INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN FLUIDS
Int. J. Numer. Meth. Fluids 2012; **68**:1497–1523
Published online 22 February 2011 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/fld.2540

A comparison of the GWCE and mixed P_1^{NC}–P_1 formulations in finite-element linearized shallow-water models

D. Y. Le Roux[1,*,†], R. Walters[2], E. Hanert[3] and J. Pietrzak[4]

[1]*Université de Lyon, CNRS, Université Lyon 1, Institut Camille Jordan, 43, blvd du 11 novembre 1918, 69622 Villeurbanne Cedex, France*
[2]*Hydrodynamics, O/RM, 6051 Hunt Rd, Vic., Canada BC V8Y 3H7*
[3]*Earth and Life Institute, Université catholique de Louvain, Croix du Sud 2/16, B-1348 Louvain-la-Neuve, Belgium*
[4]*Faculteit CiTG, TU Delft, Stevinweg 1, 2628 CN Delft, The Netherlands*

SUMMARY

The appearance of spurious pressure modes in early shallow-water (SW) models has resulted in two common strategies in the finite element (FE) community: using mixed primitive variable and generalized wave continuity equation (GWCE) formulations of the SW equations. One FE scheme in particular, the P_1^{NC}–P_1 pair, combined with the primitive equations may be advantageously compared with the wave equation formulations and both schemes have similar data structures. Our focus here is on comparing these two approaches for a number of measures including stability, accuracy, efficiency, conservation properties, and consistency. The main part of the analysis centres on stability and accuracy results via Fourier-based dispersion analyses in the context of the linear SW equations. The numerical solutions of test problems are found to be in good agreement with the analytical results. Copyright © 2011 John Wiley & Sons, Ltd.

Received 8 June 2010; Revised 17 December 2010; Accepted 27 December 2010

KEY WORDS: shallow-water equations; generalized wave continuity equation; finite elements; dispersion analysis; gravity waves

1. INTRODUCTION

The shallow-water (SW) equations have been extensively employed in environmental studies to model hydrodynamics in estuaries, lakes, coastal regions, and other applications. These equations are obtained by integrating the Navier–Stokes system over the depth of the fluid layer under Boussinesq and hydrostatic pressure assumptions. The SW system is two dimensional but it retains much of the dynamical complexity of three-dimensional flows on the rotating Earth. Indeed, the SW equations are the simplest geophysical flow model allowing the representation of two classes of wave motions which are intimately involved in basin-scale adjustment processes: inertia-gravity (Poincaré) and planetary (Rossby) waves [1, 2]. The simulation of gravity waves permits the representation of phenomena such as tsunamis or mountain waves, whereas on larger scale, slow Rossby waves play an important role in the global circulation. Owing to its inherent simplicity, the SW system is often used as a prototype of the primitive equations and is frequently employed as a benchmark for numerical schemes to be used in more complex oceanic or atmospheric models.

*Correspondence to: D. Y. Le Roux, Université de Lyon, CNRS, Université Lyon 1, Institut Camille Jordan, 43, blvd du 11 novembre 1918, 69622 Villeurbanne Cedex, France.
†E-mail: dleroux@math.univ-lyon1.fr

Early attempts to apply numerical methods to the solution of the SW equations were confronted with a number of problems. The most serious of these was the occurrence of spurious computational modes that may arise for certain choices of grids and bases due to the coupling between the momentum and continuity equations. Attention has been focused primarily on the spurious surface-elevation (pressure) modes since these were argued to be the most troublesome [3]. For example, the finite-difference (FD) B-grid (subject to the no-slip boundary condition) and the piecewise linear mixed Galerkin (for both velocity and elevation variables, i.e. the P_1–P_1 pair) discretizations are usually plagued by spurious oscillations [4–7]. Spurious Coriolis or f-modes may also exist for the C-grid FD and the finite-element (FE) pair RT_0–P_0 when the grid resolution is coarse with respect to the deformation radius [8–11]. However, such f-modes could be controlled with suitable resolution or an accurate and stable procedure when reconstructing tangential velocities for the C-grid [12]. More recently, spurious velocity modes have been identified for mixed FE discretizations using the so-called BDM_1–P_0 and BDM_1–P_1 FE pairs, and named CD-modes in [13]. Such spurious pressure, f-, and velocity modes are small-scale artefacts, introduced by the spatial discretization scheme, which do not propagate but are trapped within the model grid, and associated with zero frequency. If they are left undamped, they can cause aliasing and an accumulation of energy in the smallest-resolvable scale, leading to noisy solutions. Other types of spurious modes have also been detected: propagating inertial oscillations that have no particular spatial characteristics [7, 9, 14] and modes in $O(1/h)$ (where h is a representative meshlength parameter that measures resolution) in discontinuous Galerkin approximations of the SW equations [15]. The latter modes seem to be damped in linear models; however, their impact on the numerical solution is still unclear for non-linear models. The appearance of the above-mentioned spurious modes is mainly due to an inappropriate placement of variables on the grid and/or bad choice of approximation function spaces.

Once the origin of the spurious pressure modes was identified, there seemed to be two obvious solutions: finding discretizations (FD stencils and FE pairs) that do not support these modes, or modifying the SW equations to remove the troublesome terms. The first option allows the use of the primitive SW equations with staggered FD grids [4, 9, 16]. The problem with FE methods proved to be more difficult to solve since conventional elements with the same approximation for surface-elevation and velocity turned out to be the worst choices. Mixed-order FE interpolation methods [17–20], equal-order elements with variables carried at sets of points staggered in space [21–23], analogous to staggered FDs, stabilization and Petrov–Galerkin methods [24], yield solutions free of spurious pressure modes. Following the second option, vorticity and divergence formulations were considered [25, 26] and a wave-equation formulation was formed [27, 28], which has seen considerable development and applications [29–32]. The wave continuity equation (WCE) was primarily formed by differentiating the continuity equation in time, substituting from the momentum equation, and rearranging terms [27]. The generalized wave continuity equation (GWCE) was later obtained by introducing into the WCE a weighting parameter G which determines the balance between the primitive and pure wave form [33]. Much effort has been expanded trying to find optimal values for G [34]. The development of wave continuity-based SW models from 1979 to 1999 is reviewed in [35].

Finite volume and discontinuous Galerkin methods using upwind schemes are usually free of spurious pressure modes, but these schemes are not examined in this study. For example, five FE pairs are compared in [36] in several relevant regimes of the subcritical SW flow. Continuous, discontinuous, and partially discontinuous FE formulations are considered using Riemann solver to evaluate interface integrals. In particular, a new pair, using non-conforming linear elements (P_1^{NC}) for both velocities and elevation, is introduced and gives optimal rates of convergence in the simulated test cases.

Mixed primitive variable and wave equation formulations of the SW equations are extensively employed in environmental studies using Galerkin techniques. This is largely because of the need to represent irregular boundary geometry in many applications, using grids of variable sizes, shapes and orientation, and for local mesh refinement. Fourier and dispersion relation analyses have been performed for both mixed primitive variable [7, 13, 37–39] and wave equation

DOI: 10.1002/fld

[27, 28, 34, 40, 41] Galerkin formulations to detect the eventual presence of spurious modes as well as the dissipative/dispersive nature of the formulations. From all these studies, there have emerged a small number of FE schemes, namely the P_0–P_1, P_1^{NC}–P_1, and RT_0–P_0 pairs that are suitable for use with the primitive SW equations, and a considerable amount of analysis and applications of the wave equation formulation. In particular, the three FE pairs cited above are not subject to spurious pressure (surface-elevation) modes.

The wave equation formulations have longstanding problems with mass conservation, stability when advection is important, and consistency with scalar transport equations. The use of suitable FE pairs with the primitive equations may solve these issues and is discussed in this paper. One FE scheme in particular, the P_1^{NC}–P_1 pair, combined with the primitive equations may be advantageously compared with the GWCE formulation and both schemes have similar data structures. Our focus here is on comparing these two approaches for a number of measures including stability, accuracy, efficiency, conservation properties, and consistency. The main part of the analysis centres on stability and accuracy results via Fourier-based dispersion analyses in the context of the linear SW equations. For both schemes we have used existing results and supplemented these with new results derived here.

For the WCE, dispersion analyses have been performed in 1-D [27, 40, 41] neglecting Coriolis force, and in 2-D [28, 42] for several time-stepping schemes. The results show that the method does not contain spurious pressure modes and provides an accurate solution for an explicit temporal scheme. However, the solution is overdamped for implicit time-integration [43]. For the GWCE, the dispersion analysis has been done in 1-D [34] and 2-D [44] assuming that the time is continuous and neglecting Coriolis effects. In this paper we extend the dispersion analysis of the GWCE in 2-D by including the Coriolis terms and a general 2-level time discretization scheme. Finally, for the P_1^{NC}–P_1 FE pair, the dispersion analysis performed in [14] is extended here by considering a bottom friction term and analysing the time discretization results.

The paper is developed as follows: The model equations, time, and Galerkin FE discretization schemes are presented in Section 2. The dispersion relations are computed and analysed in Sections 3 and 4, respectively. Numerical tests are performed in Section 5, followed by a discussion on the method's properties in Section 6. Conclusions are drawn in Section 7.

2. SPATIAL AND TEMPORAL DISCRETIZATIONS

2.1. Governing equations

Let Ω be the model domain with boundary Γ. The linear SW system is expressed in Cartesian coordinates [45] as

$$\mathbf{M} \equiv \frac{\partial \widetilde{\mathbf{u}}}{\partial t} + f\,\mathbf{k}\times\widetilde{\mathbf{u}} + \tau\widetilde{\mathbf{u}} + g\,\nabla\widetilde{\eta} = 0, \tag{1}$$

$$C \equiv \frac{\partial \widetilde{\eta}}{\partial t} + H\,\nabla\cdot\widetilde{\mathbf{u}} = 0, \tag{2}$$

where $\widetilde{\mathbf{u}}(\mathbf{x}, t) = (\widetilde{u}, \widetilde{v})$ is the velocity field with $\mathbf{x} = (x, y)$, $\widetilde{\eta}(\mathbf{x}, t)$ is the surface-elevation with respect to the reference level $z = 0$, g and τ are the gravitational acceleration and the bottom friction coefficient, $\mathbf{k}$ is a unit vector in the vertical direction, leading to $\mathbf{k}\times\widetilde{\mathbf{u}} = (-\widetilde{v}, \widetilde{u})$ by considering that the vertical component of $\widetilde{\mathbf{u}}$ is zero, and the mean depth H and the Coriolis parameter f are assumed constant. Periodic boundary conditions are considered in this study.

By using operator notation, where $\mathbf{M}$ and C represent the momentum and continuity equations (1) and (2), respectively, the GWCE is given as

$$\frac{\partial C}{\partial t} - H\nabla\cdot\mathbf{M} + GC = 0, \tag{3}$$

DOI: 10.1002/fld

and we obtain

$$\frac{\partial^2\widetilde{\eta}}{\partial t^2}+G\frac{\partial\widetilde{\eta}}{\partial t}+fH\mathrm{rot}\,\widetilde{\mathbf{u}}+H(G-\tau)\nabla\cdot\widetilde{\mathbf{u}}-gH\nabla^2\widetilde{\eta}=0, \tag{4}$$

where $\mathrm{rot}\,\widetilde{\mathbf{u}}\equiv\widetilde{v}_x-\widetilde{u}_y$, and (1) and (4) are solved for the dependent variables $(\widetilde{\eta},\widetilde{\mathbf{u}})$. The parameter G was originally introduced to control the numerical properties of the solution, particularly to reduce the errors in the primitive continuity equation with a relaxation time $1/G$. Indeed, G determines the balance between the primitive and pure wave form in (4). The higher the magnitude of G, the more the GWCE (4) approaches the primitive continuity equation (2). Finally, as originally presented in [27], the WCE can be obtained from (4) by setting $G=\tau$.

2.2. Temporal discretization

We consider real parameters α, β, γ and θ belonging to the interval $[0,1]$ and let

$$\Phi^{1,\theta}=\theta\Phi^{n+1}+(1-\theta)\Phi^n,\quad \Phi^{2,\theta}=\theta\Phi^{n+1}+(1-2\theta)\Phi^n+\theta\Phi^{n-1},$$

where $\Phi=\widetilde{\mathbf{u}}$ or $\widetilde{\eta}$. For a given time step $\Delta t=t^{n+1}-t^n$, with $t^n=n\Delta t$, $n=0,1,2,\ldots$, we introduce a general 2-time-level discretization of the SW system (1)–(2) of the form

$$\frac{\widetilde{\mathbf{u}}^{n+1}-\widetilde{\mathbf{u}}^n}{\Delta t}+f\mathbf{k}\times\widetilde{\mathbf{u}}^{1,\gamma}+\tau\widetilde{\mathbf{u}}^{1,\beta}+g\nabla\widetilde{\eta}^{1,\alpha}=0, \tag{5}$$

$$\frac{\widetilde{\eta}^{n+1}-\widetilde{\eta}^n}{\Delta t}+H\nabla\cdot\widetilde{\mathbf{u}}^{1,\alpha}=0. \tag{6}$$

Observe that the standard choices $\alpha=\beta=\gamma=0,\frac{1}{2},1$ yield the respective forward Euler, trapezoidal Crank–Nicolson, and backward Euler-type schemes.

