Conservation of Energy for Schemes Applied to the Propagation of Shallow-Water Inertia-Gravity Waves in Regions with varying Depth *

Terje O. Espelid
Department of Informatics
University of Bergen, Norway

Jarle Berntsen
Department of Mathematics
University of Bergen, Norway

Knut Barthel
Department of Geophysics
University of Bergen, Norway

SUMMARY

The linear equations governing the propagation of inertia-gravity waves in geophysical fluid flows are discretized on the Arakawa C-grid using centered differences in space. In contrast to the constant depth case it is demonstrated that varying depth may give rise to increasing energy (and loss of stability) using the natural approximations for the Coriolis terms found in many well known codes. This is true no matter which numerical method is used to propagate the equations. By a simple trick based on a modified weighting that ensures that the propagation matrices for the spatially discretized equations become similar to skew-symmetric matrices, this problem is removed and the energy is conserved in regions with varying depth too. We give a number of examples both of model problems and large scale problems in order to illustrate this behavior.

In real applications diffusion, explicit through frictional terms or implicit through numerical diffusion, is introduced both for physical reasons, but often also in order to stabilize the numerical experiments. The growing modes associated with varying depth, the C-grid and equal weighting may force us to enhance the diffusion more than we would like from physical considerations. The modified weighting offers a simple solution to this problem.

KEY WORDS: Shallow water equations; inertia-gravity waves; finite differences; C-grid; numerical stability; Coriolis; energy; skew-symmetric matrix.

*Published by: *Int. J. Numer. Engng 2000; 49:1521-1545*

1. INTRODUCTION

The propagation of inertia-gravity waves is a central issue in many geophysical fluid models. The reason for this is that large scale atmospheric and oceanic motions roughly obey the geostrophic equilibrium and that the dynamics of tides and storm surges are dominated by the propagation of external inertia-gravity waves.

It is well known that when propagating the solution of partial differential equations with numerical models, instabilities may often occur. In particular steep topography often in combination with nonlinearities may cause an energy cascade towards the shortest resolvable wavelengths of the computational grid, see Arakawa and Lamb [4] and Adcroft *et al.* [1].

The stability properties of the numerical methods are often studied with the Fourier or the von Neumann method. This method is relatively easy to apply to small subsystems of equations, but gives only necessary conditions for stability. The method is only applicable to linearized equations and constant parameters i.e. constant depth of the ocean. See for instance Grammeltvedt [11] , Mesinger and Arakawa [15], Schoenstadt [18], Foreman [9], Tanguay and Robert [20], Fox-Rabinowitz [10] or the recent study by Wang [21] where a theorem of roots for polynomials is applied to extend previous studies to more complex systems.

When applying the energy method, one requires that invariants associated with the continuous equations, total energy, vorticity and/or entrophy, are maintained in the discrete representation of the integral constraints. This method may be applied to nonlinear systems to produce sufficient conditions for stability. The task is to construct discrete spatial operators such that the contributions to the discrete invariants cancel when summarized over the model region. This is often done in the space domain, see Haltiner and Williams [12], Arakawa [2] and Arakawa and Lamb [4]. Lilly [14] also applied the technique to the vorticity equation studying the invariants in the discrete Fourier domain.

The conservative schemes produced by applying the energy method are rather complicated involving large computational stencils and may therefore be difficult to use in practise, see Navon [16] and Sasaki [17]. Simpler schemes like the leapfrog scheme or the forward-backward scheme, Haltiner and Williams [12], are thus still interesting for propagating the external inertia-gravity waves in ocean models (Slagstad *et al.* [19]; Berntsen [6]). To ensure stability some kind of filtering, in time and/or space, often by adding enough viscosity and diffusivity, is often required. The application of variational techniques to enforce conservation of invariants is also suggested (Navon [16]; Sasaki [17]).

Ocean models often define the variables on staggered grids. Many regional scale models apply the C-grid, Mesinger and Arakawa [15], because of its satisfactory

properties provided that the grid resolution is high relative to the deformation radius, see Mesinger and Arakawa [15], Fox-Rabinowitz [10], and/or Arakawa and Lamb [3].

Beckers and Deleersnijder [5] analyze the important forward-backward (explicit) scheme applied to such a problem on various grids. They use a von Neumann stability analysis (implying constant depth and free waves) to throw some new light on the bounds on the time step for the various grids. Adcroft *et al.* [1] address problems arising from the spatial averaging of the Coriolis terms when using the C-grid and suggest to augment the C-grid variables with D-grid velocity variables.

In a recent paper Espelid and Berntsen [8] focus on the stability using the C-grid centered differences in space and using different numerical time-stepping methods on problems with varying depth. They demonstrate that varying depth may imply that the stability is lost no matter how small time step one uses. We will in this paper show that it is possible to avoid this problem by a simple modification in the approximations. The method is based on studying the eigenvalues of the propagation matrices of the semi-discretized partial differential equations, discretization only in space. For purely hyperbolic problems the propagation matrices should have only purely imaginary eigenvalues. Skew-symmetric matrices and matrices similar to skew-symmetric matrices have this property. It is shown that with varying depth and the normal Coriolis term averaging used in C-grids the propagation matrices also get a symmetric part giving rise to growing and damping computational modes. A Coriolis weighting that ensures that the propagation matrix is similar to a skew-symmetric matrix is suggested.

In the next section we will present some of the results from [8], then we will present the necessary modifications and finally demonstrate both the problem of varying depth and the effect of the modification on a number of problems. First we follow [8] and limit the discussion to a model problem which has some general features.

2. A MODEL PROBLEM

Let f be the Coriolis parameter, g the gravity constant, U and V the depth integrated transports in x and y directions, respectively, η the sea surface elevation and $H(x, y)$ the undisturbed water depth. The external linearized inertia-gravity equations in Cartesian coordinates (x, y) may then be written

$$\begin{aligned} \frac{\partial U}{\partial t} &= -gH\frac{\partial \eta}{\partial x} + fV \\ \frac{\partial V}{\partial t} &= -gH\frac{\partial \eta}{\partial y} - fU \end{aligned} \tag{1}$$

$$\frac{\partial \eta}{\partial t} = -\frac{\partial U}{\partial x} - \frac{\partial V}{\partial y}.$$

In order to discuss the numerical stability we restrict ourselves to a bounded model region, Ω, with a boundary $\partial\Omega$ consisting of connected lines parallel to the two axes (we allow interior island(s) in Ω with a similar boundary form). Assume furthermore closed boundaries with no-flow boundary conditions; thus either $U = 0$ or $V = 0$ at each point on the boundary $\partial\Omega$ (this has to be true for the boundary of any island as well).

Define $u = U/H$ and $v = V/H$ everywhere in the interior of Ω and put $u = 0$ or $v = 0$ at each boundary point. The total energy of (1) over Ω, assuming uniform density ρ, may then be written

$$E(t) = \frac{1}{2}\rho \int\int_{\Omega} [H(u^2 + v^2) + g\eta^2] dxdy. \tag{2}$$

The total energy obviously has to be conserved.

2.1 THE METHOD OF LINES AND THE C-GRID

Assume that a uniform subdivision is possible in each spatial direction giving a grid Ω_Δ with spacing Δx and Δy in the two directions, possibly different. We will concentrate on the Arakawa C-grid, with η evaluated in the center of each grid cell, implying that only U and V are evaluated at the boundary grid points $\partial\Omega_\Delta$.

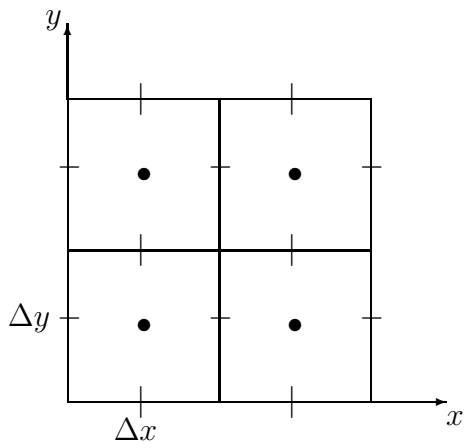


Figure 1: The Arakawa C-grid: spatial distribution of the variables (η, U, V): η is in the center of each grid cell: $\bullet$, U is at the horizontal bars on the boundary of each grid cell, and V is similarly at the vertical bars.

We follow [8] and approximate (1) in space only, using central differences. Define

$$\mathbf{w}(t) \;=\; \begin{pmatrix} U_\Delta(t) \\ V_\Delta(t) \\ \eta_\Delta(t) \end{pmatrix}$$

By $U_\Delta(t)$ we mean a vector of functions, each element of the vector defined at an interior U-grid point in Ω_Δ, similarly for $V_\Delta(t)$ and $\eta_\Delta(t)$. Let us now approximate (1) by second order space differences at time t using the vector $\mathbf{w}(t)$

$$\frac{d}{dt}\mathbf{w}(t) = (A+B)\mathbf{w}(t). \tag{3}$$

Here we have collected all terms containing the Coriolis parameter f in the matrix A and the rest of the elements in the matrix B. Both matrices are sparse. Thus we have to solve a linear system of ordinary differential equations, given the initial values $\mathbf{w}(0)$. Formally the solutions of (3) may be written $\mathbf{w}(t) = [\exp((A + B)t)]\mathbf{w}(0)$ and are approximate solutions (second order in space) to the true solution of (1).

