INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN FLUIDS
Int. J. Numer. Meth. Fluids 2003; **42**:607–622 (DOI: 10.1002/fld.528)

Dispersion analysis of the least-squares finite-element shallow-water system

D. Y. Le Roux[1,*,†] and G. F. Carey[2]

[1] *Département de Mathématiques et de Statistique, Université Laval, Québec, QC, Canada G1K 7P4*
[2] *Department of Aerospace Engineering Mechanics, University of Texas at Austin, Austin, TX 78712, U.S.A.*

SUMMARY

The frequency or dispersion relation for the least-squares mixed formulation of the shallow-water equations is analysed. We consider the use of different approximation spaces corresponding to co-located and staggered meshes, respectively. The study includes the effect of Coriolis, and the dispersion properties are compared analytically and graphically with those of the mixed Galerkin formulation. Numerical solutions of a test problem to simulate slow Rossby modes illustrate the theoretical results. Copyright © 2003 John Wiley & Sons, Ltd.

KEY WORDS: least squares; finite elements; shallow-water equations; gravity and Rossby waves

1. INTRODUCTION

The finite-element method is widely applied to shallow-water simulation for environmental applications as well as other flow problems where the shallow-water long wavelength assumption applies [1–5]. The most popular schemes are based on the scalar second-order wave equation formulation with numerical or artificial dissipation [6]. There have also been studies and applications using the mixed Galerkin formulation applied directly to the corresponding first-order system. One of the key issues that arises, especially when dealing with the mixed method, is the possible presence of spurious modes associated with folding of the dispersion relation. This question is also intimately tied to the question of consistency of the approximation spaces and the use of either co-located or staggered meshes in defining these spaces for the respective velocity and elevation variables. The dispersion relation for the mixed Galerkin method was first studied by Walters and Carey [7] who examined the folding of the relation and the behaviour of spurious modes for certain choices of basis on co-located

* Correspondence to: D. Le Roux, Département de Mathématiques et de Statistique, Université Laval, Québec, QC, Canada G1K 7P4.
† E-mail: Daniel.Leroux@mat.ulaval.ca

Received 24 September 2001
Revised 30 July 2002

meshes i.e. on meshes that are not 'staggered'. Subsequently, there have been several studies of the behaviour of mixed Galerkin methods in both one and two dimensions [3, 8–11].

While most finite-element methods (mixed or otherwise) are based on some variant of the Galerkin method, other weighted residual type finite-element formulations such as finite-element residual collocation and least-squares residual finite-element schemes are of interest [12]. The least-squares approach involves functional minimization and leads to symmetric positive algebraic systems. Moreover, this formulation is not subject to the usual consistency condition that is required of the mixed Galerkin saddle point problem (the LBB or inf sup condition) [13–16]. Finally, the least-squares finite-element scheme is equivalent to a Petrov–Galerkin scheme involving the adjoint operator and is 'naturally' dissipative with dissipation proportional to time step Δt in the time-discretized implicit formulation [17]. Hence, it is logical to explore the dispersion relation and properties of the least-squares mixed shallow-water problem and that is the main objective of the present study.

The paper is organized as follows. The model equations and the time-discretization scheme are presented in Sections 2 and 3, respectively. The least-squares mixed formulation follows in Section 4. The finite-element spaces for co-located and staggered grids are introduced and representative nodal equations for the least-squares mixed system with piecewise linear elements are presented. A dispersion analysis with graphical representation of the damping effect of the least-squares scheme and comparison to the corresponding Galerkin treatment is included both in the absence and presence of Coriolis. In the latter case, we are particularly interested in the dissipative effects for Rossby waves in ocean and atmosphere modelling applications and the stability (and damping) of gravity waves. In Section 5, this stability/dispersion analysis is followed by numerical tests to simulate slowly propagating Rossby modes. Some concluding remarks complete the study.

2. GOVERNING EQUATIONS

For an enclosed domain of length L, where $0<x<L$, the one dimensional, inviscid, linearized form of the shallow-water equations may be expressed in Cartesian coordinates [18] as

$$u_t - f\,v + g\,\eta_x = 0 \tag{1}$$

$$v_t + f\,u = 0 \tag{2}$$

$$\eta_t + H\,u_x = 0 \tag{3}$$

where u, v are the velocity components, η is the surface elevation with respect to the reference level $z=0$, g is the gravitational acceleration and the Coriolis parameter f and the mean depth H are assumed constant. Boundary and initial data complete the specification of the problem.

Equations (1)–(3) are rewritten in a non-dimensional form in terms of the variables

$$x' = x/L, \quad t' = t/T, \quad u' = u/U, \quad v' = v/U, \quad \eta' = \eta/H \tag{4}$$

with $L = aE^{-1/4}$, $T = E^{1/4}/(2\Omega)$ and $U = \sqrt{gH}$, where $E = 4\Omega^2 a^2/(gH)$ is the Lamb number, a is the radius of the earth, and Ω is the angular frequency of the earth's rotation.

Substituting the dimensionless variables from (4) into (1)–(3), and omitting the primes, we obtain

$$u_t - f_0 v + \eta_x = 0 \tag{5}$$

$$v_t + f_0 u = 0 \tag{6}$$

$$\eta_t + u_x = 0 \tag{7}$$

where $f_0 = fT$ is a non-dimensional parameter.

3. TEMPORAL DISCRETIZATION

For a given time step $\Delta t = t^{n+1} - t^n$ we introduce a general two-level time discretization of (5)–(7) of the form

$$u^{n+1} - (1-\gamma) f_0 \Delta t v^{n+1} + (1-\alpha)\Delta t \eta_x^{n+1} = u^n + \gamma f_0 \Delta t v^n - \alpha \Delta t \eta_x^n \tag{8}$$

$$v^{n+1} + (1-\gamma) f_0 \Delta t u^{n+1} = v^n - \gamma f_0 \Delta t u^n \tag{9}$$

$$\eta^{n+1} + (1-\alpha)\Delta t u_x^{n+1} = \eta^n - \alpha \Delta t u_x^n \tag{10}$$

where α and γ are real parameters, such that $0 \leqslant \alpha \leqslant 1$ and $0 \leqslant \gamma \leqslant 1$. Observe that the standard choices $\alpha = \gamma = 1, 1/2, 0$ yield the respective forward Euler, trapezoidal Crank–Nicolson and backward Euler type schemes.

Equations (8)–(10) are now spatially discretized using the least-squares method.