For the GWCE (4) a 3-time-level scheme centred at n is used

$$\frac{\widetilde{\eta}^{n+1}-2\widetilde{\eta}^n+\widetilde{\eta}^{n-1}}{\Delta t^2}+G\frac{\widetilde{\eta}^{n+1}-\widetilde{\eta}^{n-1}}{2\Delta t}+fH\mathrm{rot}\,\widetilde{\mathbf{u}}^{2,\mu}+H(G-\tau)\nabla\cdot\widetilde{\mathbf{u}}^{2,\nu}-gH\nabla^2\widetilde{\eta}^{2,\delta}=0, \tag{7}$$

where δ, μ, and ν are real parameters belonging to the interval $[0,1]$, and the system (5) and (7) is solved. As in [28], a 2-time-level scheme is employed in (5) in order to prevent the appearance of numerical artefacts. Such spurious solutions are present in the system (5) and (7) when a 3-time-level scheme is used in (5).

Because (5), (6), and (7) are linear equations with constant coefficients, we seek periodic solutions of the form

$$\widetilde{\mathbf{u}}^n=\mathbf{u}(\mathbf{x})\mathrm{e}^{\mathrm{i}\omega t^n},\qquad \widetilde{\eta}^n=\eta(\mathbf{x})\mathrm{e}^{\mathrm{i}\omega t^n},\quad n=1,2,3,\ldots, \tag{8}$$

where $\mathbf{u}(\mathbf{x})$ and $\eta(\mathbf{x})$ are the amplitudes of the velocity field and surface-elevation, respectively. By inserting the Fourier expansions (8) into (5), (6) and (7), and letting $E=\mathrm{e}^{\mathrm{i}\omega\Delta t}$, $F=f\Delta t$, $K=\tau\Delta t$, $L=G\Delta t$, $E_{1,\theta}=\theta E+1-\theta$, and $E_{2,\theta}=\theta E^2+(1-2\theta)E+\theta$, we obtain, respectively,

$$(E-1)\mathbf{u}+FE_{1,\gamma}\mathbf{k}\times\mathbf{u}+KE_{1,\beta}\mathbf{u}+g\Delta tE_{1,\alpha}\nabla\eta=0, \tag{9}$$

$$(E-1)\eta+H\Delta tE_{1,\alpha}\nabla\cdot\mathbf{u}=0, \tag{10}$$

$$\left((E-1)^2+\frac{L}{2}(E^2-1)\right)\eta+FH\Delta tE_{2,\mu}\mathrm{rot}\,\mathbf{u}+(L-K)H\Delta tE_{2,\nu}\nabla\cdot\mathbf{u}-gH\Delta t^2E_{2,\delta}\nabla^2\eta=0. \tag{11}$$

2.3. Spatial discretization

2.3.1. The weak formulations. The Sobolev space $H^1(\Omega)$ denotes the space of functions in the square-integrable space $L^2(\Omega)$ whose first derivatives belong to $L^2(\Omega)$. The two following weak formulations correspond to the mixed primitive variable and GWCE cases.

DOI: 10.1002/fld

For the mixed primitive variable formulation, let η be in a subspace Q of $H^1(\Omega)$ and let each component of the velocity field belong to a subspace V of $L^2(\Omega)$. We multiply (9) and (10) by test functions $\boldsymbol{\varphi}(\mathbf{x})$ (whose x- or y-component is formally denoted by φ) and $\phi(\mathbf{x})$ belonging to V^2 and Q, respectively, and we integrate over the domain Ω to obtain

$$((E-1)+KE_{1,\beta})\int_\Omega \mathbf{u}\cdot\boldsymbol{\varphi}\,d\mathbf{x}+FE_{1,\gamma}\int_\Omega(\mathbf{k}\times\mathbf{u})\cdot\boldsymbol{\varphi}\,d\mathbf{x}+g\Delta tE_{1,\alpha}\int_\Omega\nabla\eta\cdot\boldsymbol{\varphi}\,d\mathbf{x}=0, \tag{12}$$

$$(E-1)\int_\Omega\eta\phi\,d\mathbf{x}-H\Delta tE_{1,\alpha}\int_\Omega\mathbf{u}\cdot\nabla\phi\,d\mathbf{x}=0, \tag{13}$$

where the second term in the left-hand side (LHS) of (13) has been integrated by parts by applying the periodic boundary conditions used in this paper.

For the GWCE formulation, we assume each component of $\mathbf{u}$ and η belong to Q. We multiply (9) and (11) by test functions $\boldsymbol{\phi}(\mathbf{x})$ (whose x- or y-component is formally denoted by ϕ) and $\phi(\mathbf{x})$ belonging to Q^2 and Q, respectively, and integrating over Ω yields

$$((E-1)+KE_{1,\beta})\int_\Omega \mathbf{u}\cdot\boldsymbol{\phi}\,d\mathbf{x}+FE_{1,\gamma}\int_\Omega(\mathbf{k}\times\mathbf{u})\cdot\boldsymbol{\phi}\,d\mathbf{x}+g\Delta tE_{1,\alpha}\int_\Omega\nabla\eta\cdot\boldsymbol{\phi}\,d\mathbf{x}=0, \tag{14}$$

$$\left((E-1)^2+\frac{L}{2}(E^2-1)\right)\int_\Omega\eta\phi\,d\mathbf{x}+FH\Delta tE_{2,\mu}\int_\Omega\mathrm{rot}\,\mathbf{u}\,\phi\,d\mathbf{x}$$
$$+(L-K)H\Delta tE_{2,\nu}\int_\Omega\nabla\cdot\mathbf{u}\,\phi\,d\mathbf{x}+gH\Delta t^2E_{2,\delta}\int_\Omega\nabla\eta\cdot\nabla\phi\,d\mathbf{x}=0, \tag{15}$$

where the last term in the LHS of (15) results from integration by parts using the boundary conditions.

2.3.2. Galerkin FE discretizations. The Galerkin method approximates the solutions of (12)–(13) and (14)–(15) in finite-dimensional subspaces. Consider an FE triangulation $\mathcal{T}_h$ of the polygonal domain Ω, where h is a representative meshlength parameter that measures resolution.

For the GWCE formulation we denote by Q_h the finite-dimensional subspace of Q, defined to be the set of functions whose restriction on a triangle K of $\mathcal{T}_h$ belongs to $P_1(K)$, the set of polynomials of degree 1 defined on K. The discrete solutions $\mathbf{u}_h$ and η_h belong to $Q_h\times Q_h$ and Q_h, respectively. Piecewise linear continuous interpolating functions $\phi_h(\mathbf{x})$ of degree 1 belonging to Q_h are used to approximate the velocity components and surface-elevation at triangle vertices, i.e the velocity/surface-elevation pair denoted by P_1–P_1 in the FE literature is used. We thus have

$$\mathbf{u}_h=\sum_{j\in S_K}\mathbf{u}_j\phi_j,\quad \eta_h=\sum_{j\in S_K}\eta_j\phi_j, \tag{16}$$

where S_K denotes the set of vertex nodes of K. Higher-order interpolation can be used to obtain higher convergence rates but at the expense of greater computational overhead and may or may not be useful [46].

For the mixed primitive variable formulation, the discrete solutions $\mathbf{u}_h$ and η_h sought belong to $V_h\times V_h$ and Q_h, respectively, where V_h is a finite-dimensional subspace of V. Again, V_h is defined to be the set of functions whose restriction on K belongs to $P_1(K)$, and the interpolating function $\varphi_h(\mathbf{x})$ in V_h is linear. However, $\varphi_h(\mathbf{x})$ now approximates the velocity components on the element's two-triangle support at triangle edge midpoints [14, 22, 47]. Since this particular representation of velocity is only continuous across triangle boundaries at midedge points, and discontinuous everywhere else around a triangle boundary, it is termed nonconforming (NC). The velocity/surface-elevation pair denoted by $P_1^{NC}-P_1$ in the FE literature is then used, and for φ_h in V_h and ϕ_h in Q_h we have

$$\mathbf{u}_h=\sum_{i\in M_K}\mathbf{u}_i\varphi_i,\quad \eta_h=\sum_{j\in S_K}\eta_j\phi_j, \tag{17}$$

where M_K denotes the set of midedge points (or midside nodes) of K.

DOI: 10.1002/fld

Replacing φ and ϕ by the corresponding FE test functions φ_h and ϕ_h in (12) and (13), respectively, yield the mixed primitive variable FE formulation

$$((E-1)+KE_{1,\beta})\sum_{K\in\mathscr{T}_h}\int_K \mathbf{u}_h\cdot\boldsymbol{\varphi}_h\,\mathrm{d}\mathbf{x}+FE_{1,\gamma}\sum_{K\in\mathscr{T}_h}\int_K(\mathbf{k}\times\mathbf{u}_h)\cdot\boldsymbol{\varphi}_h\,\mathrm{d}\mathbf{x}$$

$$+g\Delta tE_{1,\alpha}\sum_{K\in\mathscr{T}_h}\int_K\nabla\eta_h\cdot\boldsymbol{\varphi}_h\,\mathrm{d}\mathbf{x}=0, \tag{18}$$

$$(E-1)\sum_{K\in\mathscr{T}_h}\int_K\eta_h\,\phi_h\,\mathrm{d}\mathbf{x}-H\Delta tE_{1,\alpha}\sum_{K\in\mathscr{T}_h}\int_K\mathbf{u}_h\cdot\nabla\phi_h\,\mathrm{d}\mathbf{x}=0, \tag{19}$$

where $\mathbf{u}_h$ and η_h are defined in (17).

The GWCE FE formulation is obtained by replacing ϕ by ϕ_h in (14) and (15)

$$((E-1)+KE_{1,\beta})\sum_{K\in\mathscr{T}_h}\int_K \mathbf{u}_h\cdot\boldsymbol{\phi}_h\,\mathrm{d}\mathbf{x}+FE_{1,\gamma}\sum_{K\in\mathscr{T}_h}\int_K(\mathbf{k}\times\mathbf{u}_h)\cdot\boldsymbol{\phi}_h\,\mathrm{d}\mathbf{x}$$

$$+\Delta tE_{1,\alpha}\sum_{K\in\mathscr{T}_h}\int_K\nabla\eta_h\cdot\boldsymbol{\phi}_h\,\mathrm{d}\mathbf{x}=0, \tag{20}$$

$$\left((E-1)^2+\frac{L}{2}(E^2-1)\right)\sum_{K\in\mathscr{T}_h}\int_K\eta_h\,\phi_h\,\mathrm{d}\mathbf{x}+FH\Delta tE_{2,\mu}\sum_{K\in\mathscr{T}_h}\int_K\mathrm{rot}\,\mathbf{u}_h\,\phi_h\,\mathrm{d}\mathbf{x}$$

$$+(L-K)H\Delta tE_{2,\nu}\sum_{K\in\mathscr{T}_h}\int_K\nabla\cdot\mathbf{u}_h\,\phi_h\,\mathrm{d}\mathbf{x}+gH\Delta t^2E_{2,\delta}\sum_{K\in\mathscr{T}_h}\int_K\nabla\eta_h\cdot\nabla\phi_h\,\mathrm{d}\mathbf{x}=0, \tag{21}$$

where $\mathbf{u}_h$ and η_h are defined in (16).

Owing to the orthogonality property of the P_1^{NC} basis functions [47], the velocity mass and Coriolis matrices in (18) are 'naturally' diagonal. As shown later in Section 6.1, such a desirable and unusual property of the FE method greatly enhances computational efficiency for the mixed primitive variable FE formulation using the P_1^{NC}–P_1 pair.

3. COMPUTATION OF THE DISPERSION RELATIONS

In the continuum case, the free modes of (1) and (2) for the mixed primitive variable formulation, and of (1) and (4) for the GWCE case, are examined by perturbing about the basic state $u=v=\eta=0$. We seek solutions of (1), (2), and (4) of the form $(\widetilde{u},\widetilde{v},\widetilde{\eta})=(\widehat{u},\widehat{v},\widehat{\eta})\mathrm{e}^{\mathrm{i}(kx+ly+\omega t)}$, where k and l are the wave numbers in the x- and y-directions, respectively, and ω is the angular frequency. Substitution into (1) and (2) for the mixed primitive variable formulation, and into (1) and (4) for the GWCE case, then leads to 3×3 square matrix systems for the Fourier amplitudes $(\widehat{u},\widehat{v},\widehat{\eta})$. For a nontrivial solution to exist, the determinant of the matrix must equal zero, and this constraint leads to a relationship between the wave numbers k and l and the frequency ω, the so-called dispersion relation.

For the mixed primitive variable formulation, the resulting dispersion relation is a polynomial of degree three in ω of the form

$$\omega^3-2\mathrm{i}\tau\omega^2-[gH(k^2+l^2)+f^2+\tau^2]\omega+\mathrm{i}gH\tau(k^2+l^2)=0. \tag{22}$$

One solution is the geostrophic mode, and it would correspond to the slow Rossby mode on a β-plane, while the other two solutions correspond to the free-surface inertia-gravity modes. For the GWCE case, the dispersion relation is identical to (22) except that it yields an additional root $\omega=\mathrm{i}G$, an analytical artefact resulting from differentiation with time to obtain (4). Note that for the WCE case, i.e. $G=\tau$, we retrieve the result obtained in [28, Equation (12)].