The matrix A has elements 0 or $\pm f/4$ due the average over neighbor U- and V- points in the C-grid. Thus, the equation for a U-point component of $\mathbf{w}(t)$ will involve at most four neighbor V-point components of $\mathbf{w}(t)$ all with weight $f/4$. This U-point component gets weight $-f/4$ in the equation for the V-point component. This implies that the matrix A, independent of the numbering of the unknowns, always will be skew-symmetric, that is $A^T = -A$, where the T means the transpose of the matrix.

The matrix B may be written $B = B_x + B_y$, splitting the Δx and Δy terms. The equation for every U-component of $\mathbf{w}(t)$ connects to neighbor η-components, through matrix B_x, with weights $\pm gH/\Delta x$. The equation for every η-component connects two neighbor U-components with weights $\pm 1/\Delta x$. Due to these neighbor relations it turns out that all matrices B_x, B_y and B are "sign skew-symmetric", that is: $b_{ii} = 0$, for $\forall i$; for $\forall i \neq j$ then if $b_{ij} = 0$ then $b_{ji} = 0$ otherwise $b_{ij}b_{ji} < 0$. Furthermore, it is always possible to make all three matrices skew-symmetric by the same similarity transformation $\hat{B} = D^{-1}BD$ using a diagonal matrix D. Written in block form separating U_Δ, V_Δ and η_Δ (the diagonal block matrices are square) we have

$$D = \begin{pmatrix} D_U & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & D_V & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & I \end{pmatrix}.$$

The diagonal elements in D_U are $\sqrt{gH_u}$, where H_u is evaluated at the (x, y)-coordinates of the corresponding U-point. Similarly for D_V at the V-points.

Let $\mathbf{z}$ be an eigenvector to the matrix $A+B$ with eigenvalue μ. Putting $\mathbf{w}(0) = \mathbf{z}$ then we have a solution to (3)

$$\mathbf{w}(t) = \exp(\mu t)\mathbf{z}. \tag{4}$$

Thus the eigenvalues of the matrix $A + B$ will characterize the different solutions to (3). We observe that the real part of μ, $\Re\mu > 0$, implies a solution where all non-zero components of $\mathbf{z}$ will increase in absolute value with t, while $\Re\mu < 0$ implies decrease in absolute values. Either of these situations means a change in energy as t increases. To illustrate the effect on the energy we approximate the double integral in (2)

$$\frac{1}{2}\rho \int\int_\Omega [H(u^2 + v^2) + g\eta^2]dxdy \approx \frac{1}{2}\rho g \Delta x \Delta y ||D^{-1}\mathbf{w}(t)||_2^2 =$$
$$E_\Delta(t) = \frac{1}{2}\rho \Delta x \Delta y (\sum U^2/H_u + \sum V^2/H_v + g\sum \eta^2) \tag{5}$$

Here the three sums are taken over all wet U-points, V-points and η-points respectively. This approximation will be second order in space.

For the solution (4) we get

$$||D^{-1}\mathbf{w}(t)||_2^2 = \exp(2t\Re\mu)||D^{-1}\mathbf{z}||_2^2,$$

implying conservation of energy for solutions associated with eigenvalues with zero real part.

In order to find the eigenvalues of $A + B$ the following lemma is useful

Lemma 1 *Let S be a real skew-symmetric matrix, $S^T = -S$, then*
(a) all eigenvalues of S are on the imaginary axis,
(b) the matrix $I - S$ is non-singular and
(c) the matrix $(I - S)^{-1}(I + S)$ is orthogonal, implying $||(I - S)^{-1}(I + S)||_2 = 1$,
(d) the matrix $\exp(S)$ is a unitary matrix.

These are well known results.

The main results in [8] are given in the following theorem

Theorem 1 *The total energy*

$$E_\Delta = \frac{1}{2} g\rho\Delta x\Delta y||D^{-1}\mathbf{w}(t)||_2^2$$

of any solution $\mathbf{w}(t)$ of (3)
(a) will be conserved when $f = 0$ (the depth $H(x, y)$ may vary over Ω),
(b) will be conserved when $f \neq 0$ and the depth is constant over Ω.
(c) When $f \neq 0$ and the depth $H(x, y)$ varies over Ω, we may have eigenvalues μ whose real part is non-zero giving rise to solutions with both increasing and decreasing total energy.

Following the proof found in [8] let us define the following matrices $P = (D^{-1}AD - DAD^{-1})/2$ and $S = (D^{-1}AD + DAD^{-1})/2 + D^{-1}BD$. S is skew-symmetric while the matrix P is symmetric. Using S and P we may write $D^{-1}(A + B)D = S + P$ illustrating the fact that the matrix $A + B$ is not similar to a skew-symmetric matrix when the matrix P is not a zero matrix. $f = 0$ implies that both A and P are zero matrices while H constant implies that $D^{-1}AD = DAD^{-1} = A$ independent of f and therefore $P = 0$ implying (b). It is easy to give examples that prove case (c), e. g. see the Three-cell case in Section 3.

In practice it is of interest how large the real part of an eigenvalue to the matrix $A + B$ may be. In [8] an upper bound of the real part of the eigenvalues μ of the matrix $A + B$ is found by investigating the elements of the matrix P as

$$|\Re\mu_j| \leq ||P||_2 = \rho(P) \leq ||P||_\infty \leq |f|(k - 1/k)/2, \forall j, \tag{6}$$

where k is an upper bound on the neighbor depth ratio $\sqrt{H_N/H}$. Here H is the depth at any wet U point and H_N the depth at any neighboring V-points and vice versa. Furthermore $\rho(P)$ is the spectral radius of the matrix P and the norm $||P||_\infty = \max_i \sum_j |p_{ij}|$, where $P = \{p_{ij}\}$.

In order to understand the essence of Theorem 1 we may consider a method, based on the C-grid and second order space-differences, designed to solve (1). Assume furthermore that this method may be viewed as a proper approximation to (3) too. A consequence of Theorem 1, case (c) with $\Re\mu > 0$, is then that no such convergent method can conserve the energy when Δt becomes small. We will only consider methods having the property that they can be viewed as approximations to (3) as well as to (1). The forward-backward scheme, the leapfrog scheme and the Crank-Nicolson method are all methods having this property. These methods are frequently used in ocean codes.

We observe that the source to this lack of energy-conservation using the C-grid is the approximation of the Coriolis terms. In the next section we will modify this approximation used in (3) in order to conserve the energy.

2.2 A MODIFIED APPROXIMATION: CONSERVING THE ENERGY

In order to approximate the Coriolis terms in (1) we need an estimate of V in each U-point and vice versa. In (3) this is simply done as follows

$$\overline{V}_C = (V_{NE} + V_{NW} + V_{SW} + V_{SE})/4 \approx V_C, \tag{7}$$

where C is a wet U-point and NE, NW, SW and SE are the four neighbor V-points, some of these may be boundary points. This approximation is second order in space to the true value V_C. We may multiply and divide V_C by any positive space function $w(x, y)$ in order to create a different approximation to V_C, e.g.

$$\overline{V}_C = \frac{w_C}{4}\left(\frac{V_{NE}}{w_{NE}} + \frac{V_{NW}}{w_{NW}} + \frac{V_{SW}}{w_{SW}} + \frac{V_{SE}}{w_{SE}}\right) \approx V_C. \tag{8}$$

Using this approximation in the Coriolis terms in (1) we still have a second order approximation in space since (8) gives a second order approximation to V_C/w_C.

Choosing $w(x, y) = \sqrt{gH(x, y)}$ and replacing the standard approximation (7) with (8) modifies (3) as follows

$$\frac{d}{dt}\tilde{\mathbf{w}}(t) = (DAD^{-1} + B)\tilde{\mathbf{w}}(t). \tag{9}$$

From this we see that $D^{-1}\tilde{\mathbf{w}}(t)$ satisfies

$$\frac{d}{dt}D^{-1}\tilde{\mathbf{w}}(t) = (A + D^{-1}BD)D^{-1}\tilde{\mathbf{w}}(t), \tag{10}$$

and now the coefficient matrix $S = A + D^{-1}BD$ is skew symmetric and therefore the matrix $\exp(St)$ becomes unitary implying that $||D^{-1}\tilde{\mathbf{w}}(t)||_2 = ||D^{-1}\tilde{\mathbf{w}}(0)||_2 = ||D^{-1}\mathbf{w}(0)||_2$ and thus conservation of energy is achieved. We will use the terminology: unweighted run when we have used (7) (that is (8) with w constant) and weighted run when (8) is applied with $w(x,y) = \sqrt{gH(x,y)}$.