4. LEAST-SQUARES FORMULATION

The weighted least-squares residual variational functional for (8)–(10) is simply

$$\begin{aligned} I(u,v,\eta) = \int_0^1 \{ & v_1[u^{n+1} - u^n - (1-\gamma) f_0 \Delta t v^{n+1} - \gamma f_0 \Delta t v^n \\ & + (1-\alpha)\Delta t \eta_x^{n+1} + \alpha \Delta t \eta_x^n]^2 \\ & + v_2[v^{n+1} - v^n + (1-\gamma) f_0 \Delta t u^{n+1} + \gamma f_0 \Delta t u^n]^2 \\ & + v_3[\eta^{n+1} - \eta^n + (1-\alpha)\Delta t u_x^{n+1} + \alpha \Delta t u_x^n]^2 \} \, \mathrm{d}x \end{aligned} \tag{11}$$

where v_1, v_2, v_3 are scaling weights for the respective residuals in a multi-objective optimization sense. In the present work we take $v_1 = v_2 = v_3$. The problem is to minimize the least-squares functional over admissible variations of $u^{n+1}, v^{n+1}, \eta^{n+1}$. Setting $\delta I = 0$ we obtain: find $u^{n+1}, v^{n+1}, \eta^{n+1}$ satisfying the essential boundary conditions and such that the following holds

for each time step Δt:

$$\int_0^1 (au^{n+1}\varphi + (1-\alpha)^2\Delta t^2 u_x^{n+1}\varphi_x)\,\mathrm{d}x$$
$$= \int_0^1 (bu^n\varphi - \alpha(1-\alpha)\Delta t^2 u_x^n\varphi_x + f_0\Delta t v^n\varphi - \Delta t\eta_x^n\varphi)\,\mathrm{d}x \tag{12}$$

$$\int_0^1 (av^{n+1}\varphi - (1-\alpha)(1-\gamma)f_0\Delta t^2\eta_x^{n+1}\varphi)\,\mathrm{d}x$$
$$= \int_0^1 (bv^n\varphi - f_0\Delta t u^n\varphi + \alpha(1-\gamma)f_0\Delta t^2\eta_x^n\varphi)\,\mathrm{d}x \tag{13}$$

$$\int_0^1 (\eta^{n+1}\psi + (1-\alpha)^2\Delta t^2\eta_x^{n+1}\psi_x - (1-\alpha)(1-\gamma)f_0\Delta t^2 v^{n+1}\psi_x)\,\mathrm{d}x$$
$$= \int_0^1 (\eta^n\psi - \alpha(1-\alpha)\Delta t^2\eta_x^n\psi_x + \gamma(1-\alpha)f_0\Delta t^2 v^n\psi_x + \Delta t u^n\psi_x)\,\mathrm{d}x \tag{14}$$

where $a = 1 + (1-\gamma)^2 f_0^2\Delta t^2$ and $b = 1 - \gamma(1-\gamma)f_0^2\Delta t^2$, and φ and ψ are the velocity and surface elevation test functions, respectively.

Consider a uniform mesh of m intervals on $(0,1)$ and let $h = 1/m$ denote the meshlength parameter. Introducing the finite-element basis leads to a finite-element statement as in (12)–(14) but with $u^{n+1}, v^{n+1}, \eta^{n+1}$ replaced by the finite-element trial functions $u_h^{n+1}, v_h^{n+1}, \eta_h^{n+1}$ and φ, ψ replaced by the corresponding finite-element test functions φ_h, ψ_h for the selected meshes and bases.

In particular, we shall consider the case where $u_h^{n+1}, v_h^{n+1}, \eta_h^{n+1}$ belong to the space of continuous piecewise-linear polynomials, with $d = h = 1/m$ the nodal spacing. Since the least-squares formulation involves a minimization it leads to symmetric positive algebraic systems and the scheme is not subject to the usual consistency requirements associated with mixed Galerkin methods. Hence, in the present study we may employ the same piecewise-linear basis $\{\varphi_j\}$ for all field variables u_h, v_h, η_h but will also consider both co-located and staggered meshes (off-set by $d/2$) for velocity and elevation, respectively. Other possible choices of bases may be considered but those investigated here are of most practical interest.

4.1. *Pure gravity modes* ($f_0 = 0$)

4.1.1. Co-located meshes. Both field variables (u_h, η_h) have piecewise-linear Lagrange bases with nodes co-located at the same positions x_j $(j = 0, 1, \ldots, m)$ on $(0,1)$. Equations (12) and (14) then lead to

$$\frac{h}{6}(u_{j-1}^{n+1} + 4u_j^{n+1} + u_{j+1}^{n+1}) + (1-\alpha)^2\frac{\Delta t^2}{h}(-u_{j-1}^{n+1} + 2u_j^{n+1} - u_{j+1}^{n+1})$$

$$= \frac{h}{6}(u_{j-1}^n + 4u_j^n + u_{j+1}^n) - \alpha(1-\alpha)\frac{\Delta t^2}{h}(-u_{j-1}^n + 2u_j^n - u_{j+1}^n) + \frac{\Delta t}{2}(\eta_{j-1}^n - \eta_{j+1}^n) \quad \text{for } j = 1, 2, \ldots, m-1 \tag{15}$$

$$\frac{h}{6}(\eta_{j-1}^{n+1} + 4\eta_j^{n+1} + \eta_{j+1}^{n+1}) + (1-\alpha)^2\frac{\Delta t^2}{h}(-\eta_{j-1}^{n+1} + 2\eta_j^{n+1} - \eta_{j+1}^{n+1}) = \frac{h}{6}(\eta_{j-1}^n + 4\eta_j^n + \eta_{j+1}^n) - \alpha(1-\alpha)\frac{\Delta t^2}{h}(-\eta_{j-1}^n + 2\eta_j^n - \eta_{j+1}^n) + \frac{\Delta t}{2}(u_{j-1}^n - u_{j+1}^n) \quad \text{for } j = 1, 2, \ldots, m-1 \tag{16}$$

Periodic solutions of discrete system (15)–(16) corresponding to $u = \tilde{u}e^{i(kx+\omega t)}$ and $\eta = \tilde{\eta}\, e^{i(kx+\omega t)}$ are sought, where $\tilde{u}$ and $\tilde{\eta}$ are amplitudes, k is the wave number in the x-direction and ω is the angular frequency. Substituting in (15)–(16) and setting

$$p = \tfrac{1}{3}(2 + \cos kh), \quad q = \tfrac{1}{2}(1 - \cos kh), \quad s_1 = \sin kh, \quad E = e^{i\omega\Delta t} \tag{17}$$

to simplify the resulting relations, we get

$$\{p(E-1) + 4c^2q(1-\alpha)(E[1-\alpha] + \alpha)\}\tilde{u} + ics_1\tilde{\eta} = 0 \tag{18}$$

$$ics_1\tilde{u} + \{p(E-1) + 4c^2q(1-\alpha)(E[1-\alpha] + \alpha)\}\tilde{\eta} = 0 \tag{19}$$

where $c \equiv \Delta t/h$ is the gravitational Courant number.