DOI: 10.1002/fld

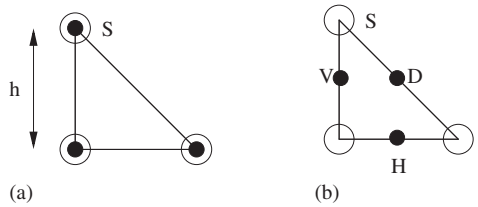

Figure 1. Typical velocity and surface-elevation node locations H, V, D, S, used in Section 3, are represented by the symbols • and ○, respectively, for the (a) $P_1 - P_1$ and (b) P_1^{NC}–P_1 pairs.

In order to compute the dispersion relation but at the discrete level, we follow the same procedure as for the continuous case. For the purpose of the following analysis we consider a uniform mesh made up of biased right isoceles triangles (as in Figure 1) and the meshlength parameter h is thus taken as a constant in the x- and y-directions. The spatially discrete operators in (18)–(19) and (20)–(21) are obtained from the stencils in [7, Figures 3.1, 3.4 and 3.5]. A set of discrete equations may then be computed at specific velocity and surface-elevation node locations.

Velocity and surface-elevation nodal unknowns are located on typical nodal sets (vertices, faces and barycenters) and because those belonging to the same set are distributed on a regular grid of size h, only selected discrete equations for each type of nodes are retained, due to symmetry reasons. For example, the discretization of (20) and (21) is obtained by using the P_1–P_1 pair and the nodal unknowns are thus only located at mesh vertices. Consequently, one discrete (vector) equation is considered for (20) and one for (21) at a typical vertex node denoted by S, shown in Figure 1(a). In the discretization of (18) and (19) using the P_1^{NC}–P_1 pair, the unknowns are located on two different sets of nodes: vertices and faces. Three discrete (vector) equations are thus considered at the three possible types of faces, denoted by H (horizontal), V (vertical) and D (diagonal), shown in Figure 1(b), and only one discrete continuity equation is retained at a typical vertex node S. For both discretizations, the typical nodes belonging to the same set are distributed on a regular grid.

As for the continuum case, the dispersion relations for the discrete schemes are found through a Fourier expansion. The discrete equations (19), (20), and (21) are first obtained at vertex node $j = S$, while (18) is computed at midside nodes $j = H, V, D$, using the stencils in [7, Figures 3.1, 3.4 and 3.5], as mentioned above. The discrete solutions corresponding to $(u_j, v_j, \eta_j) = (\widehat{u}, \widehat{v}, \widehat{\eta})\, e^{i(kx_j + ly_j)}$ are then sought, where (u_j, v_j, η_j) are the nodal unknowns that appear in the selected discrete equations at typical nodes $j = H, V, D, S$, and $(\widehat{u}, \widehat{v}, \widehat{\eta})$ are the amplitudes. The (x_j, y_j) coordinates are expressed in terms of a distance to a reference node. For both schemes, the substitution of (u_j, v_j, η_j) in the discrete equations leads to a matrix system for the Fourier amplitudes $(\widehat{u}, \widehat{v}, \widehat{\eta})$ after long and tedious algebra and only the result is given here.

3.1. The mixed primitive variable FE formulation

By using the P_1^{NC}–P_1 pair the discretization of (18) and (19) yields

$$(E - 1 + K E_{1,\beta})\widehat{u}_q - F E_{1,\gamma}\widehat{v}_q + ic\sqrt{\frac{g}{H}} E_{1,\alpha} a_{q,1}\widehat{\eta}_S = 0, \tag{23}$$

$$(E - 1 + K E_{1,\beta})\widehat{v}_q + F E_{1,\gamma}\widehat{u}_q + ic\sqrt{\frac{g}{H}} E_{1,\alpha} a_{q,2}\widehat{\eta}_S = 0, \tag{24}$$

$$(E - 1)a\widehat{\eta}_S + ic\sqrt{\frac{H}{g}} E_{1,\alpha} \sum_{q=1}^{3} (a_{q,1}\widehat{u}_q + a_{q,2}\widehat{v}_q) = 0, \tag{25}$$

DOI: 10.1002/fld

for $q=H, V, D$ in (23) and (24), and

$$a_{H,1}=2\sin\frac{kh}{2}, \qquad a_{H,2}=2\sin\frac{lh}{2}\cos\frac{(k-l)h}{2},$$
$$a_{V,1}=2\sin\frac{kh}{2}\cos\frac{(k-l)h}{2}, \qquad a_{V,2}=2\sin\frac{lh}{2},$$
$$a_{D,1}=2\sin\frac{kh}{2}\cos\frac{lh}{2}, \qquad a_{D,2}=2\sin\frac{lh}{2}\cos\frac{kh}{2},$$
$$a=1+2\cos\frac{kh}{2}\cos\frac{lh}{2}\cos\frac{(k-l)h}{2} \quad \text{with } a\geqslant\frac{3}{4} \text{ for all } kh \text{ and } lh.$$

A nontrivial solution exists for the amplitudes if the 7×7 determinant of the square matrix system (23)–(25) vanishes, which requires solving a polynomial of degree seven in E, the dispersion relation. Four roots (double conjugate), named $E^{\mathrm{MP}}_{4,5,6,7}$, where the superscript MP denotes the mixed primitive formulation, take the form of propagating inertial oscillations and have no particular spatial characteristics, i.e. they are independent of kh and lh. We obtain

$$E^{\mathrm{MP}}_{4,5,6,7}=1-\frac{F^2\gamma+K(1+K\beta)}{F^2\gamma^2+(1+K\beta)^2}\pm i\,\frac{F+FK(\beta-\gamma)}{F^2\gamma^2+(1+K\beta)^2}. \tag{26}$$

The origin of $E^{\mathrm{MP}}_{4,5,6,7}$ is a consequence of using three times more velocity nodes than surface-elevation nodes and such modes have already been encountered for the CD FD grid [9] and several FE schemes [7, 14], in the case $K=0$. The remaining three roots, named E^{MP}_j, $j=1,2,3$, are supposed to correspond to the continuous ones obtained from (22) as h and Δt tend to zero. They are computed using Maple as the solution of

$$p_3E^3+p_2E^2+p_1E+p_0=0. \tag{27}$$

3.2. The GWCE FE formulation

By using the P_1–P_1 pair, the discretization of (20) and (21) leads to

$$(E-1+KE_{1,\beta})\widehat{u}_S-FE_{1,\gamma}\widehat{v}_S+\mathrm{i}c\sqrt{\frac{g}{H}}\frac{a_2}{a_1}E_{1,\alpha}\widehat{\eta}_S=0, \tag{28}$$

$$FE_{1,\gamma}\widehat{u}_S+(E-1+KE_{1,\beta})\widehat{v}_S+\mathrm{i}c\sqrt{\frac{g}{H}}\frac{a_3}{a_1}E_{1,\alpha}\widehat{\eta}_S=0, \tag{29}$$

$$\mathrm{i}cF\sqrt{\frac{H}{g}}E_{2,\mu}\left(\frac{a_2}{a_1}\widehat{v}_S-\frac{a_3}{a_1}\widehat{u}_S\right)+\mathrm{i}c(L-K)\sqrt{\frac{H}{g}}E_{2,v}\left(\frac{a_2}{a_1}\widehat{u}_S+\frac{a_3}{a_1}\widehat{v}_S\right)$$
$$+\left((E-1)^2+\frac{L}{2}(E^2-1)+c^2\frac{a_4}{a_1}E_{2,\delta}\right)\widehat{\eta}_S=0, \tag{30}$$

with

$$a_1=\tfrac{1}{3}a, \qquad a_2=\tfrac{1}{3}(2\sin kh+\sin lh+\sin(k-l)h),$$
$$a_4=4-2\cos kh-2\cos lh, \quad a_3=\tfrac{1}{3}(2\sin lh+\sin kh-\sin(k-l)h).$$

For a nontrivial solution to exist, the 3×3 determinant of the square matrix system (28)–(30) must vanish and this leads to a polynomial of degree four in E (the dispersion relation)

$$q_4E^4+q_3E^3+q_2E^2+q_1E+q_0=0, \tag{31}$$

which is solved using Maple. Again, only three roots, named E^{GW}_j, $j=1,2,3$, where the superscript GW denotes the GWCE formulation, are supposed to correspond to the solutions of (22) for infinitesimal h and Δt parameters, whereas the fourth root, denoted by E^{GW}_4, should coincide with the analytical artefact $\omega=\mathrm{i}G$ observed earlier for the continuous model as Δt tends to zero.

DOI: 10.1002/fld

4. STABILITY/DISPERSION ANALYSIS

The solutions of the dispersion relations derived in Section 3, for both the mixed primitive variable and GWCE formulations, should be stable and well approximate the continuous modes obtained from (22). In order to compare the discrete solutions and the continuous ones in terms of stability and accuracy, Equation (22) is rewritten as

$$\Lambda^3+2K\Lambda^2+(F^2+K^2+c^2[(kh)^2+(lh)^2])\Lambda+c^2K[(kh)^2+(lh)^2]=0, \tag{32}$$

where $\Lambda=\mathrm{i}\omega\Delta t$, and $c=\sqrt{gH}\Delta t/h$ is the wave Courant number (CFL parameter). The three solutions of (32) are denoted by $\Lambda_j=\mathrm{i}\omega_j^{\mathrm{AN}}\Delta t$, and we let $E_j^{\mathrm{AN}}\equiv \mathrm{e}^{\Lambda_j}=\mathrm{e}^{\mathrm{i}\omega_j^{\mathrm{AN}}\Delta t}$, $j=1,2,3$. For the corresponding analytical artefact $\omega=\mathrm{i}G$, arising from the continuous GWCE formulation, we also let $E_4^{\mathrm{AN}}=\mathrm{e}^{-G\Delta t}=\mathrm{e}^{-L}$.

4.1. Stability of the discrete mixed primitive variable and GWCE formulations

To ensure that the discrete mixed primitive variable and GWCE formulations are stable, the dispersion relations (27) and (31) should have roots E of magnitude less than one or non-multiple roots of magnitude less than or equal to one, for all wavelengths. In order to guarantee that such a constraint is satisfied, a first approach is to solve the dispersion relations analytically for E, and then determine sufficient and necessary conditions on the meshlength, the time step, and the weighting parameters.

Practically, such an approach is only possible for the roots $E_{4,5,6,7}^{\mathrm{MP}}$ in (26) in the case of the mixed primitive variable formulation using the P_1^{NC}–P_1 scheme, and for a few roots of (31) in the case of the GWCE formulation, with $F=0$.

For the modes $E_{4,5,6,7}^{\mathrm{MP}}$ to be stable, i.e. $|E_{4,5,6,7}^{\mathrm{MP}}|<1$, it is sufficient that

$$F^2(2\gamma-1)+K^2(2\beta-1)+2K>0, \tag{33}$$

a condition which is less restrictive than the constraint $\gamma>\frac{1}{2}$ obtained in [14] when $K=0$. Indeed, F and K are usually much less than one in practice and (33) is satisfied for a wide range of parameters β and γ.

In the case of the GWCE formulation with $F=0$, the root

$$E=1-\frac{K}{1+\beta K}, \tag{34}$$

is a solution of (31) and it coincides with $E_{4,5,6,7}^{\mathrm{MP}}$ in (26) for $F=0$. However, $E_{4,5,6,7}^{\mathrm{MP}}$ in (26) is not a root of (31) for $F\neq 0$. The solution E in (34) is stable for all β in the interval $[0,1]$ provided $K(1-2\beta)\leqslant 2$, which is usually the case in practice. For infinitesimal time stepping, we obtain $E=1+\mathrm{i}\omega\Delta t+O(\Delta t^2)$ and $1-K/(1+\beta K)=1-\tau\Delta t+O(\Delta t^2)$ in (34). Hence, the discrete frequency ω asymptotes to $\mathrm{i}\tau$, and in the WCE case, i.e. for $G=\tau$, it coincides with the analytical artefact $\omega=\mathrm{i}G$ arising from the continuous GWCE formulation.

Further, when $G=\tau$ (the WCE case) and $F=0$, the roots E_1^{GW} and E_2^{GW} in (31) are solutions of the polynomial

$$q_{2,0}E^2+q_{1,0}E+q_{0,0}=0, \tag{35}$$

where

$$q_{0,0}=a_1(2-K)+2c^2a_4\delta, \qquad q_{2,0}=a_1(2+K)+2c^2a_4\delta,$$
$$q_{1,0}=2(c^2a_4(1-2\delta)-2a_1),$$

while $E_3^{\mathrm{GW}}=E_4^{\mathrm{GW}}$ coincide with E in (34). Determining the stability conditions of E_1^{GW} and E_2^{GW} in (35) is a tedious analytical exercise, even in the case of a second degree polynomial, and the first approach reaches its limits.

DOI: 10.1002/fld

A second approach is to use the Routh–Hurwitz criterion or the Liénard–Chipart modification thereof, which prescribes algebraic conditions for the coefficients of a given polynomial in order to determine whether or not the roots of the polynomial lie inside the unit circle and are stable [48]. Such conditions are described in [33] for polynomials of degree one to five, and simpler conditions which are only necessary for stability are also provided.