Applying the model to large local areas where the Coriolis parameter is varying too introduces an additional difficulty. In this case the matrix A is no longer skew-symmetric if $f(x,y)$ is evaluated in a straightforward manner: in the U-point in the moment-equation for U and similarly for V. However, by defining $w(x,y) = \sqrt{gH(x,y)/|f(x,y)|}$ in (8) we achieve that the system (10) can be written

$$\frac{d}{dt}D^{-1}\hat{\mathbf{w}}(t) = (\hat{A} + D^{-1}BD)D^{-1}\hat{\mathbf{w}}(t), \tag{11}$$

where $\hat{A}$ has elements (neglecting the sign) $\frac{1}{4}\sqrt{|f_C f_N|}$ and is skew-symmetric. Note that C denotes a U- or a V-point and that N denotes a neighbor point to this C-point.

In order to better understand the effect of a weighted approximation it is instructive to study how the Coriolis terms effect the numerical energy expression $E_\Delta(t)$, (5), over time. Let us differentiate $E_\Delta(t)$ with respect to time and use (3) to get

$$\frac{d}{dt}E_\Delta = \rho\Delta x\Delta y(\sum f_u\overline{V}U/H_u - \sum f_v\overline{U}V/H_v), \tag{12}$$

where the two sums are over all wet U-points and V-points respectively. Note that differentiating the exact energy (2) and using (1) gives a double integral over the region and two similar Coriolis terms cancelling each other pointwise: $fVU/H - fUV/H = 0$. The two approximations' ability to resemble this behavior locally is as follows: the unweighted approximation implies that for each pair of neighbor U and V points we will have two terms (one in each sum)

$$f_u\frac{1}{4}VU/H_u - f_v\frac{1}{4}UV/H_v. \tag{13}$$

These two terms will cancel if $f_u/H_u = f_v/H_v$, e. g. either $f = 0$ or both f and H are constant (Theorem 1), while otherwise we may experience increasing or decreasing energy. On the other hand the weighted approximation implies that each pair of neighbor U and V points will have two terms (one in each sum)

$$\sqrt{|f_u f_v|}\frac{1}{4}VU/\sqrt{H_u H_v} - \sqrt{|f_u f_v|}\frac{1}{4}UV/\sqrt{H_u H_v} = 0. \tag{14}$$

Thus the weighted approximation resembles the exact Coriolis effect in a perfect manner locally in contrast to the unweighted approximation. The above expression may be regarded as using the energy method to double check that the weighted approximation succeeds in conserving the energy. In retrospect this approach could have been used to develop the weighted approximation too. Due to the fact that any homogeneous linear system of ordinary differential equations will have invariant 2-norm if the matrix is skew-symmetric (Lemma 1 (c)) we consider the approach to re-establish the skew-symmetric property as a fruitful and simple idea and thus a natural path to follow in developing this approximation.

In some codes one will find that an unweighted approximation is using wet points only. Assuming that the point SE in (7) is a land point we then would get

$$\overline{V}_C = (V_{NE} + V_{NW} + V_{SW})/3 \approx V_C,$$

and a similar modification in the weighted case. We observe that this way of handling land points implies that one term will have a factor $1/4$ and the other term will have a factor $1/3$ in both (13) and (14). Thus this will in general ruin the conservation of energy property both in the new weighting case and in the unweighted, constant H, case.

It is important to notice that the weighted approach (9) guarantees that the exact solution of the system of ordinary differential equations (3) conserves the total energy of our model problem in contrast to the unweighted approximation in the varying depth case. This fact gives us the hope that the popular methods Crank-Nicolson, leapfrog and forward-backward all will benefit from this new weighting procedure.

Let us take a closer look at how Crank-Nicolson will perform on the model problem. Note that (10) and (11) are linear systems of ordinary differential equations where the matrix involved is skew-symmetric, say S. Crank-Nicolson applied to (3), using (9) is equivalent to using the trapezoidal method to (10). Given the approximation $D^{-1}\tilde{\mathbf{w}}(t)$ at time t. Using the trapeziodal method one time step from t to $t + \Delta t$, implies multiplying $D^{-1}\tilde{\mathbf{w}}(t)$ by the matrix

$$(I - 0.5\Delta t S)^{-1}(I + 0.5\Delta t S).$$

According to Lemma 1, case (c), the two norm of this matrix is 1 implying conserved energy. Thus we find that the Crank-Nicolson method will indeed conserve the energy when applied to the model problem with varying depth and the weighted approximation.

It is less obvious how the forward-backward method will perform and we will use that method in all experiments in the rest of the paper.

3. THE THREE-CELL CASE

The simplest possible C-grid ocean model involving both U, V and η points is presented in Figure 2.

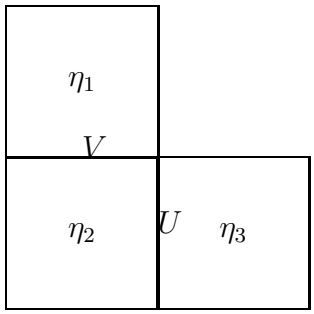


Figure 2: Variables in the Three-cell case.

The matrix-vector product $(A+B)\mathbf{w}(t)$ appearing in (3) becomes in this case

$$\begin{pmatrix} 0 & \frac{f}{4} & 0 & \frac{gH_u}{\Delta x} & -\frac{gH_u}{\Delta x} \\ -\frac{f}{4} & 0 & -\frac{gH_v}{\Delta y} & \frac{gH_v}{\Delta y} & 0 \\ 0 & \frac{1}{\Delta y} & 0 & 0 & 0 \\ -\frac{1}{\Delta x} & -\frac{1}{\Delta y} & 0 & 0 & 0 \\ \frac{1}{\Delta x} & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} U(t) \\ V(t) \\ \eta_1(t) \\ \eta_2(t) \\ \eta_3(t) \end{pmatrix}$$

The eigenvalues of $A+B$ depend on the parameters involved including the depths H_u and H_v measured in the U-point and the V-point respectively. Theorem 1 states that for $H_u = H_v$ there will be no growing/damping modes, but for $H_u \neq H_v$ some modes might experience an unphysical growth/damping. In general note that since the trace of the matrix $A+B$ is zero then the existence of damping modes always implies the existence of growing modes.

An eigenvalue μ of $A+B$ associated with a growing mode has the real part of μ, $\Re\mu$, greater than 0. In Figure 3 we plot the maximum real part of an eigenvalue as a function of H_u and keeping $H_v = 100m$ for $\Delta x = \Delta y = 20000m$. Espelid and Berntsen [8] give several upper bounds on $\Re\mu_j$ summarized in (6). In this Three-cell case the best upper bound is easily found as

$$|\Re\mu_j| \leq \rho(P) = |f|(k - 1/k)/8,$$

Figure 3 illustrates that this upper bound is, in this case, a severe overestimate

of the true value of the maximum real part.

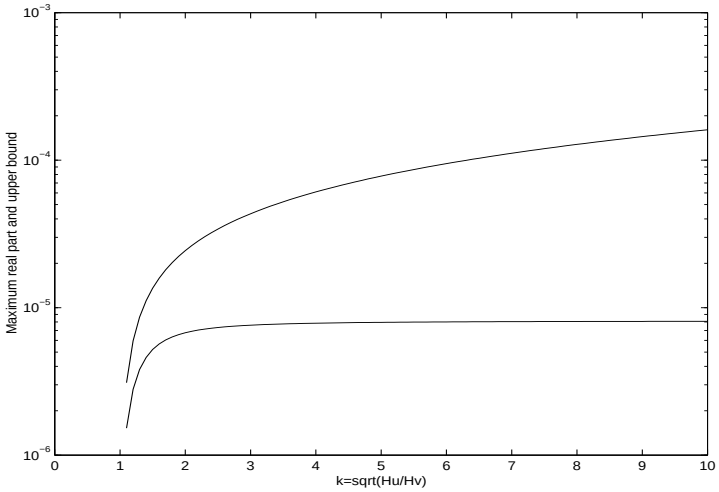


Figure 3: Maximum value of $\Re\ \mu$ in s^{-1} and its upper bound as a function of H_u.

In particular we note that in this case $\Re\mu$ approaches $.82 \times 10^{-5} s^{-1} \sim \frac{f}{16}$ as the neighbor depth ratio increases.

The propagation matrices have 5 eigenvectors/eigenvalues. One eigenvector represents the stable situation that $U = V = 0m^2s^{-1}$ and all water levels are equal. The corresponding eigenvalue is 0. Two eigenvectors appear in a pair that may be linearly combined to form the mode that represents oscillations in/out of the central cell. The corresponding eigenvalues appear in a complex pair. The real part of these eigenvalues are positive and this mode will be growing. The last two eigenvectors also appear in a pair and may be combined to form the pair that represents the fundamental oscillation in our model domain. Again the corresponding eigenvalues appear in complex pairs, but the real part of these eigenvalues are negative and this mode will therefore be damped.

For the case $H_u = 100m$ and $H_v = 200m$ we have chosen $\eta_2 = 1m$, $\eta_1 = \eta_3 = 0m$ and $U = V = 0m^2s^{-1}$ as initial values, initial values 1, and propagated the solution 150 hours in time with the forward-backward method and time step equal to 0.5s. The total energies are given in Figure 4.