For a non-trivial solution $(\tilde{u}, \tilde{\eta})$ to exist, the determinant of the coefficients in (18)–(19) must vanish. This leads to

$$E = \frac{p - 4c^2q\alpha(1-\alpha) \pm ics_1}{p + 4c^2q(1-\alpha)^2} \tag{20}$$

or equivalently,

$$E = |E|e^{i\theta} \tag{21}$$

where θ represents the change in argument (or phase change) of the numerical solution in each time step and

$$|E| = \frac{\sqrt{(p - 4c^2q\alpha[1-\alpha])^2 + c^2s_1^2}}{p + 4c^2q(1-\alpha)^2} \tag{22}$$

A necessary and sufficient condition for stability is that $|E| \leqslant 1$.

For $\alpha = 1$ (forward Euler scheme) the least-squares finite-element scheme reverts to the Galerkin finite-element scheme. Consequently, for $kh \neq 0\ (\pi)$ Equation (22) gives $|E| > 1$ and the gravity modes are unstable.

For $\alpha = 0$ and $\alpha = \frac{1}{2}$, $|E|$ is graphed in Figure 1 and the behaviour is summarized in table form, including the maximum and minimum values of $|E|$. The results obtained with the least-squares scheme (LS) are compared with those obtained with the corresponding mixed

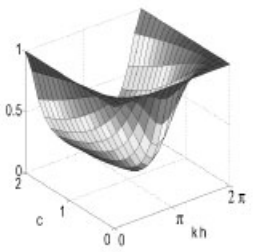
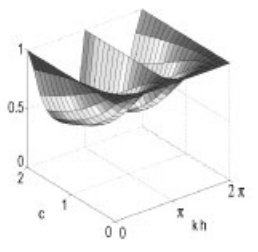
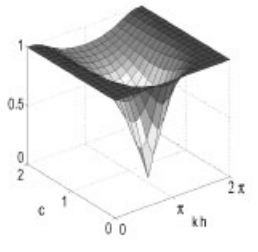
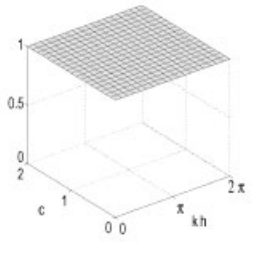
<table>
<tr>
<td rowspan="2">$\alpha = 0$</td>
<td>$$|E|^{LS} = \frac{\sqrt{p^2 + c^2 s_1^2}}{p + 4c^2 q}$$
• $|E|^{LS} < 1$ if $kh \neq 0\ (2\pi)$, the gravity modes are stable but damped.
• $|E|^{LS}_{max} = 1$ for $kh = 0\ (2\pi)$
• $|E|^{LS}_{min} = \frac{1}{(1 + 12c^2)}$
for $kh = \pi\ (2\pi)$</td>
<td>  </td>
</tr>
<tr>
<td>$$|E|^{G} = \frac{p}{\sqrt{p^2 + c^2 s_1^2}}$$
• $|E|^{G} < 1$ if $kh \neq 0\ (\pi)$, the gravity modes are stable but damped.
• $|E|^{G}_{max} = 1$ for $kh = 0\ (\pi)$
• $|E|^{G}_{min} = \frac{1}{\sqrt{1 + 3c^2}}$
for $kh = 2\pi/3$ and $4\pi/3\ (2\pi)$</td>
<td>  </td>
</tr>
<tr>
<td rowspan="2">$\alpha = \frac{1}{2}$</td>
<td>$$|E|^{LS} = \frac{\sqrt{(p - c^2 q)^2 + c^2 s_1^2}}{p + c^2 q}$$
• $|E|^{LS} < 1$ if $kh \neq 0\ (2\pi)$, the gravity modes are stable but damped.
• $|E|^{LS}_{max} = 1$ for $kh = 0\ (2\pi)$
• $|E|^{LS}_{min} = \frac{|1 - 3c^2|}{1 + 3c^2}$
for $kh = \pi\ (2\pi)$</td>
<td>  </td>
</tr>
<tr>
<td>$$|E|^{G} = 1$$
• The gravity modes are neutrally stable.</td>
<td>  </td>
</tr>
</table>

Figure 1. $|E|$ on co-located meshes for pure gravity modes.

Galerkin scheme (G). For both schemes the mode associated with $kh = 0\ (2\pi)$ and leading to $\omega = 0$ and $|E| = 1$, corresponds to the hydrostatic mode or steady solution.

For $\alpha = 0$, of special interest is the observation that in the Galerkin case waves of length $2d$, corresponding to $kh = \pi\ (2\pi)$, lead to $\omega = 0$ (and $|E| = 1$) and the waves do not propagate. This mode is identified as a spurious mode and it corresponds to a physical eigenmode of the system which has its phase speed reduced to zero by the numerical method [7]. The least-squares scheme does not suffer from the presence of spurious modes and $|E| = 1$ only for the hydrostatic mode; in particular, waves of length $2d$ normally propagate. However, the least-

Table I. Cosine and sine of the phase change θ on co-located meshes.

		$\cos\theta$	$\sin\theta$
$\alpha=0$	LS and G	$\dfrac{p}{\sqrt{p^2+c^2s_1^2}}$	$\dfrac{cs_1}{\sqrt{p^2+c^2s_1^2}}$
$\alpha=\frac{1}{2}$	LS	$\dfrac{p-c^2q}{\sqrt{(p-c^2q)^2+c^2s_1^2}}$	$\dfrac{cs_1}{\sqrt{(p-c^2q)^2+c^2s_1^2}}$
	G	$\dfrac{4p^2-c^2s_1^2}{4p^2+c^2s_1^2}$	$\dfrac{4cps_1}{4p^2+c^2s_1^2}$

squares scheme excessively damps the gravity modes as shown in Figure 1, and the damping is even more dramatic than in the Galerkin case. For example, we obtain $|E|^{\mathrm{LS}}_{\min}\simeq 0.13$ and $|E|^{\mathrm{G}}_{\min}\simeq 0.61$ for $c=0.75$.