For $G=\tau$ (the WCE case) and $F=0$, the required conditions for stability can be derived by substituting $E=(\widetilde{E}+1)/(\widetilde{E}-1)$ in (35), and this leads to a polynomial in $\widetilde{E}$ of the form $\sum_{j=0}^{2}\widetilde{q}_{j,0}\widetilde{E}^j$ where

$$\widetilde{q}_{0,0}=q_{0,0}-q_{1,0}+q_{2,0}=4\left(a_1+c^2a_4\left(\delta-\tfrac{1}{4}\right)\right),\qquad \widetilde{q}_{2,0}=q_{0,0}+q_{1,0}+q_{2,0}=c^2a_4,$$
$$\widetilde{q}_{1,0}=2(q_{2,0}-q_{0,0})=2a_1K.$$

We obtain from [33, Table 4.2] that necessary and sufficient conditions for stability, i.e. $|E|<1$, are $\widetilde{q}_{j,0}>0$, $j=0,1,2$ or equivalently $\widetilde{q}_{j,0}<0$, $j=0,1,2$. Since we have $\frac{1}{3}\leqslant a_1<1$ and $0<a_4\leqslant 8$ for $-\pi\leqslant kh, lh\leqslant\pi$, as the wavenumbers corresponding to an infinite wavelength are disregarded, we deduce $\widetilde{q}_{1,0}>0$ and $\widetilde{q}_{2,0}>0$. The necessary and sufficient condition for stability then reads $\widetilde{q}_{0,0}>0$, which is trivially satisfied for $\delta>\frac{1}{4}$, and true for $\delta<\frac{1}{4}$ if $c<1/\sqrt{6(1-4\delta)}$. For the particular case $\delta=0$, the latter condition coincide with the result obtained in [28, p. 377] for the explicit scheme, i.e. $c<1/\sqrt{6}$.

In the more general case $G\neq\tau$ (GWCE) and $F\neq 0$, the Liénard–Chipart criterion is used again, but the values of α,β and γ have to be prescribed in order to allow tractable calculations. As in [28], attention is restricted to second-order approximations in time and we choose $\alpha=\beta=\gamma=\frac{1}{2}$. The required conditions for stability are obtained by letting $E=(\widetilde{E}+1)/(\widetilde{E}-1)$ in (31) yielding a polynomial in $\widetilde{E}$ of the form $\sum_{j=0}^{4}\widetilde{q}_j\widetilde{E}^j$ with

$$\widetilde{q}_0=q_4-q_3+q_2-q_1+q_0,\qquad \widetilde{q}_3=2(2q_4+q_3-q_1-2q_0),$$
$$\widetilde{q}_1=2(2q_4-q_3+q_1-2q_0),\qquad \widetilde{q}_4=q_4+q_3+q_2+q_1+q_0,$$
$$\widetilde{q}_2=2(3q_4-q_2+3q_0),$$

After long algebra we obtain

$$\widetilde{q}_0=16a_1^2+4c^2a_1a_4(4\delta-1),$$
$$\widetilde{q}_1=8a_1^2(L+2K)+4c^2a_1a_4K(4\delta-1)+2c^2(a_2^2+a_3^2)(L-K)(4\nu-1),$$
$$\widetilde{q}_2=4a_1^2(F^2+2KL+K^2)+c^2a_1a_4(4+(L^2+K^2)(4\delta-1))$$
$$+c^2(a_2^2+a_3^2)(K(L-K)(4\nu-1)+F^2(1-4\mu)),$$
$$\widetilde{q}_3=2a_1^2L(F^2+K^2)+2c^2(2a_1a_4-a_2^2-a_3^2)K+2c^2(a_2^2+a_3^2)L,$$
$$\widetilde{q}_4=c^2(a_2^2+a_3^2)KL+c^2(a_1a_4-a_2^2-a_3^2)(F^2+K^2).$$

The set of necessary and sufficient conditions for stability, in the case of a quartic polynomial [33, Table 4.4], is again obtained by examining the sign of $\widetilde{q}_j$, $j=0,1,2,3,4$, but in addition, it requires to determine the sign of the quantity $\Delta_3=\widetilde{q}_1\widetilde{q}_2\widetilde{q}_3-\widetilde{q}_1^2\widetilde{q}_4-\widetilde{q}_3^2\widetilde{q}_0$. However, because the expression of Δ_3 contains more than 500 terms after simplification using Maple, the calculation becomes intractable and only necessary conditions are examined in the following.

Necessary conditions for stability are $\widetilde{q}_j>0$, $j=0,1,\ldots,4$ or equivalently $\widetilde{q}_j<0$, $j=0,1,\ldots,4$. Because $2a_1a_4-a_2^2-a_3^2>0$ and $a_1a_4-a_2^2-a_3^2>0$ for $-\pi\leqslant kh, lh\leqslant\pi$, we have $\widetilde{q}_3>0$ and $\widetilde{q}_4>0$, and the necessary conditions reduced to $\widetilde{q}_j>0$, $j=0,1,2$. These are obtained in Table I taking into account that $0\leqslant a_2^2+a_3^2\leqslant 2$. For a given j, $j=0,1,2$, condition A or condition B has to be fulfilled in order to have $\widetilde{q}_j>0$, $j=0,1,2$.

DOI: 10.1002/fld

Table I. Necessary stability conditions to guarantee $\widetilde{q}_j>0$, $j=0,1,2$, in the case $\alpha=\beta=\gamma=\frac{1}{2}$. For a given j, $j=0,1,2$, condition A or condition B has to be fulfilled in order to have $\widetilde{q}_j>0$.

	Condition A	Condition B
$\widetilde{q}_0>0$	$\delta\geqslant\frac{1}{4}$	$\delta<\frac{1}{4}$ and $c<1/\sqrt{6(1-4\delta)}$
$\widetilde{q}_1>0$	$\delta\geqslant\frac{1}{4}$	• $(L-K)(4\nu-1)>0$ and $\delta<\frac{1}{4}$ and $c<\sqrt{\frac{L+2K}{12K(1-4\delta)}}$
		or
	and	• $(L-K)(4\nu-1)<0$ and
		◦ either $\delta<\frac{1}{4}$ and $c<\sqrt{\frac{L+2K}{12K(1-4\delta)+\frac{9}{2}(L-K)(1-4\nu)}}$
	$(L-K)(4\nu-1)\geqslant0$	◦ or $\delta>\frac{1}{4}$ and $c<\frac{1}{3}\sqrt{\frac{2(L+2K)}{(L-K)(1-4\nu)}}$
$\widetilde{q}_2>0$	$(L-K)(4\nu-1)\geqslant0$	$(L-K)(4\nu-1)<0$ and
	and	◦ either $\mu<\frac{1}{4}$ and $c<\frac{1}{3}\sqrt{\frac{2(F^2+2KL+K^2)}{K(L-K)(1-4\nu)}}$
	$\mu\leqslant\frac{1}{4}$	◦ or $\mu>\frac{1}{4}$ and $c<\frac{1}{3}\sqrt{\frac{2(F^2+2KL+K^2)}{F^2(4\mu-1)+K(L-K)(1-4\nu)}}$

The necessary stability conditions of Table I reduce to one of the four following cases:

- *Case 1*: $\delta<\frac{1}{4}$ and $(L-K)(4\nu-1)>0$ and $\mu\leqslant\frac{1}{4}$ and $c<1/\sqrt{6(1-4\delta)}$,
- *Case 2*: $\delta<\frac{1}{4}$ and $(L-K)(4\nu-1)<0$ and
  either $\mu<\frac{1}{4}$ and

$$c<\min\left(\sqrt{\frac{L+2K}{12K(1-4\delta)+\frac{9}{2}(L-K)(1-4\nu)}},\frac{1}{3}\sqrt{\frac{2(F^2+2KL+K^2)}{K(L-K)(1-4\nu)}}\right)$$

  or $\mu>\frac{1}{4}$ and

$$c<\min\left(\sqrt{\frac{L+2K}{12K(1-4\delta)+\frac{9}{2}(L-K)(1-4\nu)}},\frac{1}{3}\sqrt{\frac{2(F^2+2KL+K^2)}{F^2(4\mu-1)+K(L-K)(1-4\nu)}}\right),$$

- *Case 3*: $\delta\geqslant\frac{1}{4}$ and $(L-K)(4\nu-1)\geqslant0$ and $\mu\leqslant\frac{1}{4}$,
- *Case 4*: $\delta\geqslant\frac{1}{4}$ and $(L-K)(4\nu-1)<0$ and
  either $\mu<\frac{1}{4}$ and

$$c<\min\left(\frac{1}{3}\sqrt{\frac{2(L+2K)}{(L-K)(1-4\nu)}},\frac{1}{3}\sqrt{\frac{2(F^2+2KL+K^2)}{K(L-K)(1-4\nu)}}\right)$$

  or $\mu>\frac{1}{4}$ and

$$c<\min\left(\frac{1}{3}\sqrt{\frac{2(L+2K)}{(L-K)(1-4\nu)}},\frac{1}{3}\sqrt{\frac{2(F^2+2KL+K^2)}{F^2(4\mu-1)+K(L-K)(1-4\nu)}}\right).$$

Note that we have $(4+(L^2+K^2)(4\delta-1))>0$ in the expression of $\widetilde{q}_2$, because $F<1$ and $K<1$ in practice. This explains why the stability conditions do not include δ in Table I for $\widetilde{q}_2$. Further, the case $(L-K)(4\nu-1)>0$ and $\mu>\frac{1}{4}$ is not examined for $\widetilde{q}_2$ in Table I, as it leads to unconditional instability. Finally, the stability conditions presented in Table I generalize previous results obtained for special cases in [27, 28, 34, 40, 41, 43, 44].

In the general case $F\neq0$, the Liénard–Chipart criterion can also be used to determine whether or not the mixed primitive variable formulation employing the P_1^{NC}–P_1 FE pair is stable. As for

DOI: 10.1002/fld

the GWCE scheme we let $\alpha=\beta=\gamma=\frac{1}{2}$ and $E=(\widetilde{E}+1)/(\widetilde{E}-1)$ is now substituted in (27). Such a procedure yields a polynomial in $\widetilde{E}$ of the form $\sum_{j=0}^{3}\widetilde{p}_j\widetilde{E}^j$ with

$$\widetilde{p}_0=p_3-p_2+p_1-p_0=8a,$$

$$\widetilde{p}_1=3p_3-p_2-p_1+3p_0=8aK,$$

$$\widetilde{p}_2=3p_3+p_2-p_1-3p_0=2c^2\sum_{j=1}^{2}(a_{H,j}^2+a_{V,j}^2+a_{D,j}^2)+2a(F^2+K^2),$$

$$\widetilde{p}_3=p_3+p_2+p_1+p_0=c^2K\sum_{j=1}^{2}(a_{H,j}^2+a_{V,j}^2+a_{D,j}^2),$$

$$\Delta_2=\widetilde{p}_1\widetilde{p}_2-\widetilde{p}_0\widetilde{p}_3=8c^2aK\sum_{j=1}^{2}(a_{H,j}^2+a_{V,j}^2+a_{D,j}^2)+16a^2K(F^2+K^2).$$

The set of necessary and sufficient conditions for stability, in the case of a cubic polynomial [33, Table 4.3], is obtained by examining the sign of $\widetilde{p}_j$, $j=0,1,2,3$, and Δ_2. Because $a\geqslant\frac{3}{4}$ for all kh and lh, we deduce that $\widetilde{p}_j>0$, $j=0,1,2,3$, and $\Delta_2>0$ for all kh and lh and hence the scheme is unconditionally stable.

4.2. Accuracy of the discrete mixed primitive variable and GWCE formulations

Although a discrete scheme may provide stable solutions, this does not necessarily mean that the solutions are accurate. In order to evaluate the ability of a numerical scheme to accurately represent the amplitude and phase speed of a wave, the complex propagation factor has been introduced by Leendertse [49]. By using the notations of the present study, the propagation factor can be define as $T_j^{\mathrm{MP}}=(E_j^{\mathrm{MP}}/E_j^{\mathrm{AN}})^{\Upsilon}$ and $T_j^{\mathrm{GW}}=(E_j^{\mathrm{GW}}/E_j^{\mathrm{AN}})^{\Upsilon}$, $j=1,2,3,4$, for the mixed primitive variable and GWCE formulations, respectively, where $\Upsilon=2\pi/(\omega_j^{\mathrm{AN}}\Delta t)=2\mathrm{i}\pi/\Lambda_j$, $j=1,2,3,4$, is the number of time steps required for the analytical wave to propagate one wavelength. The propagation factor is the ratio of the computed wave over the analytical one after the time required for the latter to propagate one wavelength. The magnitude of T represents the relative change in the wave amplitude due to the discrete approximation, while the argument of T is the phase lead or lag of the computed wave compared with the analytical one.

The computation of the propagation factor leads to large errors when evaluating $E_j^{\mathrm{MP}}/E_j^{\mathrm{AN}}$ and $E_j^{\mathrm{GW}}/E_j^{\mathrm{AN}}$ at the power Υ, even by using Maple with 40 digits. Consequently, in order to adequately measure the accuracy of the discrete mixed primitive variable and GWCE formulations, we consider the complex ratios

$$R_j^{\mathrm{MP}}=\frac{E_j^{\mathrm{MP}}}{E_j^{\mathrm{AN}}}\quad\text{and}\quad R_j^{\mathrm{GW}}=\frac{E_j^{\mathrm{GW}}}{E_j^{\mathrm{AN}}},\quad j=1,2,3,4,$$

in the present work, and compute the amplitude and argument of R_j^{MP} and R_j^{GW}, $j=1,2,3,4$.