Keeping the same depths we have tried $\eta_1 = \frac{2}{\sqrt{5}}m$, $\eta_2 = \frac{1}{\sqrt{5}}m$, $\eta_3 = 0m$ and $U = V = 0m^2s^{-1}$ as initial values, initial values 2, and propagated the solution 150 hours in time with the forward-backward method and time step equal to 0.5s. The total energies are given in Figure 4. Initially the total energy is $33.44Jm^{-3}$ for both initial values. For initial values 1 the total energy is $3309.74Jm^{-3}$ after 150 hours with an unweighted run. In the first experiment the initial values represent

the growing mode and the expected increase in energy occurs already at the start of the simulation. In the second experiment the initial values represent the damping mode and on a short time scale this is dominating and damping before the growing mode eventually builds up.

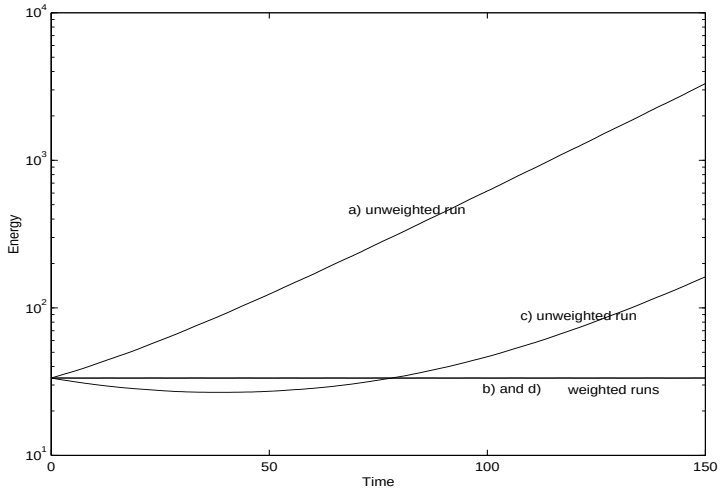


Figure 4: Total energy in Jm^{-3} for a) Initial values 1 and unweighted run b) Initial values 1 and weighted run, c) Initial values 2 and unweighted run and d) Initial values 2 and weighted run as functions of time in hours.

The modification given in the previous section implies that the matrix vector product in (10), $(A + D^{-1}BD)D^{-1}\tilde{\mathbf{w}}(t)$, becomes

$$\begin{pmatrix} 0 & \frac{f}{4} & 0 & \frac{\sqrt{gH_u}}{\Delta x} & -\frac{\sqrt{gH_u}}{\Delta x} \\ -\frac{f}{4} & 0 & -\frac{\sqrt{gH_v}}{\Delta y} & \frac{\sqrt{gH_v}}{\Delta y} & 0 \\ 0 & \frac{\sqrt{gH_v}}{\Delta y} & 0 & 0 & 0 \\ -\frac{\sqrt{gH_u}}{\Delta x} & -\frac{\sqrt{gH_v}}{\Delta y} & 0 & 0 & 0 \\ \frac{\sqrt{gH_u}}{\Delta x} & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \tilde{U}(t)/\sqrt{gH_u} \\ \tilde{V}(t)/\sqrt{gH_v} \\ \tilde{\eta}_1(t) \\ \tilde{\eta}_2(t) \\ \tilde{\eta}_3(t) \end{pmatrix}$$

3.1 BALANCING THE GROWTH WITH VISCOSITY

From Figure 3 we see that for $H_v = 100m$ and H_u larger than approximately $200m$ or smaller than $50m$, $\Re\mu$ will be greater than $0.5 \times 10^{-5}s^{-1}$. Such depth ratios are often appearing in many model areas and one may ask if such a growth of some eigenmodes is something to worry about. $\Re\mu = 0.5 \times 10^{-5}s^{-1}$ means that

this eigenmode will double its energy approximately every 20 hour.

When running 3-D models with more complete sets of equations, we often enhance the viscosity to get stable runs. We therefore consider to balance the growth above with a horizontal viscosity term modifying (1) as follows

$$\begin{aligned}
\frac{\partial U}{\partial t} &= -gH\frac{\partial \eta}{\partial x} + fV + A_H(\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2}) \\
\frac{\partial V}{\partial t} &= -gH\frac{\partial \eta}{\partial y} - fU + A_H(\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2}). \\
\frac{\partial \eta}{\partial t} &= -\frac{\partial U}{\partial x} - \frac{\partial V}{\partial y}
\end{aligned} \tag{15}$$

Assuming $U = 0$ at the boundary grid points and on land, the new spatial operators above are approximated by $A_H(-2U/\Delta x^2 - 2U/\Delta y^2)$ in this Three-cell case. This introduces the terms $-2A_H(1/\Delta x^2 + 1/\Delta y^2)$ in the diagonal elements in $(A+B)$ associated with the $U-$ and $V-$block respectively. For the case $H_u = 100m$ and $H_v = 200m$ and with initial values 1 we have run the model 150 hours and the total energies at the end of the computations are given as function of A_H in Figure 5. In order to avoid growth in energy A_H must be chosen greater than $900m^2s^{-1}$. This will represent a lower limit on the horizontal viscosity we may apply unless other frictional terms are introduced and for some applications this may be a larger value than we from physical considerations may want to apply. The energy after 150 hours in a weighted run with $A_H = 900m^2s^{-1}$ is $11.32Jm^{-3}$ as opposed to $36.68Jm^{-3}$ in the unweighted run.

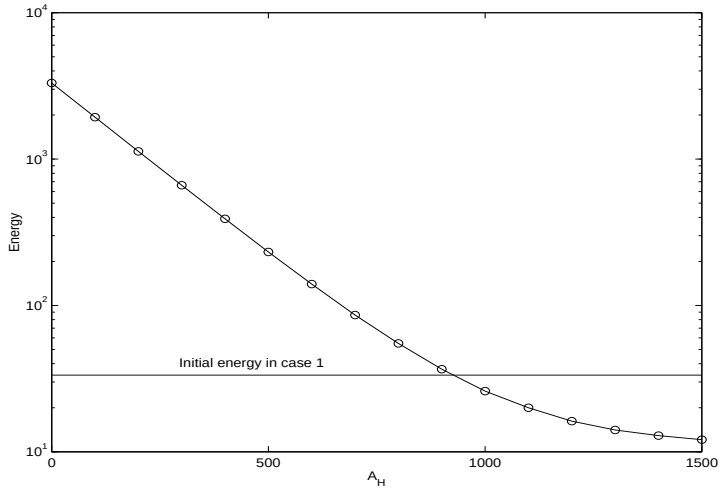


Figure 5: Total energy in Jm^{-3} after 150 hours as a function of A_H in m^2s^{-1} with initial values 1 and unweighted runs.

3.2 BALANCING THE GROWTH WITH BOTTOM FRICTION

Alternatively one may introduce bottom friction to balance the growth in energy. We therefore consider to balance the growth above with bottom friction terms

$$\begin{aligned}
\frac{\partial U}{\partial t} &= -gH\frac{\partial \eta}{\partial x} + fV - r\sqrt{U^2+V^2}\,U/H_u^2, \\
\frac{\partial V}{\partial t} &= -gH\frac{\partial \eta}{\partial y} - fU - r\sqrt{U^2+V^2}\,V/H_v^2, \\
\frac{\partial \eta}{\partial t} &= -\frac{\partial U}{\partial x} - \frac{\partial V}{\partial y},
\end{aligned}$$

where r is a bottom drag coefficient. This coefficient is usually taken in the range 2×10^{-3} to 4×10^{-3}, see Kowalik and Murty [13]. For the case $H_u = 100m$ and $H_v = 200m$ and with initial values 1 we have run the model 150 hours and the total energies at the end of the computations are given as function of r in Figure 6. In order to avoid growth in energy r must be chosen greater than 0.011 which is much larger than the values normally used in ocean models. In all experiments in this paper we have used the same approximation of V in a U-point under the square root in the friction term as for the Coriolis term and vice versa for U in a V-point.

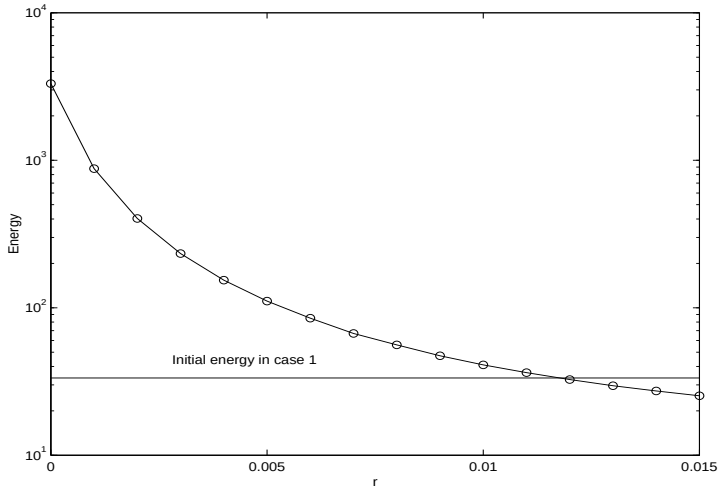


Figure 6: Total energy in Jm^{-3} after 150 hours as a function of r with initial values 1 and unweighted runs.

4. THE NORTH SEA EXPERIMENT

To investigate whether the growing modes associated with the standard Coriolis averaging may be a problem in more realistic applications an experiment for an

extended North Sea discretized in a 20 × 20 km grid with totally 80 × 70 grid cells has been performed. The bottom matrix is given in Figure 7.