For $\alpha=1/2$ the Galerkin scheme is neutrally stable for the gravity modes whereas the least-squares scheme leads to $|E|^{\mathrm{LS}}_{\min}=0$ for $kh=\pi$ (2π) (waves of length $2d$) and $c=1/\sqrt{3}$ as shown in Figure 1. In the latter case, the damping of the gravity modes is significant for $0.45\leqslant c\leqslant 0.75$ since $|E|^{\mathrm{LS}}_{\min}<0.25$.

The analytical expressions for $\cos\theta$ and $\sin\theta$ are given in Table I and the corresponding diagrams are shown in Figure 2. As expected, both the least squares and the Galerkin solutions coincide for $\alpha=0$; in particular for $kh=\pi$ (2π) we have $\cos\theta=1$, so $\theta=0$ (2π) and there is no phase change of the numerical solutions. In the case $\alpha=1/2$, both schemes mainly differ with respect to the waves of length $2d$ corresponding to $kh=\pi$ (2π) mentioned above. Indeed, in the Galerkin case $\theta=0$ (2π) again as for $\alpha=0$, whereas with the least-squares scheme $\theta=0$ (2π) for $c<1/\sqrt{3}$ but $\theta=\pi$ (2π) for $c>1/\sqrt{3}$ implying a phase change of the numerical solution.

4.1.2. Staggered meshes. In this discretization, the field variables u_h are again located at the previous nodal positions x_j $(j=0,1,\ldots,m)$ on $(0,1)$ but the nodal field variables for η_h are now located at the cell midpoints; that is, at $x_{j+1/2}$ $(j=0,1,\ldots,m-1)$ midway between the velocity nodes. Equations (12) and (14) then lead to

$$\begin{aligned}&\frac{h}{6}(u_{j-1}^{n+1}+4u_j^{n+1}+u_{j+1}^{n+1})+(1-\alpha)^2\frac{\Delta t^2}{h}(-u_{j-1}^{n+1}+2u_j^{n+1}-u_{j+1}^{n+1})\\&\quad=\frac{h}{6}(u_{j-1}^{n}+4u_j^{n}+u_{j+1}^{n})-\alpha(1-\alpha)\frac{\Delta t^2}{h}(-u_{j-1}^{n}+2u_j^{n}-u_{j+1}^{n})\\&\qquad+\frac{\Delta t}{8}(\eta_{j-3/2}^{n}+5\eta_{j-1/2}^{n}-5\eta_{j+1/2}^{n}-\eta_{j+3/2}^{n})\quad\text{for } j=2,3,\ldots,m-2\end{aligned}\tag{23}$$

$$\begin{aligned}&\frac{h}{6}(\eta_{j-1/2}^{n+1}+4\eta_{j+1/2}^{n+1}+\eta_{j+3/2}^{n+1})+(1-\alpha)^2\frac{\Delta t^2}{h}(-\eta_{j-1/2}^{n+1}+2\eta_{j+1/2}^{n+1}-\eta_{j+3/2}^{n+1})\\&\quad=\frac{h}{6}(\eta_{j-1/2}^{n}+4\eta_{j+1/2}^{n}+\eta_{j+3/2}^{n})-\alpha(1-\alpha)\frac{\Delta t^2}{h}(-\eta_{j-1/2}^{n}+2\eta_{j+1/2}^{n}-\eta_{j+3/2}^{n})\\&\qquad+\frac{\Delta t}{8}(u_{j-1}^{n}+5u_j^{n}-5u_{j+1}^{n}-u_{j+2}^{n})\quad\text{for } j=1,2,\ldots,m-2\end{aligned}\tag{24}$$

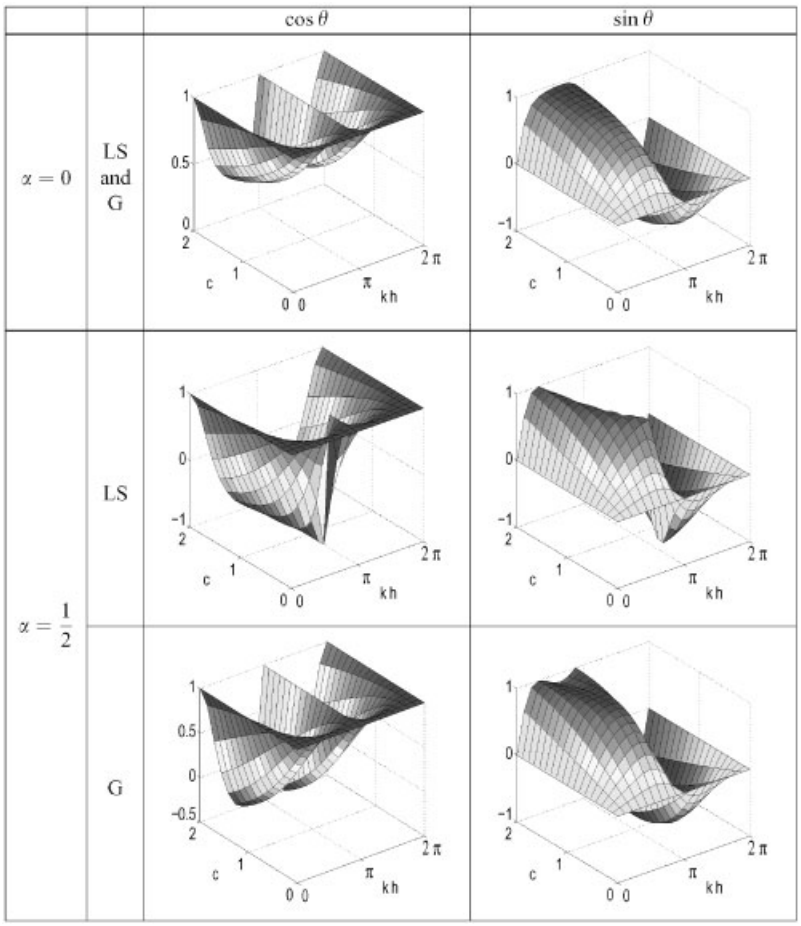


Figure 2. Cosine and sine of the phase change θ on co-located meshes.