We set $\alpha=\beta=\gamma=\frac{1}{2}$ for both the mixed primitive (using the P_1^{NC}–P_1 pair) and GWCE formulations, and hence the P_1^{NC}–P_1 scheme is unconditionally stable. As shown in Section 4.1, the stability of the GWCE formulation depends on the choice of the parameters K, L, c, δ, ν, and μ and two cases corresponding to explicit and implicit cases are investigated. Further, in both cases, we allow the parameter L to vary from K to $10K$, as suggested in [34].

4.2.1. The explicit case. For the explicit case, where $\delta=\mu=0$ and $\nu=0$, the amplitude ratio $|R_j|$, $j=1,2,3,4$, which corresponds either to $|R_j^{\mathrm{MP}}|$ or $|R_j^{\mathrm{GW}}|$, $j=1,2,3,4$, is plotted in Figure 2 as a surface function for the mixed primitive variable (using the P_1^{NC}–P_1 FE pair) and GWCE formulations. The minimum and maximum values of the surface function are specified at the bottom

DOI: 10.1002/fld

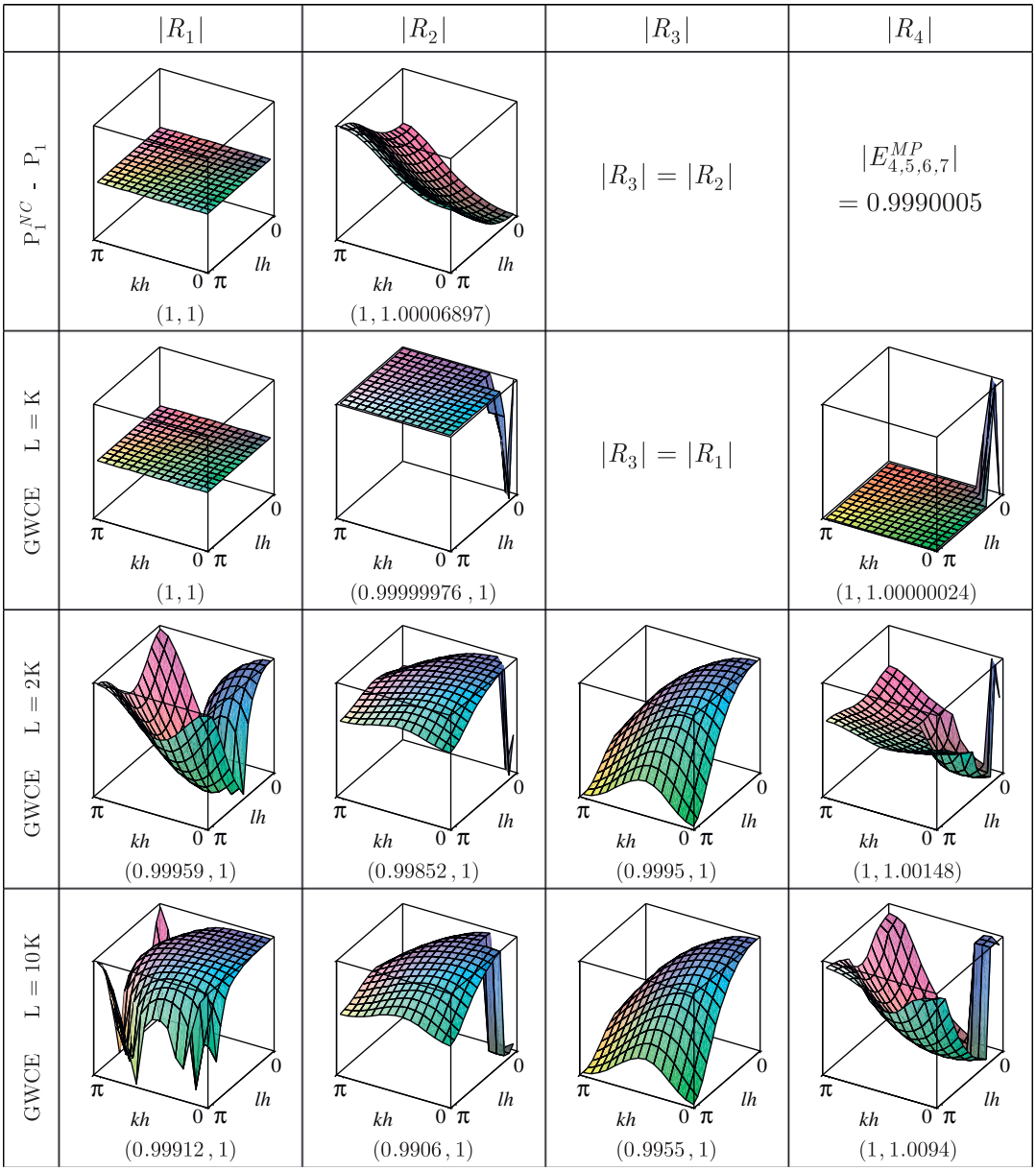

Figure 2. The amplitude ratio $|R_j|$, $j=1,2,3,4$, which corresponds either to $|R_j^{\mathrm{MP}}|$ (the mixed primitive variable case using the P_1^{NC}–P_1 FE pair) or $|R_j^{\mathrm{GW}}|$ (the GWCE formulation case), $j=1,2,3,4$, is plotted as a surface function for $\alpha=\beta=\gamma=\frac{1}{2}$. The explicit case is considered with $\delta=\mu=\nu=0$, $c=0.2$, $K=F=10^{-3}$. The minimum and maximum value of the surface function are specified at the bottom of each panel for $0\leqslant kh\leqslant\pi$ and $0\leqslant lh\leqslant\pi$.

of each panel for $0\leqslant kh\leqslant\pi$ and $0\leqslant lh\leqslant\pi$. The Coriolis and bottom friction parameters are selected on the basis of physical relevance and we choose $f=\tau=10^{-4}$. Indeed, the Coriolis parameter varies very smoothly since we have $f=2\Omega\sin\psi_0$, where $\Omega=7.29\times10^{-5}\,\mathrm{rad\,s^{-1}}$ is the angular frequency of the Earth's rotation and ψ_0 is the latitude. Consequently, we obtain $f=10^{-4}\,\mathrm{s^{-1}}$ at midlatitudes for $\psi_0\simeq45^\circ$. The choice of τ in (1), corresponding to lakes and shallow estuaries, is obtained by rewriting $\tau\tilde{\mathbf{u}}$ as $\tau\tilde{\mathbf{u}}=C_d|\tilde{\mathbf{u}}|\tilde{\mathbf{u}}/H$, where C_d is a bottom drag coefficient, to obtain the quadratic friction approximation. For test problems, C_d is commonly of the order 2.5×10^{-3}, $|\tilde{\mathbf{u}}|=0.4\,\mathrm{ms^{-1}}$ and $H=10\,\mathrm{m}$, which leads to $\tau=10^{-4}\,\mathrm{s^{-1}}$, and τ and f are thus of the same

DOI: 10.1002/fld

order of magnitude. For the explicit case we choose $\Delta t = 10\,\mathrm{s}$ and $h = 500\,\mathrm{m}$, and we thus deduce $c = \sqrt{gH}\Delta t/h \simeq 0.2$ and $F = K = 10^{-3}$.

For the parameters employed in the explicit case, $|E_j^{\mathrm{MP}}|$ and $|E_j^{\mathrm{GW}}|$, $j = 1, 2, 3, 4$, are less than 1 and the mixed primitive variable and GWCE schemes are stable. We observe that $|R_1| = 1$ in Figure 2 for the P_1^{NC}–P_1 and GWCE (with $L = K$) cases and hence, $|E_1^{\mathrm{MP}}| = |E_1^{\mathrm{GW}}| = |E_1^{\mathrm{AN}}|$ for all wavelengths, i.e. the discrete and continuous amplitudes exactly coincide for this root (in absolute value) and for both formulations. The examination of $|R_2|$ reveals a very different behaviour of the surface function near the origin $kh = lh = 0$ for the P_1^{NC}–P_1 and GWCE (with $L = K$) schemes. Indeed, as the curve for $|R_2^{\mathrm{MP}}|$ smoothly increases in the vicinity of $kh = lh = 0$, the $|R_2^{\mathrm{GW}}|$ curve exhibits an abrupt variation close to the origin. However, because these changes have very small amplitudes, no damping of the solution is expected. For $L = 2K$ and $L = 10K$, $|R_1^{\mathrm{GW}}|$ and $|R_3^{\mathrm{GW}}|$ smoothly decrease from 1 in the vicinity of the origin while $|R_2^{\mathrm{GW}}| < 1$ at $kh = lh = 0$ and hence slightly damped solutions may be expected for medium term simulations for the GWCE formulation. Finally, the artefact $|R_4^{\mathrm{GW}}|$ remains close to 1 near the origin and the root $|E_4^{\mathrm{GW}}|$ would not be expected to cause problems.

The arguments of R_j, $j = 1, 2, 3, 4$, are examined in Figure 3. For the P_1^{NC}–P_1 scheme and the arguments of R_1^{GW} and R_4^{GW}, the observed phase differences are close to zero and good phase accuracy is expected. The arguments of R_2^{GW} and R_3^{GW} are also close to 0 at the origin $kh = lh = 0$, but they rapidly decrease (or increase) to -1.912522 (or 1.912522) at $kh = lh = \pi$, i.e. $\pm 110^{\circ}$, leading to a phase difference for small to medium range wavenumbers. Note that as the ratio L/K increases, $\arg(R_2)$ and $\arg(R_3)$ are zero over a slightly increasing area near the origin.

For $\delta = \mu = 0$, $K \leqslant L \leqslant 10K$, and $0 < v \leqslant 1$ very similar results have been found for R_j^{GW} and $\arg(R_j^{\mathrm{GW}})$, $j = 1, 2, 3, 4$, compared with those obtained for $v = 0$ in Figures 2 and 3, provided $c < 1/\sqrt{6}$. Indeed, the results are nearly identical at $kh = lh = 0$ and $kh = lh = \pi$, and only minor differences occur inside the domain. However, because the stability condition $c < 1/\sqrt{6}$, obtained by setting $\delta = 0$ in $c < 1/\sqrt{6(1-4\delta)}$ (Case 1), is slightly more constraining (for realistic values of F, K, and L) than the necessary condition

$$c < \min\left(\sqrt{\frac{L+2K}{12K(1-4\delta)+\frac{9}{2}(L-K)(1-4v)}},\quad \frac{1}{3}\sqrt{\frac{2(F^2+2KL+K^2)}{K(L-K)(1-4v)}}\right)$$

obtained in Case 2, it is suggested that the necessary and sufficient condition for stability in Cases 1 and 2 is $c < 1/\sqrt{6(1-4\delta)}$. Further, when instability occurs, i.e. for $c > 1/\sqrt{6(1-4\delta)}$, the roots E_2 and E_3 are primarily affected.

4.2.2. The implicit case. For the implicit case, we choose $\delta = \frac{1}{2}$, $v = \frac{1}{2}$, $\mu = 0$, and we have $c = 2$ and $K = F = 10^{-2}$ (with $\tau = f = 10^{-4}\,\mathrm{s}^{-1}$) using $H = 10\,\mathrm{m}$, $h = 500\,\mathrm{m}$, and $\Delta t = 100\,\mathrm{s}$. Again, $|E_j^{\mathrm{MP}}|$ and $|E_j^{\mathrm{GW}}|$ are less than 1 for the parameters employed in the implicit case and the P_1^{NC}–P_1 and GWCE schemes are stable. Small amplitude variations around 1 are observed in Figure 4 for $|R_1^{\mathrm{MP}}|$ and $|R_1^{\mathrm{GW}}|$ close to the origin $kh = lh = 0$, and hence the corresponding solutions E_1^{MP} and E_1^{GW} should not suffer from damping. This is likely to be the case for E_2^{MP}, E_2^{GW} (with $L = K$), and E_3^{GW} (whatever the ratio L/K) since $|R_2^{\mathrm{MP}}|$, $|R_2^{\mathrm{GW}}|$ (with $L = K$) and $|R_3^{\mathrm{GW}}|$ are equal to 1 at $kh = lh = 0$ and then slightly increase from 1 up to roughly 1.005 at $kh = lh = \pi$. However, the situation is very different for E_2^{GW} with $L > K$ because we have $|R_2^{\mathrm{GW}}| = 0.990$ (with $L = 2K$) and $|R_2^{\mathrm{GW}}| = 0.914$ (with $L = 10K$) at the origin $kh = lh = 0$ and the GWCE solution will be damped as the ratio L/K increases. As for the explicit case, the artefact $|R_4^{\mathrm{GW}}|$ remains close to 1 near the origin and the root $|E_4^{\mathrm{GW}}|$ should not be a source of difficulties. Note that the small-scale noise observed for $|R_1^{\mathrm{GW}}|$ and $|R_4^{\mathrm{GW}}|$ when $L = K$ slightly decreases for $v = \frac{1}{4}$ (instead of $\frac{1}{2}$).