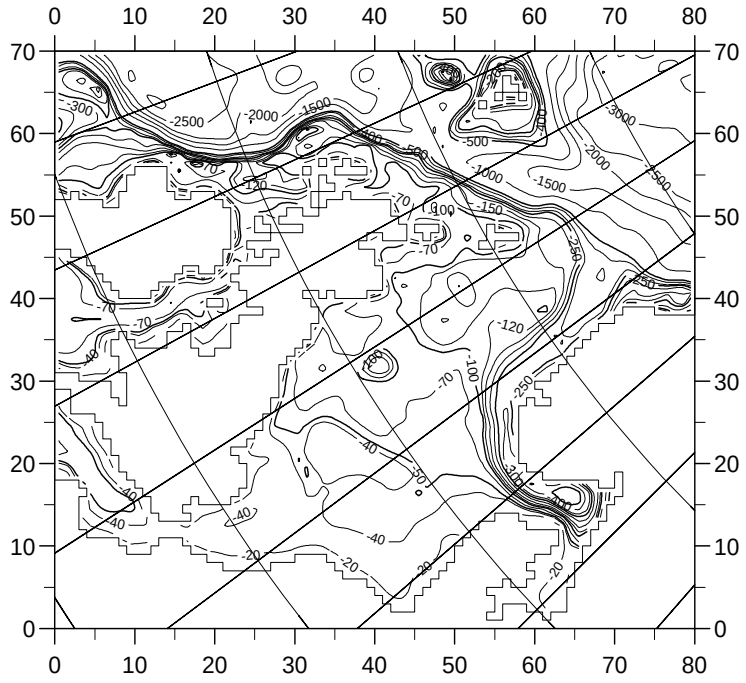


Figure 7: Bottom topography of the North Sea.

The depths for this model area are given at the cell centers(η points) and they are averaged to produce the depths in U and V points. In the numerical experiment all border cells are defined as land cells to avoid open boundary problems. Initial values of η, U and V are taken from Engedahl *et al.* [7]. The initial values for η are given in Figure 8.

The basis for this experiment is (1) with no forcing and no fluxes at the boundaries. The solution is propagated in time 600 hours with the forward-backward technique and 60s time step using both the unweighted approximation and the weighted approximation. In Figure 9 the total energies in Jm^{-3} are given for the two methods. The average growth rate for the simulation is close to $1.2 \times 10^{-6} s^{-1}$.

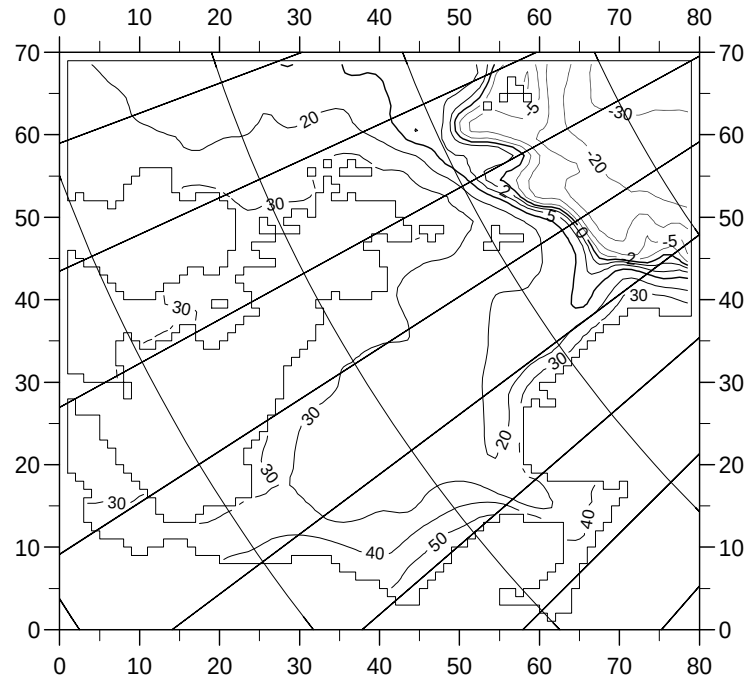


Figure 8: Initial values of η in cm.

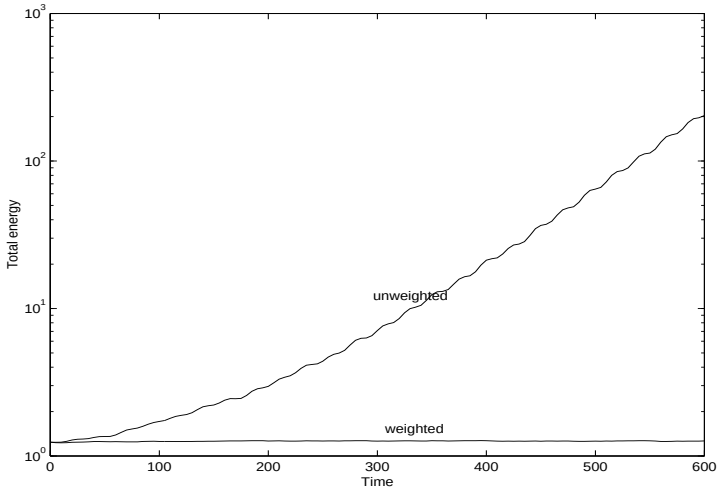


Figure 9: Total energy in Jm^{-3} as a function of time in hours for weighted and unweighted runs.

As for the Three-cell case one may consider to add viscosity terms to the equations of the form (15) and in Table 1 the total energies in Jm^{-3} after 600 hours

of simulation for the two methods are given as functions of the viscosity A_H in m^2s^{-1}. We notice that for this more realistic model problem it suffices to apply $A_H = 400m^2s^{-1}$ to balance the growth when using the unweighted approximation. We also notice that the solutions at the end of the simulations are less energetic when applying the weighted approximation for all values of A_H even if the differences are reduced as we increase A_H.

A_H	Energy unweighted case	Energy weighted case
0.	203.563	1.265
100.	33.357	0.593
200.	6.029	0.485
300.	1.425	0.445
400.	0.609	0.423
500.	0.449	0.409
600.	0.408	0.397
700.	0.392	0.389

Table 1. Total energies in Jm^{-3} after 600 hours of simulation using weighted and unweighted Coriolis as functions of the viscosity A_H in m^2s^{-1}.

In Figure 10 the η values after 600 hours simulation using the unweighted Coriolis approximation and $A_H = 200m^2s^{-1}$ are given and the corresponding values in the weighted case are given in Figure 11. Notice especially that the small scale noise in the Skagerrak is no longer present in the results produced in the weighted case. The instabilities present in Figure 10 seem to be connected to the steep topography in the Skagerrak and along the Norwegian trench. Figures 10 and 11 also suggest that it may be possible to run ocean models with smaller values of viscosity and thus allow the presence of more physical phenomena with the weighted approach.

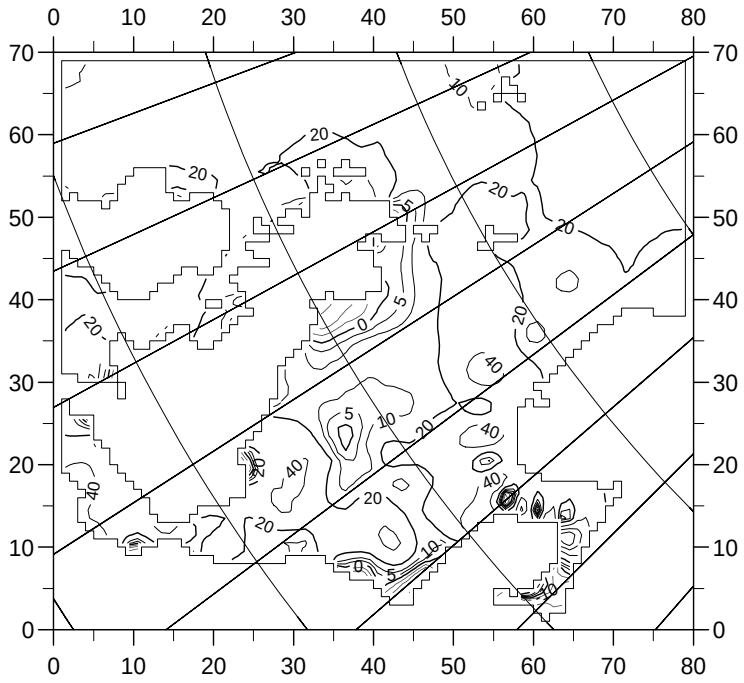


Figure 10: Values of η in cm after 600 hours of simulation using the unweighted Coriolis approximation and $A_H = 200m^2s^{-1}$.