Substituting as before for u, η and simplifying with

$$s_2 = \frac{1}{4}\left(\sin\frac{3kh}{2} + 5\sin\frac{kh}{2}\right) \tag{25}$$

Equations (23)–(24) imply

$$\{p(E-1) + 4c^2q(1-\alpha)(E[1-\alpha]+\alpha)\}\tilde{u} + \mathrm{i}cs_2\tilde{\eta} = 0 \tag{26}$$

$$\mathrm{i}cs_2\tilde{u} + \{p(E-1) + 4c^2q(1-\alpha)(E[1-\alpha]+\alpha)\}\tilde{\eta} = 0 \tag{27}$$

For a non-trivial solution $(\tilde{u}, \tilde{\eta})$ to exist, the determinant of the coefficients in (26)–(27) must vanish. This leads to

$$E = \frac{p - 4c^2 q\alpha(1-\alpha) \pm \mathrm{i}cs_2}{p + 4c^2 q(1-\alpha)^2} \tag{28}$$

and

$$|E| = \frac{\sqrt{(p - 4c^2 q\alpha[1-\alpha])^2 + c^2 s_2^2}}{p + 4c^2 q(1-\alpha)^2} \tag{29}$$

For $\alpha = 1$ (forward Euler scheme) the least-squares finite-element scheme reverts again to the Galerkin finite-element scheme. Consequently, for $kh \neq 0\ (\pi)$, Equation (29) gives $|E| > 1$ and the gravity modes are again unstable.

For $\alpha = 0$ and $\alpha = 1/2$, $|E|$ is graphed in Figure 3 and the behaviour is summarized in table form in the same manner as for the co-located meshes considered earlier. Both the least squares and the Galerkin schemes now exhibit solutions without spurious modes. However, for $\alpha = 0$ the damping of the gravity modes is significant, and slightly more pronounced in the least-squares case. For example, we obtain $|E|^{\mathrm{LS}}_{\min} \simeq 0.32$ and $|E|^{\mathrm{G}}_{\min} \simeq 0.41$ for $c = 0.75$. For $\alpha = 1/2$ the Galerkin scheme is neutrally stable for the gravity modes whereas the least-squares scheme still damps those modes but much less dramatically than in the case $\alpha = 0$. $|E|^{\mathrm{LS}}_{\min}$ is reached for $kh = \pi\ (2\pi)$ (waves of length $2d$) and $c = 1/\sqrt{3}$ with $|E|^{\mathrm{LS}}_{\min} = \sqrt{3}/2$.

The analytical expressions corresponding to $\cos\theta$ and $\sin\theta$ are obtained from Table I by replacing s_1 by s_2. The corresponding diagrams are not shown. As for the co-located case the Galerkin and least-squares schemes give identical results for $\alpha = 0$, and insignificant differences are observed between both schemes for $\alpha = 1/2$.

4.2. *Inertia-gravity and Rossby modes* ($f_0 \neq 0$)

Because of the previous results showing excessive damping of the gravity modes on co-located meshes, the inertia-gravity and Rossby modes are only considered here in the case of a staggered mesh for $\alpha = 1/2$. Applying now the discretization described in Section 4.1.2 for the field variables u_h, v_h and η_h, Equations (12)–(14) lead to

$$\begin{aligned} a\frac{h}{3}(u_{j-1}^{n+1} + 4u_j^{n+1} + u_{j+1}^{n+1}) + \frac{\Delta t^2}{2h}(-u_{j-1}^{n+1} + 2u_j^{n+1} - u_{j+1}^{n+1}) \\ = b\frac{h}{3}(u_{j-1}^{n} + 4u_j^{n} + u_{j+1}^{n}) - \frac{\Delta t^2}{2h}(-u_{j-1}^{n} + 2u_j^{n} - u_{j+1}^{n}) \\ + f_0\Delta t\frac{h}{3}(v_{j-1}^{n} + 4v_j^{n} + v_{j+1}^{n}) + \frac{\Delta t}{4}(\eta_{j-3/2}^{n} + 5\eta_{j-1/2}^{n} - 5\eta_{j+1/2}^{n} - \eta_{j+3/2}^{n}) \end{aligned} \tag{30}$$

$$\begin{aligned} a\frac{h}{3}(v_{j-1}^{n+1} + 4v_j^{n+1} + v_{j+1}^{n+1}) + (1-\gamma)f_0\frac{\Delta t^2}{8}(\eta_{j-3/2}^{n+1} + 5\eta_{j-1/2}^{n+1} - 5\eta_{j+1/2}^{n+1} - \eta_{j+3/2}^{n+1}) \\ = b\frac{h}{3}(v_{j-1}^{n} + 4v_j^{n} + v_{j+1}^{n}) - f_0\Delta t\frac{h}{3}(u_{j-1}^{n} + 4u_j^{n} + u_{j+1}^{n}) \\ - (1-\gamma)f_0\frac{\Delta t^2}{8}(\eta_{j-3/2}^{n} + 5\eta_{j-1/2}^{n} - 5\eta_{j+1/2}^{n} - \eta_{j+3/2}^{n}) \quad \text{for } j = 2, 3, \ldots, m-2 \end{aligned} \tag{31}$$

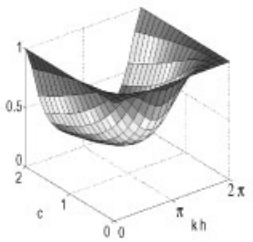
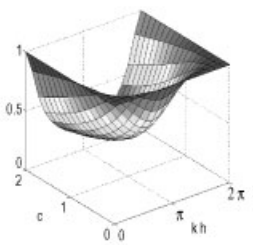
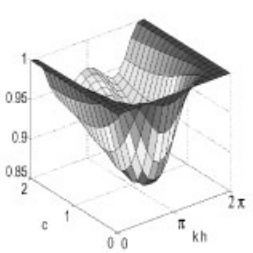
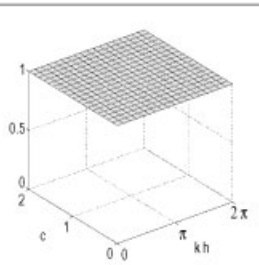
$\alpha = 0$	$\lvert E\rvert^{LS} = \dfrac{\sqrt{p^2 + c^2 q r}}{p + 4c^2 q}$ • $\lvert E\rvert^{LS} < 1$ if $kh \neq 0\ (2\pi)$, the gravity modes are stable but damped. • $\lvert E\rvert^{LS}_{\max} = 1$ for $kh = 0\ (2\pi)$ • $\lvert E\rvert^{LS}_{\min} = \dfrac{\sqrt{1 + 9c^2}}{(1 + 12c^2)}$ for $kh = \pi\ (2\pi)$	
	$\lvert E\rvert^{G} = \dfrac{p}{\sqrt{p^2 + c^2 q r}}$ • $\lvert E\rvert^{G} < 1$ if $kh \neq 0\ (\pi)$, the gravity modes are stable but damped. • $\lvert E\rvert^{G}_{\max} = 1$ for $kh = 0\ (\pi)$ • $\lvert E\rvert^{G}_{\min} = \dfrac{1}{\sqrt{1 + 9c^2}}$ for $kh = \pi\ (2\pi)$	
$\alpha = \frac{1}{2}$	$\lvert E\rvert^{LS} = \dfrac{\sqrt{(p - c^2 q)^2 + c^2 q r}}{p + c^2 q}$ • $\lvert E\rvert^{LS} < 1$ if $kh \neq 0\ (2\pi)$, the gravity modes are stable but damped. • $\lvert E\rvert^{LS}_{\max} = 1$ for $kh = 0\ (2\pi)$ • $\lvert E\rvert^{LS}_{\min} = \sqrt{1 - \dfrac{3c^2}{(1 + 3c^2)^2}}$ for $kh = \pi\ (2\pi)$	
	$\lvert E\rvert^{G} = 1$ • The gravity modes are neutrally stable.	