The arguments of R_j, $j = 1, 2, 3, 4$, are examined in Figure 5 for the implicit case. As for the explicit case, the arguments of R_1 and R_4 for both formulations are close to zero and consequently, good phase accuracy is expected. The arguments of R_2 and R_3 are also close to 0 at the

DOI: 10.1002/fld

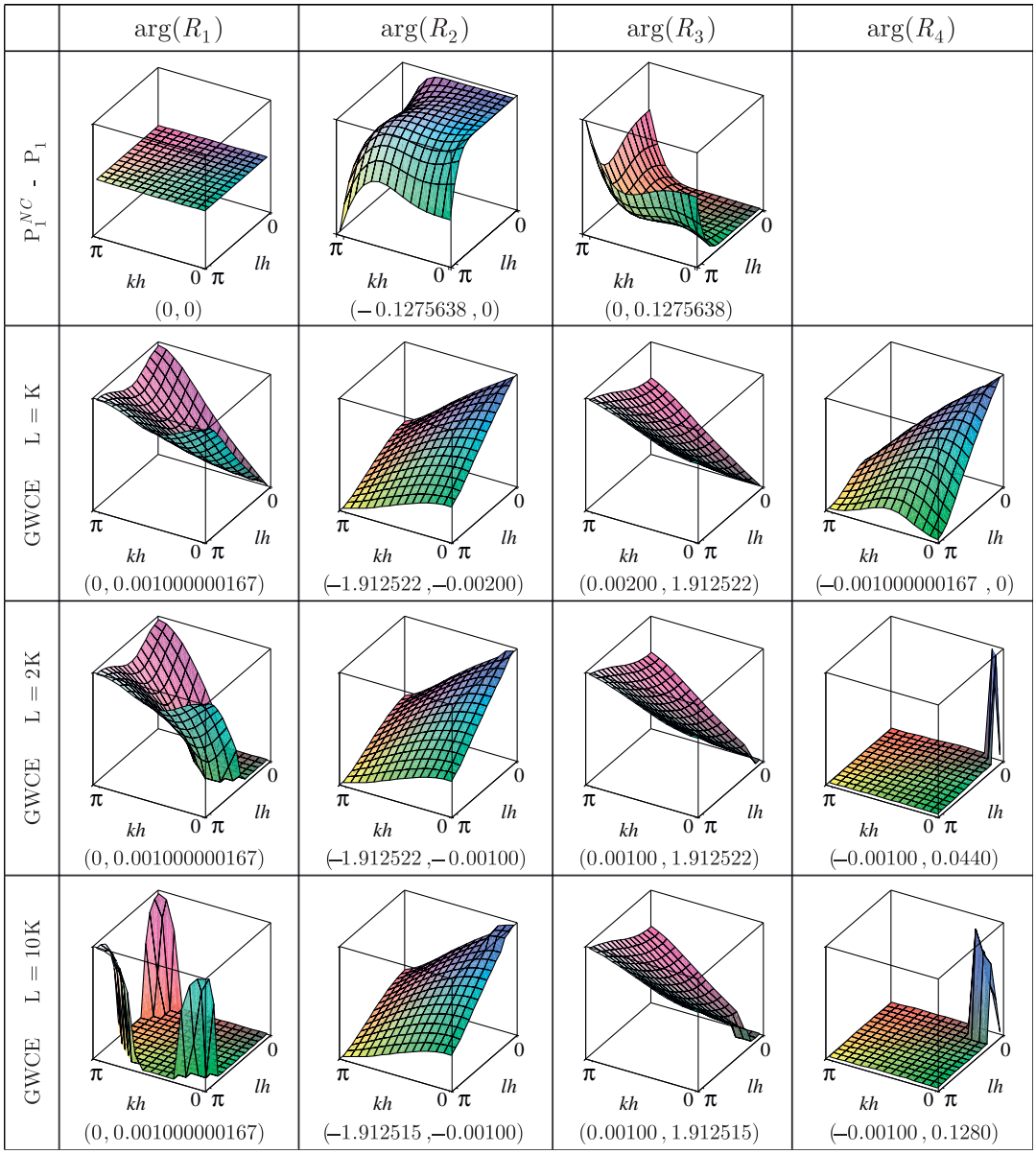

Figure 3. As for Figure 2 but for the argument (in radians) of R_j, $j=1,2,3,4$.

origin $kh=lh=0$, however, they rapidly decrease (or increase) to $-\pi$ (or π) and significant phase difference may be expected compared with the continuous solution, particularly for the GWCE formulation for which the decrease (or increase) is more rapid than in the P_1^{NC}–P_1 case. Finally, the variation of L with K has a minor impact on the phase, contrary to the explicit case.

Calculations of $|R_j^{\mathrm{GW}}|$, $j=1,2,3,4$, have shown that the necessary conditions of Case 3 are likely sufficient conditions. In Case 4, with $\delta=\frac{1}{2}$ and $v=0$ two choices of μ are possible. For $\mu<\frac{1}{4}$, the necessary condition for stability

$$c<\min\left(\frac{1}{3}\sqrt{\frac{2(L+2K)}{L-K}},\frac{1}{3}\sqrt{\frac{2(F^2+2KL+K^2)}{K(L-K)}}\right)$$

DOI: 10.1002/fld

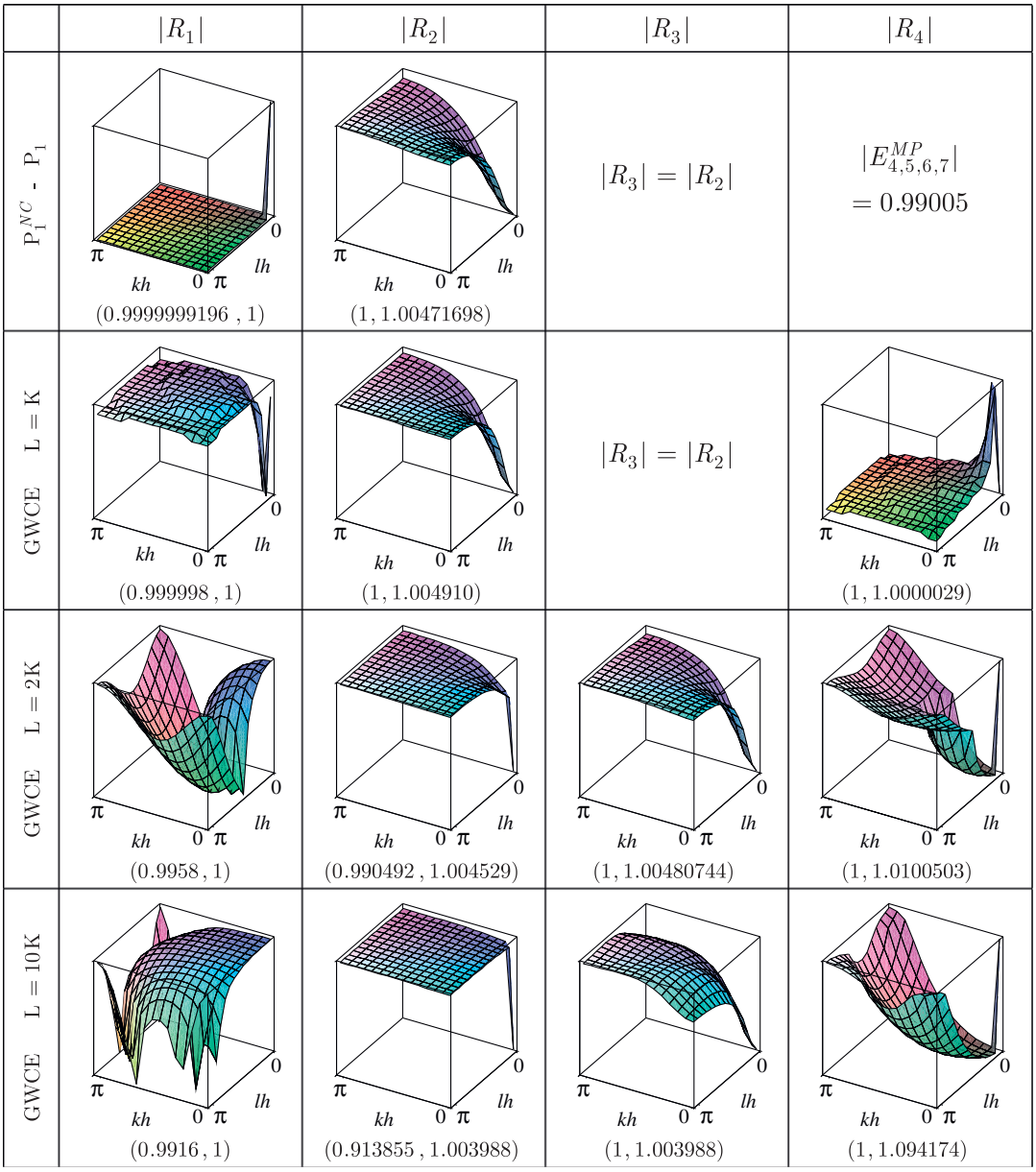

Figure 4. As for Figure 2 but for the implicit case with $\delta = v = \frac{1}{2}$, $\mu = 0$, $c = 2$, $K = F = 10^{-2}$.

yields $c < \min(0.943, 1.155)$ for $L = 2K$ and $c < \min(0.544, 0.737)$ for $L = 10K$, with $F = K$ as it is the case in this study. However, we obtain $|E_2^{\mathrm{GW}}| > 1$ and $|E_3^{\mathrm{GW}} > 1$ as soon as $c > 0.68$ for $L = 2K$ and $c > 0.39$ for $L = 10K$. For $\mu > \frac{1}{4}$, the necessary condition for stability

$$c < \min\left(\frac{1}{3}\sqrt{\frac{2(L+2K)}{L-K}}, \frac{1}{3}\sqrt{\frac{2(F^2+2KL+K^2)}{F^2(4\mu-1)+K(L-K)}}\right)$$

leads to $c < \min(0.943, 0.817)$ for $L = 2K$ and $c < \min(0.544, 0.699)$ for $L = 10K$, with $F = K$ and $\mu = \frac{1}{2}$. Again the calculations show that the roots E_2^{GW} and E_3^{GW} are greater than 1 as soon as $c > 0.68$ for $L = 2K$ and $c > 0.39$ for $L = 10K$, i.e. the same stability conditions than for the case $\mu < \frac{1}{4}$. This suggests that necessary conditions found for stability in Case 4 are slightly less

DOI: 10.1002/fld

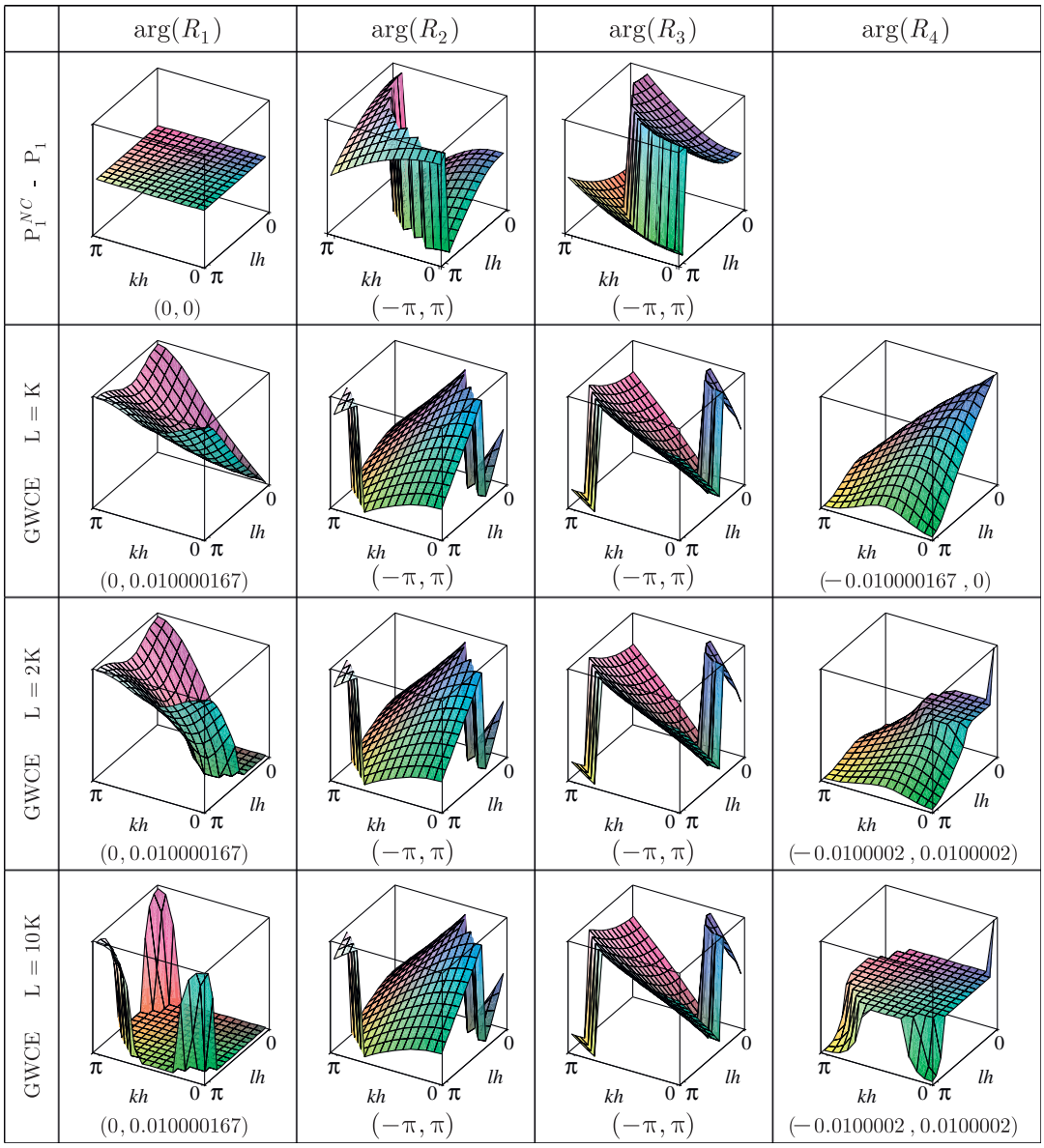

Figure 5. As for Figure 3 but for the implicit case.

restrictive than the necessary and sufficient conditions on $|R_j^{\mathrm{GW}}|$, $j = 1, 2, 3, 4$. Finally, note that the stability constraints obtained for the explicit and implicit cases are close, particularly for large L.