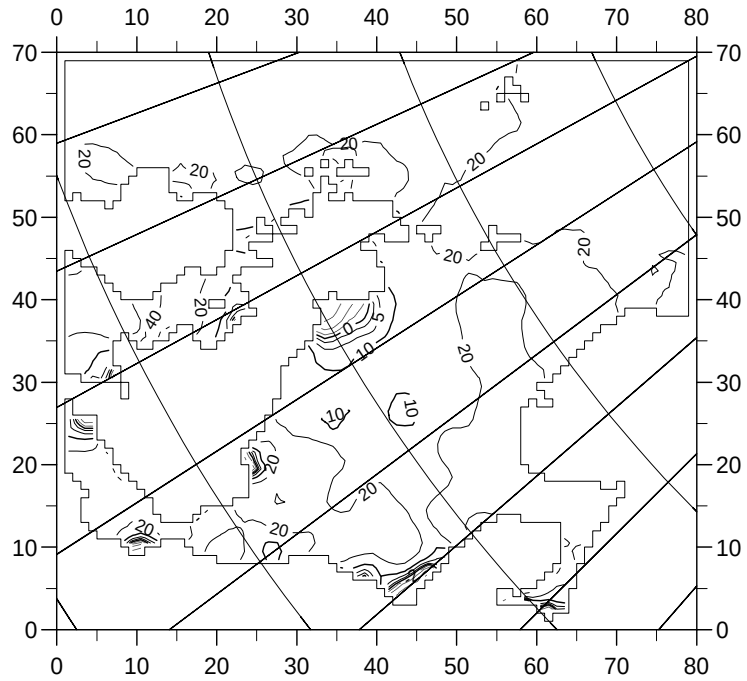


Figure 11: Values of η in cm after 600 hours of simulation using the weighted Coriolis approximation and $A_H = 200m^2s^{-1}$.

Alternatively bottom friction terms may be added to the problem and in Table 2 the total energies in Jm^{-3} after 600 hours of simulation for the two methods are given as functions of the bottom friction parameter r.

r	Energy unweighted	Energy weighted
0.000	203.563	1.265
0.001	0.729	0.452
0.002	0.531	0.402
0.003	0.464	0.382
0.004	0.431	0.372
0.005	0.411	0.365
0.006	0.397	0.360

Table 2. Total energies in Jm^{-3} after 600 hours of simulation after an unweighted run and a weighted run as functions of the bottom friction parameter r.

It suffices to apply $r = 0.001$ to balance the growth when using the unweighted Coriolis approximation which is a smaller value than normally is applied. However, the solutions at the end of the simulations are less energetic in the weighted case for all values of r even if the differences are smaller for larger values of r.

5. THE BIG LAKE EXPERIMENT

In order to investigate how the varying depth problem is experienced in a case where both wind forcing and bottom friction are included we have studied a regularly shaped closed sea (a big lake).

Modifying the depth integrated linearized shallow water equations (1) to include bottom friction and constant wind forcing we may write

$$\begin{aligned} \frac{\partial U}{\partial t} &= -gH\frac{\partial \eta}{\partial x} + fV + \rho^{-1}(\tau_x^{(0)} - \tau_x^{(B)}), \\ \frac{\partial V}{\partial t} &= -gH\frac{\partial \eta}{\partial y} - fU + \rho^{-1}(\tau_y^{(0)} - \tau_y^{(B)}), \\ \frac{\partial \eta}{\partial t} &= -\frac{\partial U}{\partial x} - \frac{\partial V}{\partial y}. \end{aligned} \tag{16}$$

The stress terms are non-linear, and given by

$$\rho^{-1}(\tau_x^{(0)}, \tau_y^{(0)}) = \lambda\sqrt{W_x^2 + W_y^2}(W_x, W_y), \tag{17}$$

$$\rho^{-1}(\tau_x^{(B)}, \tau_y^{(B)}) = r\sqrt{U^2 + V^2}(U, V)/H^2. \tag{18}$$

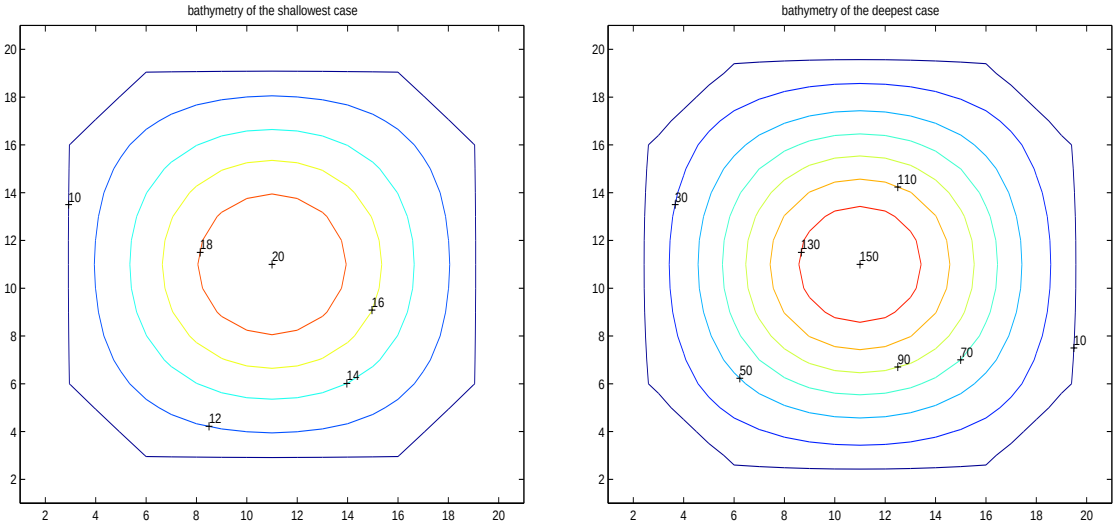
W_x, W_y are the wind velocity components, ms^{-1}, in the x- and y-direction, respectively. λ is equal to the product of the drag coefficient and the ratio between the densities of air and water. λ is set to 3.2×10^{-6}. r is a bottom drag coefficient, which is set equal to 3.0×10^{-3}. ρ is the density, $\tau_{(x,y)}^{(0)}$ is the wind stress and $\tau_{(x,y)}^{(B)}$ is the bottom stress.

The depth integrated shallow water equations with bottom friction and constant wind forcing are discretized, and integrated from a basic state at rest using the forward-backward method.

We let the maximum depth vary from $20m$ to $150m$. We choose $\Delta t = 3$ minutes for all runs, so that the CFL-criteria is fulfilled for a maximum depth of $150m$. The sea has a closed boundary, and the transport normal to the boundary must vanish.

After forty days we have stationary current and surface displacement fields. In the stationary state the left hand sides of (16) vanish. This gives a divergence free transport and a surface displacement field, where the horizontal pressure gradient is balancing the stress terms and the Coriolis force.

The lake bathymetry is given in the figures 12a and 12b for the shallowest and deepest case, respectively. The area outside the $10m$ contour is land (Figure 12a). The unit along axes and the grid resolution are $10km$.



(a) shallowest case

(b) deepest case

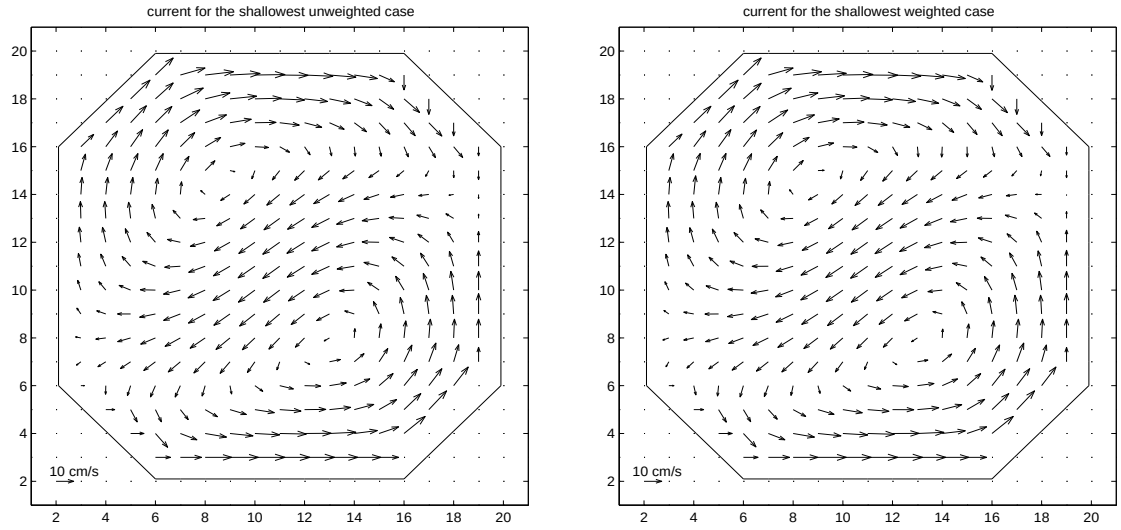
Figure 12: Bottom topography of the Big Lake. The unit along both axes is $10km$