Figure 3. $|E|$ on staggered meshes for pure gravity modes.

$$\frac{h}{3}(\eta^{n+1}_{j-1/2} + 4\eta^{n+1}_{j+1/2} + \eta^{n+1}_{j+3/2}) + \frac{\Delta t^2}{2h}(-\eta^{n+1}_{j-1/2} + 2\eta^{n+1}_{j+1/2} - \eta^{n+1}_{j+3/2})$$
$$-(1-\gamma) f_0 \frac{\Delta t^2}{8}(v^{n+1}_{j-1} + 5v^{n+1}_{j} - 5v^{n+1}_{j+1} - v^{n+1}_{j+2})$$
$$= \frac{h}{3}(\eta^{n}_{j-1/2} + 4\eta^{n}_{j+1/2} + \eta^{n}_{j+3/2}) - \frac{\Delta t^2}{2h}(-\eta^{n}_{j-1/2} + 2\eta^{n}_{j+1/2} - \eta^{n}_{j+3/2})$$

$$+\gamma f_0 \frac{\Delta t^2}{8}(v_{j-1}^n + 5v_j^n - 5v_{j+1}^n - v_{j+2}^n)$$

$$+\frac{\Delta t}{4}(u_{j-1}^n + 5u_j^n - 5u_{j+1}^n - u_{j+2}^n) \quad \text{for } j = 1,2,\ldots,m-2 \tag{32}$$

Periodic solutions of (30)–(32) are again sought with $v = \tilde{v}e^{i(kx+\omega t)}$, and we obtain

$$(p[aE-b] + c^2 q[E+1])\tilde{u} - f_0 \Delta t p \tilde{v} + ics_2\tilde{\eta} = 0 \tag{33}$$

$$f_0 \Delta t p \tilde{u} + (p[aE-b] + c^2 q[E+1])\tilde{v} - ics_2 \frac{f_0 \Delta t}{2}(1-\gamma)(E+1)\tilde{\eta} = 0 \tag{34}$$

$$ics_2\tilde{u} + ics_2 \frac{f_0 \Delta t}{2}(E[1-\gamma]+\gamma)\tilde{v} + (p[E-1] + c^2 q[E+1])\tilde{\eta} = 0 \tag{35}$$

For a non-trivial solution $(\tilde{u}, \tilde{v}, \tilde{\eta})$ to exist, the determinant of the coefficients in (33)–(35) must vanish. Simplifying with

$$r = \frac{1}{4}(3 + \cos kh)^2, \quad s_2^2 = qr \tag{36}$$

leads to

$$\{p(aE-b) + qc^2(E+1)\} \times \left\{ p^2(aE-b)(E-1) + pqc^2(aE-b)(E+1) \right.$$

$$\left. - qrc^2 \frac{(f_0\Delta t)^2}{4}(1-\gamma)(E+1)(E[1-\gamma]+\gamma) \right\}$$

$$- p(f_0\Delta t)^2\{-p^2(E-1) - pqc^2(E+1) + \frac{qr}{2}c^2(E[1-\gamma]+\gamma)\}$$

$$+ pqrc^2 \left\{ (aE-b) - \frac{(f_0\Delta t)^2}{2}(1-\gamma)(E+1) \right\} = 0 \tag{37}$$

We let

$$A = (ap + qc^2)\left\{ ap^2 + \left(ap - \frac{[f_0\Delta t]^2}{4}[1-\gamma]^2 r \right) qc^2 \right\}$$

$$B = -a(a+2b)p^3 + \{(a^2 - 2ab - b)p + (2a-b)qc^2\}pqc^2$$

$$- \frac{(f_0\Delta t)^2}{4}(1-\gamma)\{(a + b[\gamma-1])p + (2-\gamma)qc^2\}qrc^2$$

$$C = (b^2 + 2ab + [f_0\Delta t]^2)p^3 - \frac{(f_0\Delta t)^2}{4}(1-\gamma)\{(a\gamma - b + 4)p + (1+\gamma)qc^2\}qrc^2$$

$$+ \{(b^2 - 2ab - a + [f_0\Delta t]^2)p + (a-2b)qc^2 + ar\}pqc^2$$

$$D = -(b^2 + [f_0\Delta t]^2)p^3 + \frac{(f_0\Delta t)^2}{4}\gamma(1-\gamma)(bp - qc^2)qrc^2 + \left\{(b^2 + b + [f_0\Delta t]^2)p - bqc^2 - \left(b + \frac{[f_0\Delta t]^2}{4}\right)r\right\} pqc^2$$

and (37) reduces to

$$AE^3 + BE^2 + CE + D = 0 \tag{38}$$

Two roots in (38) are complex conjugate and correspond to the inertia-gravity modes. For these modes, $|E|$ is graphed in Figure 4 in the case $\alpha = 1/2$ for $\gamma = 0$ and $1/2$, respectively, and the least squares and the Galerkin solutions are compared for two values of $f_0\Delta t$ (0.1 and 0.5). For $\gamma = 0$ and $1/2$ the least-squares solutions are comparable to those obtained in Figure 3 for gravity modes with no rotation ($f_0 = 0$). However, for $\gamma = 0$ as the time step progressively increases, the least-squares solution tends to draw nearer to the Galerkin case. For $\gamma = 1/2$ the Galerkin scheme is neutrally stable for the inertia-gravity modes, whereas in the case $\gamma = 0$ these modes are damped and the damping increases with increasing values of $f_0\Delta t$. For $\gamma = 1$ we obtain $|E|^{\mathrm{LS}}_{\max} = |E|^{\mathrm{G}}_{\max} = \sqrt{1 + (f_0\Delta t)^2} > 1$ for $kh = 0$ (2π), and hence the inertia-gravity modes (not shown) are unstable.