5. NUMERICAL SIMULATIONS

We have shown in Section 4 that the GWCE solution is expected to be out of phase compared with the P_1^{NC}–P_1 results in the explicit case, because $\arg(R_2^{\mathrm{GW}})$ and $\arg(R_3^{\mathrm{GW}})$ rapidly depart from 0 as kh and lh increase in Figure 3. Further, the GWCE results should be damped as the ratio L/K increases. Such a damping is moderate in the explicit case but it appears much more severe in the implicit one (see $|R_2^{\mathrm{GW}}|$ in Figures 2 and 4). Finally, phase difference may be expected for both schemes compared with the continuous solution in the implicit case, affecting more severely the GWCE scheme (see $\arg(R_2)$ and $\arg(R_3)$ in Figure 5).

DOI: 10.1002/fld

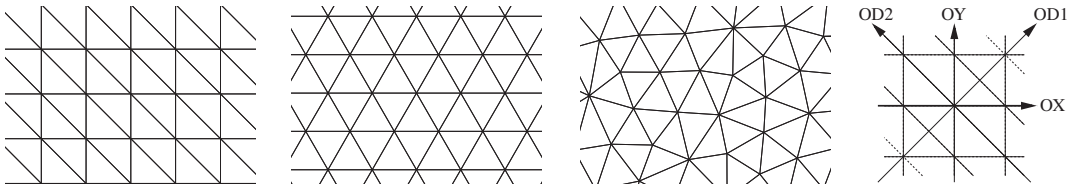

Figure 6. A window of meshes 1, 2 and 3, made up of: (a) biased right isoceles triangles; (b) equilateral triangles; (c) unstructured triangles with smoothing; and (d) definition of the directions OX, OY, OD1, and OD2 on Mesh 1.

Table II. Minimum and maximum values of the numerically simulated surface elevations (in meters) for the $P_1^{NC}-P_1$ and GWCE schemes in the explicit case at different stages of the propagation, up to $t=1000$ s, on Mesh 1 for $\xi=5.5h$.

		GWCE		
Time (s)	$P_1^{NC}-P_1$	$L=K$	$L=2K$	$L=10K$
200	0, 0.14970	0, 0.14766	0, 0.14764	0, 0.14751
400	−0.13368, 0.10771	−0.12843, 0.10901	−0.12811, 0.10891	−0.12611, 0.10823
600	−0.07523, 0.09086	−0.07281, 0.09289	−0.07268, 0.09278	−0.07196, 0.09201
800	−0.05454, 0.07980	−0.05296, 0.08194	−0.05292, 0.08182	−0.05262, 0.08099
1000	−0.04448, 0.07172	−0.04307, 0.07390	−0.04304, 0.07378	−0.04279, 0.07291

In order to validate such results, the propagation and dispersion of gravity waves are simulated in a square basin. The square domain extent is $40\,\text{km}\times 40\,\text{km}$ and it is first discretized on Mesh 1, made up of biased right isoceles triangles, shown in Figure 6(a), with a uniform node spacing $h=470$ m. Note that Mesh 1 corresponds to the mesh used in Sections 3 and 4 for the computation and analysis of the dispersion relations. Consequently, due to the smallness of the domain we set $f=0$ and hence $F=0$. A flat bottom with mean depth $H=10$ m is assumed, and $\tau=10^{-4}\,\text{s}^{-1}$. The explicit case is run with a time step of 10 s leading to $c=\sqrt{gH}\Delta t/h\simeq 0.2$. For the implicit case, we choose $\Delta t=100$ s and hence $c\simeq 2$. The parameters used in the numerical experiment thus correspond to those employed for the explicit and implicit cases in Section 4. The initial conditions are $\widetilde{\mathbf{u}}(\mathbf{x},t=0)=\mathbf{0}$ and $\widetilde{\eta}(\mathbf{x},t=0)=\sigma e^{-\zeta(x^2+y^2)}$. We let $\sigma=0.5$ m and hence the initial perturbation amplitude represents 5% of the total depth. The e-folding radius of the initial Gaussian, noted by ξ in the following, is defined as the distance from the origin for which $\widetilde{\eta}=\sigma e^{-1}$. By choosing $\zeta=1.6\times 10^{-7}\,\text{m}^{-2}$, we have $\xi=\zeta^{-1/2}=2.5$ km, i.e. $\xi\simeq 5.5h$ and the e-folding radius is thus resolved by about six surface-elevation nodes.

In [14], by letting $f=0$ and $\tau=0$, the analytical solution of (1) and (2) is obtained in a circular basin by exploiting the symmetry and good agreement is obtained between the discrete $P_1^{NC}-P_1$ results and the analytical solution since the error is about 1%. However, the analytical solution is not available here for comparison in the case $\tau\neq 0$.

The minimum and maximum values of the numerically simulated surface elevations for the $P_1^{NC}-P_1$ and GWCE schemes are shown in Tables II and III, for the explicit and implicit cases, respectively, at different stages of the propagation, up to $t=1000$ s on Mesh 1. Hence, since the phase speed of the surface gravity waves is $\sqrt{gH}\simeq 10\,\text{ms}^{-1}$ in the case $\tau=0$, the numerical experiment is stopped after the wave has propagated over approximately 10 km, much before it reaches the boundary.

The GWCE solution is found to be slightly lagging behind in the explicit case (Table II) compared with the $P_1^{NC}-P_1$ results, as far as the maximum values of the surface elevation are concerned. Further, as the ratio L/K increases the solution is slightly damped (as expected from $|R_2^{GW}|$ in Figure 2) and/or the phase lag decreases (see $\arg(R_2^{GW})$ and $\arg(R_3^{GW})$ close to the origin

DOI: 10.1002/fld

Table III. As for Table II, but for the implicit case.

		GWCE		
Time (s)	P_1^{NC}–P_1	$L=K$	$L=2K$	$L=10K$
200	0, 0.15218	0, 0.13176	0, 0.13172	0, 0.13145
400	−0.16621, 0.10578	−0.21682, 0.10326	−0.21358, 0.10283	−0.19086, 0.09980
600	−0.09281, 0.08730	−0.11177, 0.08363	−0.11029, 0.08321	−0.10068, 0.08038
800	−0.07143, 0.07483	−0.09077, 0.07040	−0.09021, 0.07004	−0.08548, 0.06762
1000	−0.06135, 0.06561	−0.07897, 0.06036	−0.07855, 0.06006	−0.07468, 0.05803

Table IV. As for Table II, but at $t=1000$ s for different values of ξ.

		GWCE		
ξ (h)	P_1^{NC}–P_1	$L=K$	$L=2K$	$L=10K$
6.5	−0.05133, 0.07751	−0.05021, 0.07897	−0.05017, 0.07887	−0.04990, 0.07817
5.5	−0.04448, 0.07172	−0.04307, 0.07390	−0.04304, 0.07378	−0.04279, 0.07291
4.5	−0.03832, 0.06544	−0.03634, 0.06813	−0.03632, 0.06798	−0.03608, 0.06689

Table V. As for Table IV, but for the implicit case.

		GWCE		
ξ (h)	P_1^{NC}–P_1	$L=K$	$L=2K$	$L=10K$
6.5	−0.06351, 0.07381	−0.07869, 0.06996	−0.07826, 0.06961	−0.07433, 0.06732
5.5	−0.06135, 0.06561	−0.07897, 0.06036	−0.07855, 0.06006	−0.07468, 0.05803
4.5	−0.06074, 0.05541	−0.07789, 0.04877	−0.07748, 0.04852	−0.07375, 0.04683

in Figure 3). Finally, whatever the ratio L/K, the minima of the GWCE surface elevation exhibit an error (a phase lead ?) compared with the P_1^{NC}–P_1 solution.

In the implicit case (Table III), the phase lead and/or damping of the GWCE solution is significant compared with the P_1^{NC}–P_1 one as far as the maxima of the surface elevation are concerned, while it is the opposite situation for the minima. Further, as the ratio L/K increases, the GWCE solution is progressively damped. For $L=10\,K$ the damping is about 4% at $t=1000$ s compared with the case $L=K$. Since we have found in Section 4.2 (Figure 5) that the variation of L with K has a minor impact on the phase, the observed damping should be essentially due to the behaviour of $|R_2^{GW}|$ close to the origin in Figure 4.

The minimum and maximum values of the numerically simulated surface elevation for both schemes are also shown in Tables IV and V, for the explicit and implicit cases, respectively, at $t=1000$ s for $\xi=6.5h$ and $\xi=4.5h$. The results are in line with those obtained for $\xi=5.5h$ and the above comments are still valid for both the explicit and implicit cases. As expected, compared with the P_1^{NC}–P_1 pair, the phase/dispersion problems observed with the GWCE scheme are slightly attenuated for $\xi=6.5h$ while they increase in the case $\xi=4.5h$.

Finally, the minimum and maximum values of the discrete surface elevation for both schemes are shown in Tables VI and VII, for the explicit and implicit cases, respectively, at $t=1000$ s and $\xi=5.5h$ on two other meshes, named Meshes 2 and 3, shown in Figure 6(b) and (c), respectively. Indeed, in [7, 13, 42, 44], it is shown that the dispersion properties of the discrete equations are dependent on the grid configuration and the direction of the wave propagation. This is the reason why Mesh 2, made up of equilateral triangles, and Mesh 3, an unstructured triangulation, are employed in the present study. The mean node spacing in the $-x$-direction is still $h=470$ m for these two meshes.

DOI: 10.1002/fld

Table VI. As for Table II, but at $t=1000$ on Meshes 1, 2 and 3.

		GWCE		
Mesh	$P_1^{NC}-P_1$	$L=K$	$L=2K$	$L=10K$
Mesh 1	−0.04448, 0.07172	−0.04307, 0.07390	−0.04304, 0.07378	−0.04279, 0.07291
Mesh 2	−0.04447, 0.07169	−0.04275, 0.07211	−0.04272, 0.07205	−0.04242, 0.07163
Mesh 3	−0.04457, 0.07208	−0.04288, 0.07254	−0.04284, 0.07247	−0.04252, 0.07202

Table VII. As for Table VI, but for the implicit case.

		GWCE		
Mesh	$P_1^{NC}-P_1$	$L=K$	$L=2K$	$L=10K$
Mesh 1	−0.06135, 0.06561	−0.07897, 0.06036	−0.07855, 0.06006	−0.07468, 0.05803
Mesh 2	−0.06135, 0.06559	−0.07865, 0.05910	−0.07826, 0.05882	−0.07472, 0.05698
Mesh 3	−0.06159, 0.06577	−0.07902, 0.05931	−0.07864, 0.05903	−0.07508, 0.05717

Again, in terms of phase/dispersion problems and dissipation, the results are in line with those obtained in Tables II–V, except that as the ratio L/K increases, the GWCE maxima are now close to the $P_1^{NC}-P_1$ ones on Meshes 2 and 3 in Table VI. However, compared with the $P_1^{NC}-P_1$ minima, the minimum values of the GWCE solutions suggest the persistence of an error (a phase lead ?) for the GWCE scheme on the three meshes in the explicit case (Table VI). It is remarkable to note that the $P_1^{NC}-P_1$ results are nearly identical on Meshes 1 and 2 in Tables VI and VII while this is not the case regarding the GWCE scheme. Indeed, the GWCE maxima represent a noticeable departure on Mesh 1 compared with those obtained on Meshes 2 and 3, whereas the GWCE minima are reasonably close on the three meshes. Finally, note that for each scheme close results are observed in Tables VI and VII on Meshes 2 and 3.

