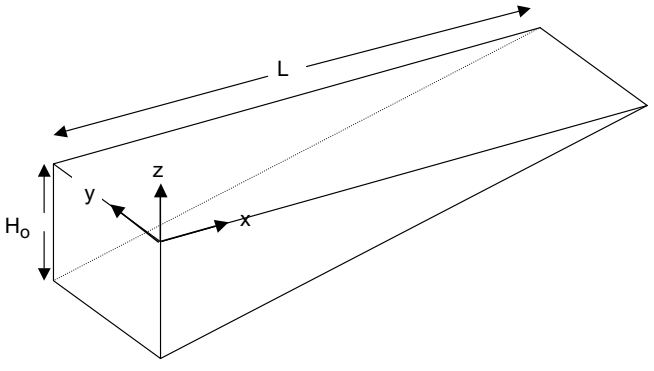

Fig. 2. Domain and coordinate system for semi-enclosed channel.

Table 4
Convergence factors for the tidally-forced channel

Δx (km)	$\frac{T_p}{\Delta t}$	1/4	1/2	1	2	4
p/p_{opt}						
10	5	0.57 (4.0)	0.37 (2.4)	0.30 (1.9)	1.4 (2.3)	0.45 (2.9)
1.25	5	0.62 (4.8)	0.70 (6.5)	0.76 (8.5)	0.80 (9.8)	0.80 (9.9)
10	40	0.46 (3.0)	0.39 (2.4)	0.42 (2.7)	0.36 (2.3)	0.35 (2.2)
1.25	40	0.52 (3.6)	0.31 (2.0)	0.32 (2.1)	0.33 (2.0)	0.44 (2.9)

A range of p/p_{opt} is tested in a preconditioned w-cycle multigrid iteration. The numbers in parentheses are the number of iterations required to reduce the error one order of magnitude.

wave-speed-based CFL $\left(\frac{\Delta t\sqrt{gh}}{\Delta x}\right)$ ranges from 1 to 100 based on mesh resolution and time step. The convective CFL ranges from 0.01 to 1.

Table 4 shows the convergence factors for the smallest and largest time step (displayed in time steps per forcing period, $\frac{T_p}{\Delta t}$) and coarsest and finest mesh resolutions for a range of preconditioning multiplicative factors p/p_{opt}. The results shown are averages of the convergence factor for time steps at quarter period intervals. Every case run converged. The fine mesh resolution, large time step case (results row 2), corresponds to a near unity convective CFL, similar to conditions from case 3 in the idealized numerical tests (Table 1). The low mesh resolution, small time step (results row 3) is similar to the conditions from case 4 in the idealized tests, with an intermediate time step. Convergence factors from the tidally-forced channel show that with the exception of the high resolution, large time step case, all cases converged rapidly. For the convective time step case, the smallest value of preconditioner multiplier was found to be optimal. Further reduction of the preconditioner to $p = \frac{1}{8}p_{opt}$ led to divergence. For this test case as well as both relevant idealized tests (case 3 and 4), an increase in wave speed CFL led to a decrease in convergence rate. As mentioned previously, this is because of the difficulty in eliminating slip modes. Our personal experience is that these modes become less problematic when the flow is not perfectly aligned with the mesh or the mesh is unstructured.

Assuming an order of magnitude convergence per time step is sufficient, the implicit formulation would most likely be more efficient than an explicit approach. For the fine grid with the large time step, 170 fine grid relaxations would be needed per period (8.5 cycles per time step, 4 fine grid relaxations per cycle, 5 time steps per period). An explicit scheme would require 5000 time steps per period based on the wave speed limited explicit time step. Again this comparison is rough but it does indicate that this approach can have significant benefits depending on the problem.

8. Conclusions

We have investigated preconditioning for the shallow water equations with inclusion of Coriolis terms and bottom friction. A formulation of the preconditioner parameter p was determined for use in a multigrid iteration and tested for dynamical limits relevant to the system. Coupling the preconditioner to a multigrid cycle, we demonstrated good convergence rates for most conditions. For all conditions, our choice of preconditioner parameter, p, was within a factor of 2 of optimum and most of the test cases converged rapidly. However, with this simple explicit preconditioner it was not possible to achieve good convergence for subcritical steady, Coriolis-dominated, and supercritical friction-dominated cases. Of these, the most important is the Coriolis-dominated case, and the preconditioner was not very robust for these conditions. Possible solutions may be to include a point-implicit treatment of the Coriolis terms in the iteration, but this was not examined here.

Two test cases were employed to examine the performance of the preconditioner for more realistic flows using a range of time steps and mesh resolutions. The first, a Rossby equatorial soliton tested the preconditioner for planetary-scale problems with strong rotational effects. For this test case, the multigrid iteration diverged at small grid scale Rossby number with $p = p_{opt}$. This is again because of the difficulties associated with Coriolis effects. All cases were found to converge using a more conservative $p = \frac{1}{2}p_{opt}$. Rough estimates indicate that the proposed fully implicit approach would be moderately slower than a standard explicit scheme for this problem. This is because the problem does not have much inherent stiffness.