The third root in (38) is real and corresponds to the slow mode, the Rossby mode. $|E|$ is graphed for this mode in Figure 5 in the case $\alpha = 1/2$ with $\gamma = 0$ and $1/2$, and both the least squares and the Galerkin solutions are again compared with for the values $f_0\Delta t = 0.1$ and 0.5. For $\gamma = 0$ and $1/2$ the Galerkin scheme is neutrally stable for the Rossby modes whereas some damping occurs in the least-squares case, and this damping is enhanced with increasing values of $f_0\Delta t$.

The diagrams representing $\cos\theta$ and $\sin\theta$ (not shown) exhibit insignificant differences when the least squares and the Galerkin solutions are compared for both the inertia-gravity and the Rossby modes.

In the next section, some numerical tests are performed to evaluate the level of damping observed in Figure 5 in the least-squares discretization of the Rossby modes (the modes of interest), for $\alpha = 1/2$ and $\gamma = 1/2$.

5. NUMERICAL RESULTS

The slowly propagating Rossby modes are simulated as the one-dimensional part of an ideal eddy at midlatitudes. The least-squares finite-element discretized equations (30)–(32) (for $\gamma = 1/2$) are solved as a reduced-gravity model with parameters set to correspond to the first internal vertical mode of a baroclinic model. A second passive layer is implicitly assumed that is infinitely deep and at rest. This formulation precludes any influence of the bathymetry. The choice $H = 1.63$ m results in a phase speed for gravity waves of $\sqrt{gH} \approx 4\ \mathrm{m\ s^{-1}}$. Such a small equivalent depth is unusual for atmospheric studies, but it is pertinent in the oceanic context for the adjustment under gravity of a density-stratified fluid (e.g. see Reference [19]).

A Gaussian distribution is prescribed at initial time and the initial velocity field is taken to be in geostrophic equilibrium, a balance between the Coriolis and the pressure gradient

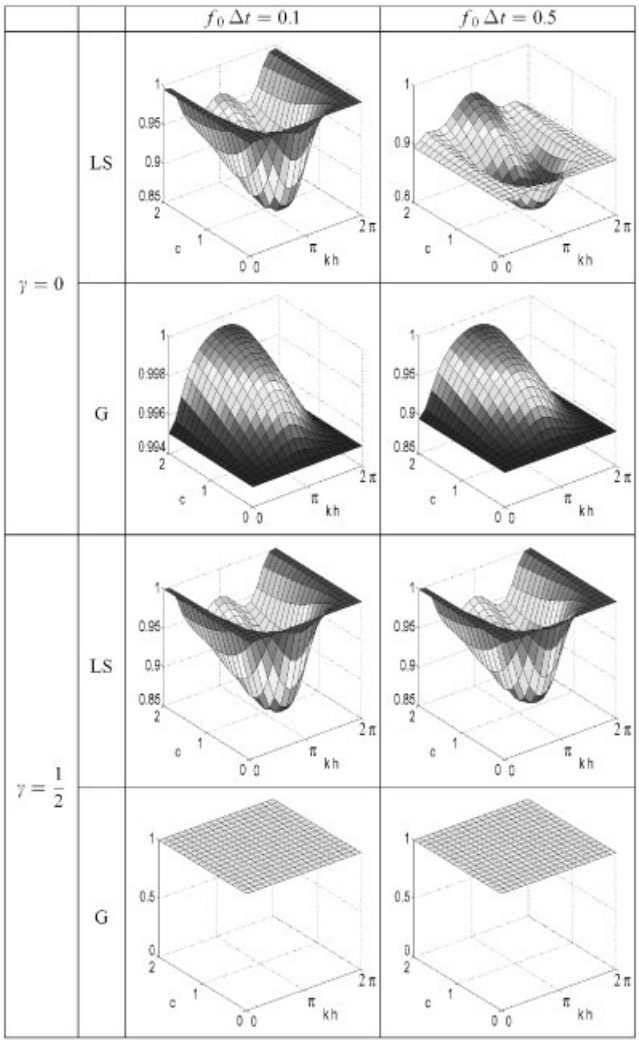


Figure 4. $|E|$ on staggered meshes in the case $\alpha = 1/2$, for inertia-gravity modes.

terms, i.e.

$$\eta(x,0) = \mu \mathrm{e}^{-\nu x^2} \tag{39}$$

$$u(x,0) = 0 \tag{40}$$

$$v(x,0) = \frac{1}{f_0}\,\eta_x(x,0) \tag{41}$$

The distance between two velocity nodes is taken to be 30 km and the domain is chosen sufficiently long to prevent the Gaussian from approaching the boundaries in the time interval

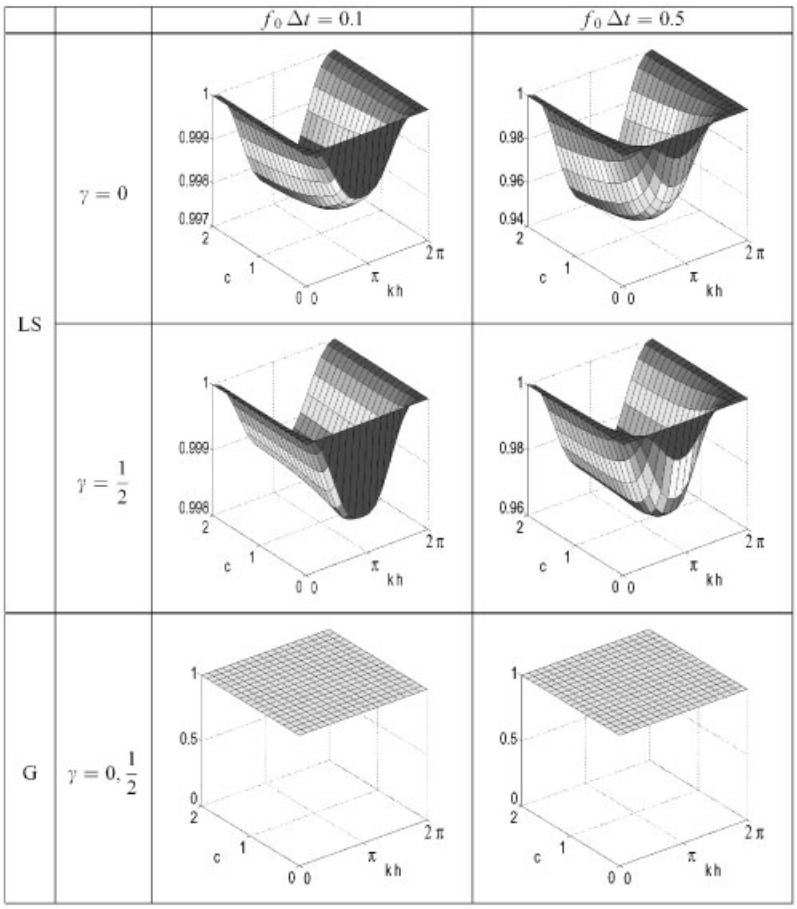


Figure 5. $|E|$ on staggered meshes in the case $\alpha = 1/2$, for Rossby modes.