The two outermost grid-boxes around the sea are land-boxes where $H = 0$. Six boxes arranged symmetrically at each corner are also land-boxes. Else the topography is given by the formula

$$H_{k,l} = 10 + (H_{\max} - 10)(\cos(\pi(k-11)/20)cos(\pi(l-11)/20))^2$$

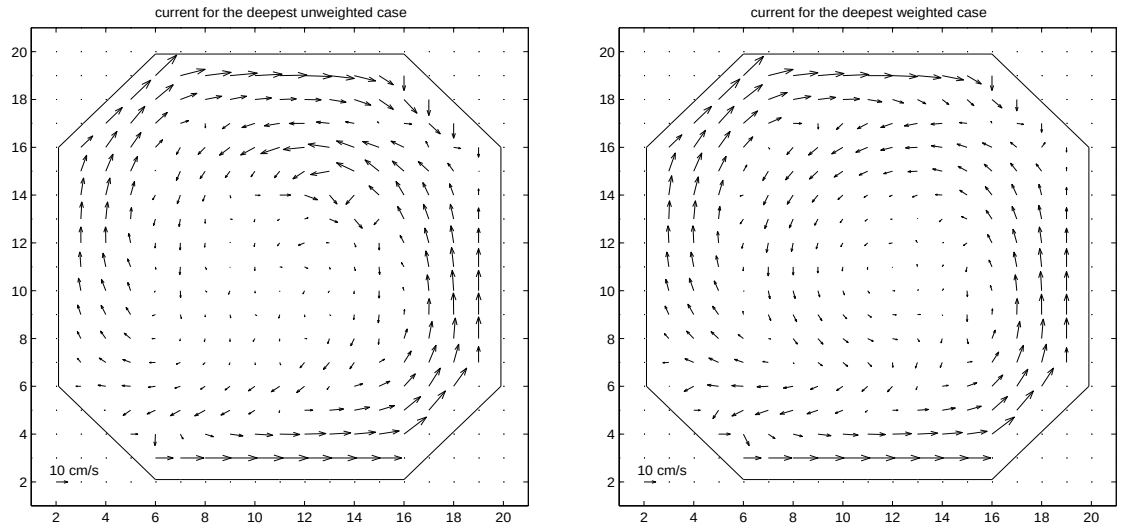
In the experiment the maximum depth $H_{\max}$ is varied from $20m$ to $150m$ in steps of $10m$. Thus in the shallowest case where the undisturbed depth sloped from $10.48m$ at the bays to $20m$ at midsea, we ran the model first with an unweighted Coriolis approximation and thereafter with weighting. The wind was westerly $10ms^{-1}$, that is $W_x = 10ms^{-1}$, $W_y = 0ms^{-1}$. Figures 13a and b show the resulting current fields (in cms^{-1}). It is hard to see any differences between the two runs for this shallow case. Neither are there any differences in the surface

displacement (not shown)



(a) unweighted (b) weighted

Figure 13: Current for the shallowest case.



(a) unweighted (b) weighted

Figure 14: Current for the deepest case.

The runs were repeated thirteen times. The current plots of the deepest case are shown in the figures 14a and b. Differences are clearly seen, for example around

point $(k, l) = (12, 13)$. The stronger slope (from $16.7m$ at the bays to $150m$ midsea) makes the effect of the weighting stronger.

The surface displacement is unaffected (not shown). The maximum current difference between unweighted and weighted runs as a function of maximum depth is plotted in Figure 15. The effect of an increased slope is evident. The maximum difference increases from only $2cms^{-1}$ in the shallowest case to $14cms^{-1}$ in the deepest case. The currents are slower in the deepest case (compare figures 14 and 15), and $14cms^{-1}$ is a pronounced effect here.

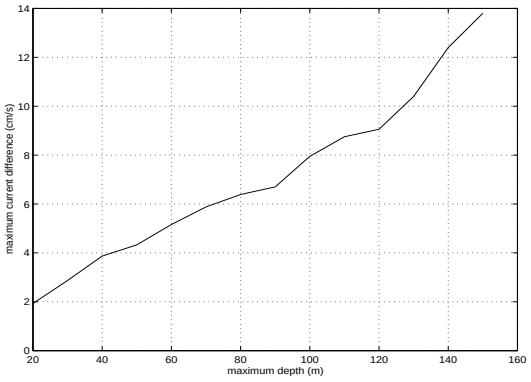


Figure 15: The maximum current difference between weighted and unweighted runs.

Next we will study the energy of the model system. The total kinetic energy of the sea is given by

$$K = \frac{1}{2}\rho\Delta x\Delta y \sum_{k,l}(U_{k,l}^2/H_{k,l} + V_{k,l}^2/H_{k,l})$$

The results of the fourteen different cases are plotted in Figure 16. The total kinetic energy increases with maximum depth. This increase is weaker, when weighting is applied than if it is not applied. The difference is plotted in Figure 17 as a function of depth. It increases weakly from $0.08 \times 10^{11}J$ for $20m$ (less than 5%) to $1.81 \times 10^{11}J$ for $80m$ (3.6%), then more and more strongly up to $8.13 \times 10^{11}J$ for $150m$ (11.2%).

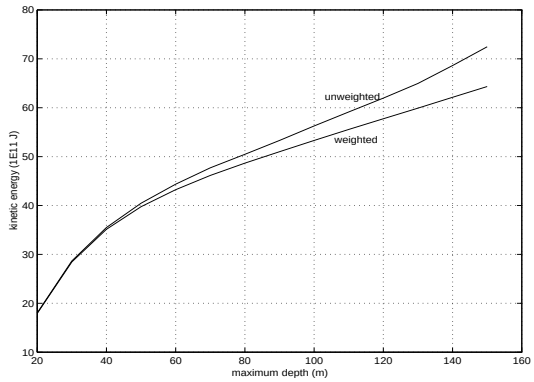


Figure 16: Total kinetic energy as a function of maximum depth. The unweighted runs gave the upper curve, while the weighted runs gave the lower curve.

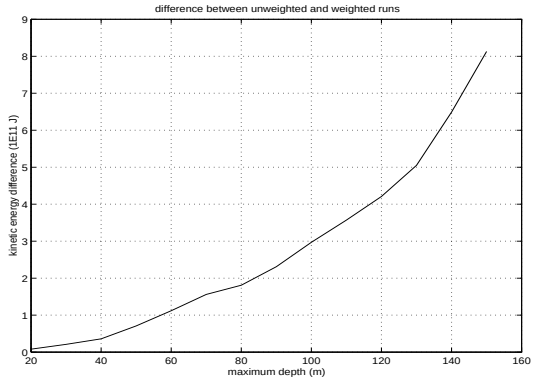


Figure 17: Total kinetic energy difference between unweighted and weighted runs as a function of maximum depth.

The total potential energy of the ocean is given by

$$P = \frac{1}{2}\rho g \Delta x \Delta y \sum_{k,l} \eta_{k,l}^2$$

It decreases with increasing maximum depth from about $13.2 \times 10^{11} J$ for $20m$ to about $1.1 \times 10^{11} J$ for $150m$(not shown), and the effect of weighting is very small, see Figure 18. It is interesting that the weighting gives an increase in potential energy, whereas it gave a decrease in kinetic energy.

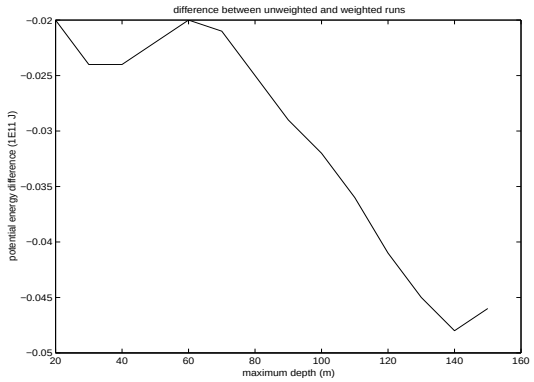


Figure 18: Total potential energy difference between unweighted and weighted runs as a function of maximum depth.

The experiments with the Big Lake have rather modest $k = \max \sqrt{H_u/H_v}$ values close to 1 , see Figure 3. To see what happens when the k value is larger than 2, we made an experiment with the shallowest case where two $100m$ deep troughs along i=11 and along j=11 were added giving $k \approx \sqrt{10}$, see Figure 19 .