The second test case, a tidally-forced semi-enclosed channel examined the preconditioner's performance for time-dependent low-Froude number flows at scales relevant to coastal oceanography. For this case, the preconditioner proved to be quite robust and good convergence was obtained for all time steps and mesh resolutions. A rough comparison to the work required for an explicit simulation indicated that more than an order of magnitude increase in efficiency is possible using the implicit approach.

Acknowledgments

B.T. Helenbrook was supported in part by the National Science Foundation under Award No. DMS-0513380. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect those of the National Science Foundation. G. Cowles was supported by the Massachusetts Marine Fisheries Institute (MFI) through NOAA Grants DOC/NOAA/NA04NMF4720332 and DOC/NOAA/NA05NMF4721131.

References

[1] K. Bryan, A numerical method for the study of the circulation of the world ocean, J. Comput. Phys. 4 (1969) 347–376.
[2] W. Gates, A numerical study of transient Rossby waves in a wind-driven homogenous ocean, J. Atmos. Sci. 25 (1968) 3–22.
[3] V. Casulli, Semi-implicit finite difference methods for the two-dimensional shallow water equations, J. Comput. Phys. 86 (1990) 56–74.
[4] J. Dukowicz, R. Smith, Implicit free-surface method for the Bryan–Cox–Semtner ocean model, J. Geophys. Res. 99 (C4) (1994) 7991–8014.
[5] B. Hodges, Accuracy order of Crank–Nicolson discretization for hydrostatic free-surface flow, J. Eng. Mech. 130 (2004) 8.
[6] K.M. Dresback, K.L. Kolar, An implicit time-marching algorithm for 2-D GWC shallow water models, in: Bentley et al., (Eds.), Computational Methods in Water Resources XIII, vol. 2, 2000, pp. 913–920.
[7] D. Le Roux, A. Staniforth, C.A. Lin, Finite elements for shallow-water equation ocean models, Mon. Weather Rev. 126 (7) (1998) 1931–1951.
[8] L. Begnudelli, B.F. Sanders, Unstructured grid finite-volume algorithm for shallow-water flow and scalar transport with wetting and drying, J. Hydraul. Eng. 132 (4) (2006) 371–384.
[9] S.F. Bradford, B.F. Sanders, Finite-volume model for shallow-water flooding of arbitrary topography, J. Hydraul. Eng. 128 (3) (2002) 289–298.
[10] T.H. Yoon, S.-K. Kang, Finite volume model for two-dimensional shallow water flows on unstructured grids, J. Hydraul. Eng. 130 (7) (2004) 678–688.
[11] R. Peyret, Unsteady evolution of a horizontal jet in a stratified fluid, J. Fluid Mech. 78 (1) (1976) 49–63.
[12] C.L. Merkle, M. Athavale, Time-accurate unsteady incompressible flow algorithms based on artificial compressibility, in: The 8th AIAA Computational Fluid Dynamics Conference, AIAA-1987-1137, Honolulu, Hawaii, 1987.
[13] W.Y. Soh, J.W. Goodrich, Unsteady solution of incompressible Navier–Stokes equations, J. Comput. Phys. 79 (1988) 113–134.
[14] S. Rogers, D. Kwak, Upwind differencing scheme for the time-accurate incompressible Navier–Stokes equations, AIAA J. 28 (2) (1990) 253–262.
[15] J. Farmer, L. Martinelli, A. Jameson, Fast multigrid method for solving incompressible hydrodynamic problems with free surfaces, AIAA J. 32 (6) (1994) 1175–1182.
[16] M. Breuer, D. Hänel, A dual time-stepping method for 3-D, viscous, incompressible vortex flows, Comput. Fluids 22 (4/5) (1993) 467–484.
[17] C. Sheng, L.K. Taylor, D.L. Whitfield, Multigrid algorithm for three-dimensional incompressible high-reynolds number turbulent flows, AIAA J. 33 (11) (1995) 2073–2079.
[18] C.W. Gear, Numerical Initial Value Problems in Ordinary Differential Equations, Prentice-Hall, Englewood Cliffs, 1971.
[19] T.J.R. Hughes, M. Mallet, A new finite element formulation for computational fluid dynamics: III. the generalized streamline operator for multidimensional advective diffusive systems, Comput. Methods Appl. Mech. Eng. 58 (1986) 305–328.
[20] B.T. Helenbrook, A two-fluid spectral element method, Comput. Methods Appl. Mech. Eng. 191 (3–5) (2001) 273–294.
[21] D.J. Mavriplis, Multigrid techniques for unstructured meshes, Tech. Rep. 95-27, ICASE, April 1995.
[22] L. Martinelli, Calculations of viscous flow with a multigrid method, Ph.D. Thesis, Princeton University, Princeton, NJ, October 1987.
[23] B.T. Helenbrook, Preconditioning for incompressible flows with free-surfaces and two-fluid interfaces, J. Comput. Phys. 207 (1) (2005) 282–308.
[24] J. Boyd, Equatorial solitary waves, Part I: Rossby solitons, J. Phys. Oceanogr. 10 (1980) 1699–1717.