for the simulation carried out. The Coriolis parameter is evaluated at 25°N and the radius of deformation is thus $R_d \equiv \sqrt{gH}/f \approx 65\,\text{km}$. The parameters μ and ν are chosen such that the e-folding radius of the initial Gaussian is resolved by four velocity nodes and the initial maximum velocity is $1\,\text{m}\,\text{s}^{-1}$. A simple scaling can provide more realistic atmospheric parameters without substantially modifying the results.

The evolution of the maximum values of the computed surface-elevation field is shown in Table II at different stages of propagation (10, 20 and 30 days) and for different time steps (15, 30 and 60 min). In all cases the maximum value of the surface-elevation field is scaled to 1 at the beginning of the experiment. Table II shows that the Rossby modes are damped, as anticipated, by the dispersion relation analysis, and the damping is found to be proportional both to the length of the time step and the duration of the simulation. For a two-dimensional experiment the damping is expected to be even more important.

Table II. Maximum values of the computed surface-elevation field at different stages of the propagation (10, 20 and 30 days) for different time steps (15, 30 and 60 min) in simulating the slowly propagating Rossby modes for $\alpha = 1/2$ and $\gamma = 1/2$ with the least-squares method on staggered meshes.

	10 days	20 days	30 days
15 min	0.993	0.986	0.980
30 min	0.986	0.973	0.961
60 min	0.973	0.950	0.929

It has also been computationally verified that the maximum value of the computed surface-elevation field remained unchanged in the Galerkin case, whatever the length of the time step and the duration of the experiment.

6. CONCLUDING REMARKS

This appears to be the first study of the dispersion relation and mode behaviour for finite-element solutions of the shallow-water equations based on the least-squares mixed approach. The graphical display of the dispersion relation implies that the least-squares scheme can again be applied without the usual restrictions on the bases associated with mixed Galerkin schemes. Moreover, the method does not exhibit folding of the dispersion relation. However, the analysis also shows that both the inertia-gravity and Rossby modes are over-damped. The latter behaviour might have been anticipated based on the dissipative nature of the least-squares scheme observed in previous studies.

The damping of the Rossby mode might be an issue. Indeed, a high-order spatial accuracy is usually desirable for the treatment of the slow Rossby modes, and has proven practical and beneficial for atmospheric and oceanic prediction models. The damping observed for the Rossby modes in the present simulations, which is expected to be even more important for two-dimensional experiments, suggests the least-squares approach should be used with care (if at all) and restricted to short time steps and short simulations.

REFERENCES

1. Carey GF (ed.). *Finite Element Modeling of Environmental Problems*. Wiley: UK, 1995.
2. Côté J, Staniforth A. An accurate and efficient finite-element global model of the shallow-water equations. *Monthly Weather Review* 1990; **118**:2707–2717.
3. Hua BL, Thomasset F. A noise-free finite-element scheme for the two-layer shallow-water equations. *Tellus* 1984; **36A**:157–165.
4. Le Roux DY, Lin CA, Staniforth A. A semi-implicit semi-Lagrangian finite-element shallow-water ocean model. *Monthly Weather Review* 2000; **128**:1384–1401.
5. Walters RA, Cheng RT. Accuracy of an estuarine hydrodynamic model using smooth elements. *Water Resources Research* 1980; **16**(1):187–195.
6. Lynch DR, Gray WG. A wave-equation model for finite-element tidal computations. *Computers and Fluids* 1979; **7**:207–228.
7. Walters RA, Carey GF. Analysis of spurious oscillation modes for the shallow-water and Navier–Stokes equations. *Computers and Fluids* 1983; **11**:51–68.
8. Le Provost C, Bernier C, Blayo E. An intercomparison of two numerical methods for integrating a quasi-geostrophic multilayer model of ocean circulations. *Journal of Computational Physics* 1993; **110**:341–359.
9. Le Roux DY, Staniforth A, Lin CA. Finite elements for shallow-water equation ocean models. *Monthly Weather Review* 1998; **126**:1931–1951.

10. Myers PG, Weaver AJ. A diagnostic barotropic finite-element ocean circulation model. *Journal of Atmospheric and Oceanic Technology* 1995; **12**:511–526.
11. Walters RA, Casulli V. A robust, finite-element model for hydrostatic surface water flows. *Communications in Numerical Methods in Engineering* 1998; **14**:931–940.
12. Finlayson BA. *The Method of Weighted Residuals and Variational Principles*. Academic Press: New York, 1972.
13. Carey GF, Shen Y. Convergence studies of least-squares finite elements for first-order systems. *Communications in Applied Numerical Methods* 1989; **5**:427–434.
14. Carey GF, Pehlivanov AI, Shen Y, Bose A, Wang KC. Least-squares finite elements for fluid flow and transport. *International Journal for Numerical Methods in Fluids* 1998; **27**:97–107.
15. Pehlivanov AI, Carey GF. Error estimates for least-squares mixed finite elements. *RAIRO Mathematical Models in Numerical Analysis* 1994; **28**:499–516.
16. Pehlivanov AI, Carey GF, Lazarov RD. Least-squares mixed finite elements for second order elliptic problems. *SIAM Journal on Numerical Analysis* 1994; **31**:1368–1377.
17. Carey GF, Pehlivanov AI, Shen Y. Least-squares mixed finite elements. In *Finite Element Methods: Fifty Years of the Courant Method*, Krizek M, Neittaanmaki P, Stenberg R (eds). Marcel Dekker: New York, 1994; 105–117.
18. LeBlond PH, Mysak LA. *Waves in the Ocean*. Elsevier: Amsterdam, 1978.
19. Gill AE. *Atmosphere-Ocean Dynamics*. International Geophysics Series, vol. 31. Academic Press: San Diego, 1982.