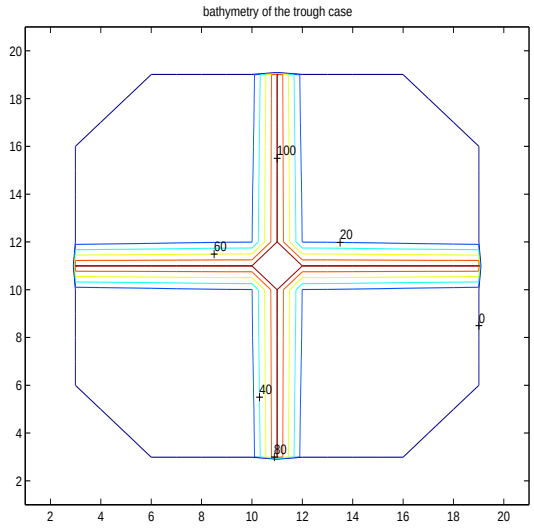
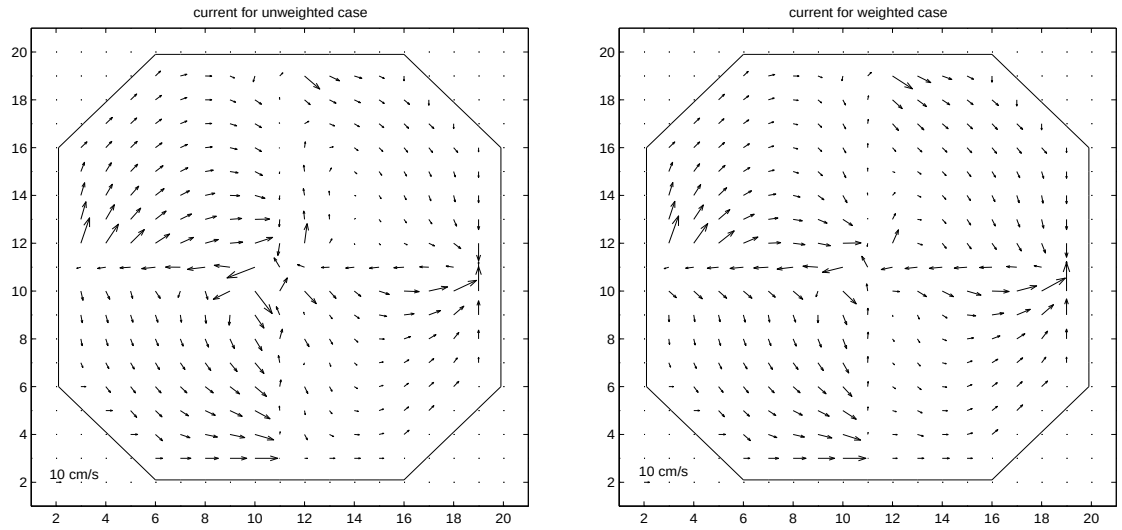


Figure 19: Bottom topography with deep troughs in the Big Lake.



(a) unweighted (b) weighted

Figure 20: Currents with deep troughs in the Big Lake.

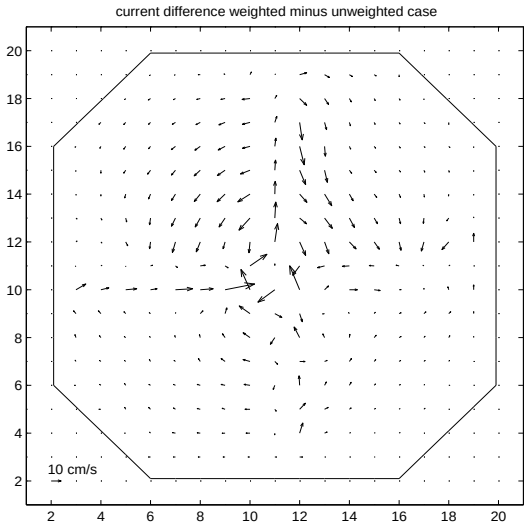


Figure 21: The maximum current difference between weighted and unweighted runs.

The results show much larger differences between unweighted and weighted runs. The total kinetic energy changed from $133.3 \times 10^{11} J$ in the unweighted run to around $105 \times 10^{11} J$ in the weighted, which is a substantial 20% change. The

largest current change was about $30cms^{-1}$, which is also substantial, see figures 20 and 21.

6. CONCLUSIONS

Linear analysis shows that in the constant depth case the Arakawa C-grid may be an attractive choice for spatial discretization when solving the shallow water equations. In the present paper the varying depth case is analyzed and it is shown that the unweighted approximation of the Coriolis terms will give rise to growing modes and instability unless frictional terms are added to the equations. A new weighted approximation that guarantees no growth or damping for the purely hyperbolic case is suggested. These two approaches are compared for a simple Three-cell case, an enclosed North Sea and for a wind driven Big Lake. In the experiments it is demonstrated that in the unweighted case we may get unphysical growth in energy and artificial flows.

In real applications diffusion, explicit through frictional terms or implicit through numerical diffusion, is introduced both for physical reasons, but often also in order to stabilize the numerical experiments. The growing modes associated with varying depth and the C-grid using the unweighted approximation may force us to enhance the diffusion more than we would like from physical considerations. The suggested weighting procedure offers a simple solution to this problem.

The method used to develop the energy conservative spatial discretization is based on producing propagation matrices for the semi-discretized partial differential equations that are similar to skew-symmetric matrices. The method may be applied to equations with non-constant parameters as in this case. To extend the analysis to 3-D equations including only the same terms as in the present study, neglecting bottom friction and viscosity, is straightforward if the prognostic equations for the momentum are written on transport form, $\partial(uH)/\partial t$, where u is the velocity. This is not surprising for the problem with the present terms is basically 2-D. If the moment equations are given for the velocities, $\partial u/\partial t$, a slight revision is required. The similarity matrix D will get terms of the type $\sqrt{g/H_u}$ on the diagonal and the Coriolis weighting has to be adjusted accordingly.

We also hope to extend the study by including non-linear terms. The task will be to produce consistent approximations to these terms that also give propagation matrices that are similar to skew-symmetric matrices.

ACKNOWLEDGEMENTS

The authors thank three anonymous reviewers for constructive comments, which lead to a substantial improvement of the paper.

References

[1] A.J. Adcroft, C.N. Hill, and J.C. Marshall. A New Treatment of the Coriolis Terms in C-Grid Models at Both High and Low Resolutions. *Mon. Wea. Rev.*, 127:1928–1936, 1999.

[2] A. Arakawa. Computational design for long-term numerical integration of the equations of fluid motion: Two-dimensional incompressible flow. Part I. *J. Comput. Phys.*, 1:119–143, 1966.

[3] A. Arakawa and V.R. Lamb. Computational design of the basic dynamical processes of the UCLA general circulation model. In *Methods in Computational Physics. Vol. 17. General Circulation Models of the Atmosphere.* Academic-Press, 1977.

[4] A. Arakawa and V.R. Lamb. A Potential Enstrophy and Energy Conserving Scheme for the Shallow Water Equations. *Mon. Wea. Rev.*, 109:18–36, 1981.

[5] J.-M. Beckers and E. Deleersnijder. Stability of a FBTCS scheme applied to the propagation of shallow water inertia-gravity waves on various space grids. *J.Comp.Phys.*, 108:95–104, 1993.

[6] J. Berntsen. A sensitivity study of a baroclinic model for the North Sea with focus on the Utsira-Orkneys transect. Technical Report Fisken og Havet Nr. 16, Institute of Marine Research, 1998.

[7] H. Engedahl, B. Ådlandsvik, and E.A. Martinsen. Production of monthly mean climatological archives of salinity, temperature, current and sea level for the Nordic Seas. *J. Mar. Syst.*, 14:1–26, 1998.

[8] T. O. Espelid and J. Berntsen. Stability of different schemes applied to the propagation of shallow-water inertia-gravity waves in regions with varying depth. Technical Report 136, Department of Informatics, University of Bergen, Norway, 1997.

[9] M.G.G. Foreman. A two-dimensional dispersion analysis of selected methods for solving the linearized shallow water equations. *J.Comp.Phys.*, 56:287–323, 1984.

[10] M.S. Fox-Rabinowitz. Computational Dispersion Properties of Horizontal Staggered Grids for Atmospheric and Ocean Models. *Mon. Wea. Rev.*, 119:1624–1639, 1991.

[11] A. Grammeltvedt. A survey of finite-difference schemes for the primitive equations for a barotropic fluid. *Mon. Wea. Rev.*, 97:384–404, 1969.

[12] G.J. Haltiner and R.T. Williams. *Numerical prediction and dynamic meteorology*. John Wiley, 1980.

[13] Z. Kowalik and T.S. Murty. *Numerical Modeling of Ocean Dynamics*, volume 5 of *Advanced Series on Ocean Engineering*. World Scientific, 1993. ISBN-981-02-1333-6.

[14] D.K. Lilly. On the computational stability of numerical solutions of time-dependent non-linear geophysical fluid dynamics problems. *Mon. Wea. Rev.*, 93:11–26, 1965.

[15] F. Mesinger and A. Arakawa. Numerical methods used in atmospheric models, Volume I, 1976. WMO/ICSU Joint Organizing Committee, Garp Publication Series No. 17.

[16] I.M. Navon. Implementation of a posteriori methods for enforcing conservation of potential enstrophy and mass in discretized shallow-water equations models. *Mon. Wea. Rev.*, 109:946–958, 1981.

[17] Y. Sasaki. Variational design of finite-difference schemes for initial value problems with an internal invariant. *J.Comput.Phys.*, 21:270–278, 1976.

[18] A.L. Schoenstadt. The effect of spatial discretization on the steady state and transient behavior of a dispersive wave equation. *J.Comp.Phys.*, 23:364–379, 1977.

[19] D. Slagstad, K. Støle-Hansen, and H. Loeng. Density driven currents in the Barents Sea calculated by a numerical model. *Modeling, Identification and Control*, 11:181–190, 1991.

[20] M. Tanguay and A. Robert. Elimination of the Helmholtz Equation Associated with the Semi-Implicit Scheme in a Grid Point Model of the Shallow Water Equations. *Mon. Wea. Rev.*, 114:2154–2162, 1986.

[21] J. Wang. Global Linear Stability of the Two-Dimensional Shallow-Water Equations: An Application of the Distributive Theorem of Roots for Polynomials on the Unit Circle. *Mon. Wea. Rev.*, 124:1301–1310, 1996.