The results of Tables II–VII illustrate the phase/dispersion problems found in Section 4 for the GWCE scheme in the explicit case, and phase difference and damping effects in the implicit one, compared with the $P_1^{NC}-P_1$ pair. However, such results only give information about the minima and maxima values of the surface elevation. In order to visualize the above-mentioned problems, the simulated surface elevation is shown in Figure 7 on Mesh 1 for the $P_1^{NC}-P_1$ and GWCE (with $L=K$) schemes at $t=1000$ s, after the wave has propagated over 10 km, and for $\xi=4.5h, 5.5h$, and $6.5h$, i.e. the e-folding radius of the initial Gaussian is resolved by about 5, 6, and 7 surface-elevation nodes, respectively. Only the explicit case is considered in the following as it is of most practical interest. Further, as shown in Figure 6(d) for Mesh 1, the directions OX and OY correspond to waves travelling in the x- and y-directions, for $lh=0$ and $kh=0$, respectively, whereas the directions $OD1$ and $OD2$ correspond to waves travelling along the diagonal axes, for $kh=lh$ and $lh=-kh$, respectively.

Since $f=0$, the solution of (1) and (2) should preserve the symmetry of the initial solution in time. In Figure 7, the most interesting feature is the lack of symmetry observed for the minimum values of the surface elevation (Min) in the GWCE case, except along the $OD2$ axis. Such a problem is already noticeable for $\xi=6.5h$ but it becomes more and more significant as ξ decreases to $4.5h$. Simulations (not shown) employing other values of L and meshes have also been performed. When $L=2K$ and $10K$ the lack of symmetry is still observed on Mesh 1 (except with respect to the $OD2$ axis) but with a slightly weaker amplitude than in Figure 7. However, the GWCE results appear symmetric on Meshes 2 and 3, although the GWCE minima have approximately the same error on the three meshes compared with the $P_1^{NC}-P_1$ minima (around 5% when $\xi=4.5h$). Note that the $P_1^{NC}-P_1$ solution preserves the symmetry of the initial solution for $\xi=4.5h, 5.5h$, and $6.5h$, not only on Meshes 2 and 3 but also on Mesh 1 (Figure 7).

Because the lack of symmetry, observed on Mesh 1 in Figure 7 for the GWCE scheme (except with respect to the $OD2$ axis) should be enhanced as the wave propagates over longer distances,

DOI: 10.1002/fld

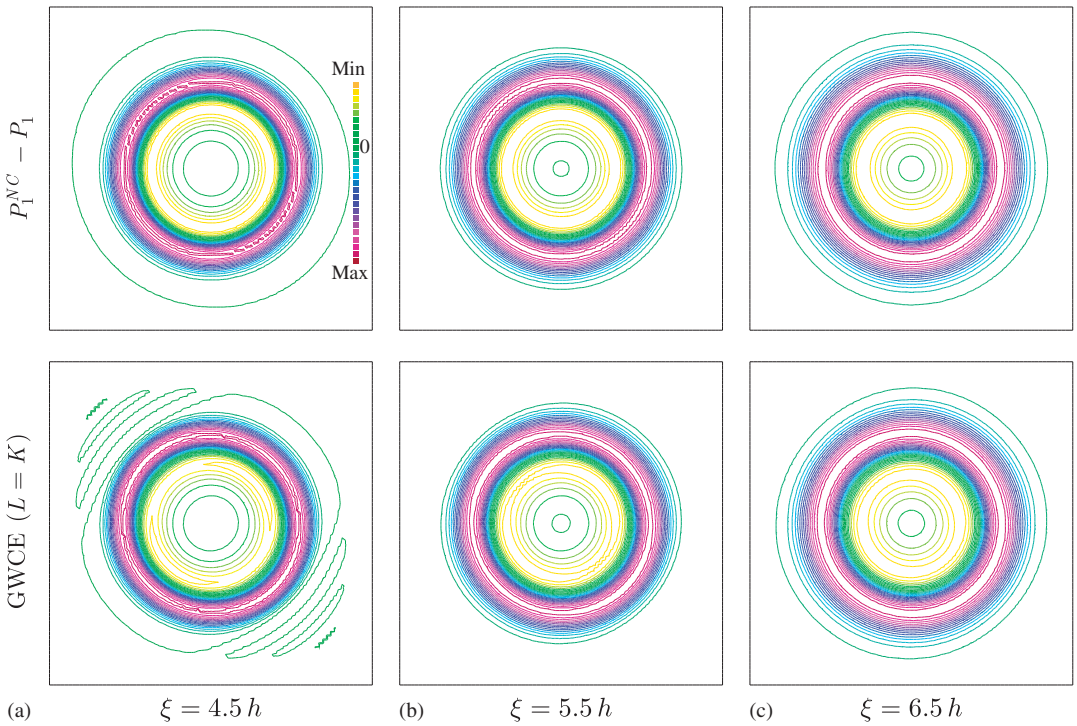

Figure 7. Surface elevation on Mesh 1, for the $P_1^{\mathrm{NC}} - P_1$ and GWCE (with $L = K$) schemes in the explicit case, at $t = 1000$ s after the wave has propagated over approximately 10 km. The maximum (Max) and minimum (Min) values and the contour interval (CI) in meters are: (a) Min $= -0.03832$ m, Max $= 0.06813$ m, CI $= 0.004258$ m; (b) Min $= -0.04448$ m, Max $= 0.07390$ m, CI $= 0.004735$ m; and (c) Min $= -0.05133$ m, Max $= 0.07897$ m, CI $= 0.005212$ m.

Table VIII. As for Table VI, but at $t = 5000$ s and for $\xi = 4.5h$ and $L = K$.

Mesh	P_1^{NC}–P_1	GWCE $(L = K)$
Mesh 1	−0.01212, 0.02500	−0.00975, 0.02577
Mesh 2	−0.01210, 0.02483	−0.00919, 0.02568
Mesh 3	−0.01200, 0.02529	−0.00896, 0.02604

the simulation of Figure 7 is carried on in the explicit case on Meshes 1, 2 and 3 for $\xi = 4.5h$ up to $t = 5000$ s, i.e. the wave propagates over approximately 50 km. The domain extent is now 120 km × 120 km and again, the numerical experiment is stopped before the wave reaches the boundary.

For both schemes, with $L = K$, the minima and maxima values of the discrete surface elevation are shown in Table VIII. The GWCE minima now exhibit a significant error ranging from 20 to 25% compared with the P_1^{NC}–P_1 results, while the difference between the GWCE and P_1^{NC}–P_1 maxima is no more than 3%, a value close to the maximum bound (4%) computed at time $t = 1000$ s for $\xi = 4.5h$ on the three meshes.

The simulated surface elevation is displayed in Figure 8 on Mesh 1 for both schemes. Again, the P_1^{NC}–P_1 solution preserves the symmetry of the initial solution, while this is only the case with respect to the $OD2$ axis for the GWCE scheme. In addition, the phase lead is now conspicuous for the GWCE minima along the $OD2$ axis (in both directions) outside the wave front, where the maximum values of the surface elevation are located (Max), an area where the solution should be roughly at rest (close to zero) as is the case for the P_1^{NC}–P_1 solution.

Similar results have been obtained for $L = 2K$ and $L = 10K$. Further, the phase lead along the $OD2$ axis and the lack of symmetry are also observed for the GWCE scheme in the case

DOI: 10.1002/fld

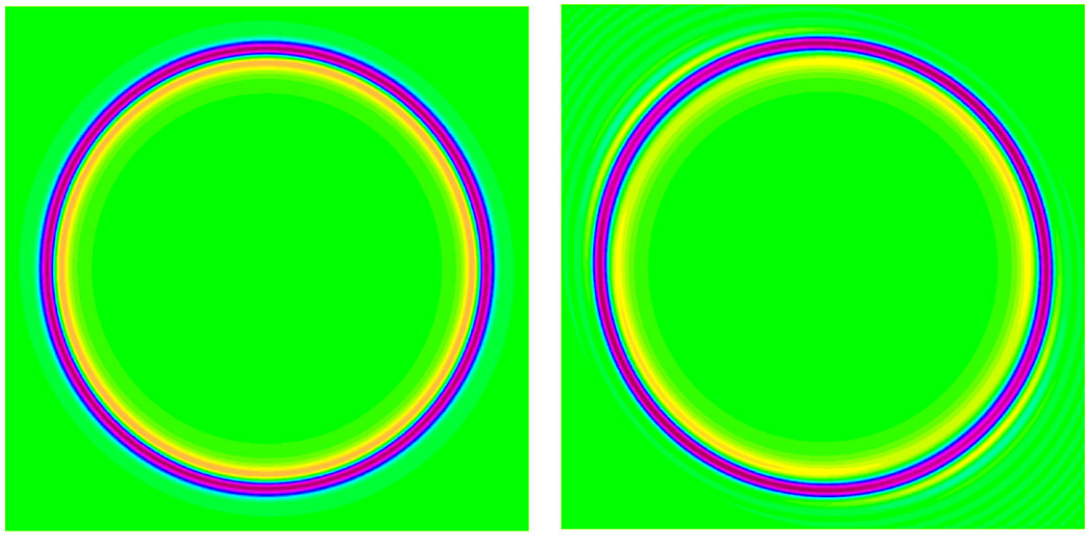

Figure 8. As for Figure 7 but at $t = 5000$ s after the wave has propagated over approximately 50 km and for $\xi = 4.5h$. We have Min $= -0.01212$ m, Max $= 0.02577$ m, CI $= 0.001516$ m.

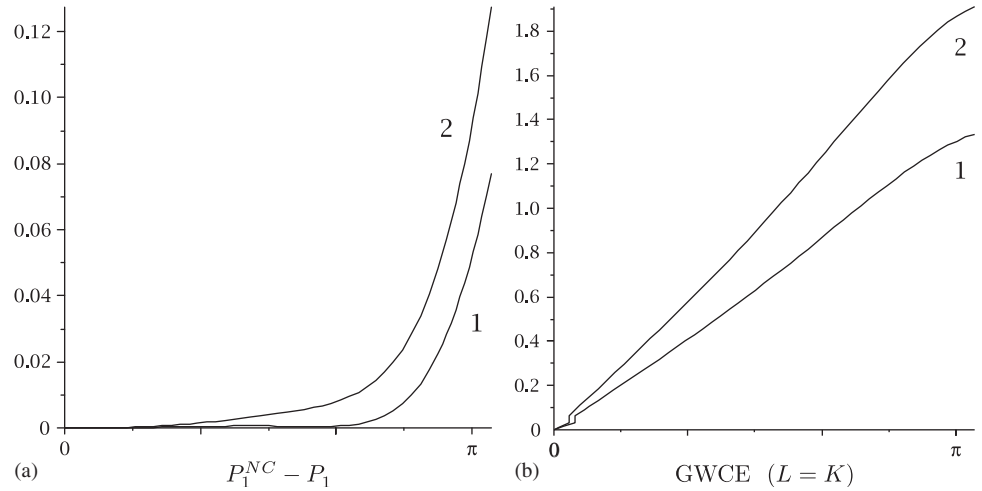

Figure 9. The argument (in radians) of R_3 in Figure 3 along selected axes: OX and OY (curve 1), OD1 and OD2 (curve 2), for the (a) $P_1^{\mathrm{NC}} - P_1$ and (b) GWCE (with $L = K$) schemes.

$\xi = 5.5h$ and $6.5h$ but with a slightly weaker amplitude. Note that the GWCE minima still exhibit a significant error of approximately 10% on Meshes 1, 2, and 3 when $\xi = 6.5h$, compared with the $P_1^{\mathrm{NC}} - P_1$ results. Additional simulations have also been performed on a mesh obtained from Mesh 1 by a rotation of $\pi/2$. The $P_1^{\mathrm{NC}} - P_1$ results are unchanged on such a mesh while the GWCE solution is shifted from $\pi/2$ compared with the solution obtained on Mesh 1, and the symmetry is now only preserved with respect to the $OD1$ axis.

The phase lead observed in the front propagation along the $OD2$ axis (on Mesh 1) for the GWCE scheme in Figure 8 is likely due to the behaviour of $\arg(R_3^{\mathrm{GW}})$ which rapidly departs from 0 for increasing values of kh and lh in Figure 3. The results of Figure 3 are hence emphasized and detailed in Figure 9(b), where $\arg(R_3^{\mathrm{GW}})$ is shown along the directions OX, OY, $OD1$, and $OD2$, and $\arg(R_3^{\mathrm{MP}})$ is given for comparison in Figure 9(a). While $\arg(R_3^{\mathrm{GW}})$ rapidly increases

DOI: 10.1002/fld

with increasing wavenumbers (decreasing wavelengths) arg(R_3^{MP}) remains close to zero for a large part of the spectrum.

On Meshes 2 and 3 the solution is found to preserve the initial symmetry for both schemes, as it was the case up to $t = 1000$ s. However, as previously mentioned, the minimum and maximum values of the surface elevation, shown in Table VIII, reveal an error of 25% for the GWCE minima compared with the P_1^{NC}–P_1 ones. Consequently, a symmetric phase lead is still observed outside the wave front, although slightly less visible than on Mesh 1, and this behaviour reflects the lack of preferential direction on Meshes 2 and 3.

6. PROPERTIES OF THE METHODS

6.1. *Efficiency*

For efficiency reasons when using a primitive equation formulation, the discrete momentum equations are generally diagonalized in some manner and substituted into the discrete continuity equations to derive a discrete form of a wave equation. This discrete wave equation is solved for sea level first. Then the velocity equations are solved by a back calculation. Forming the wave equation at the discrete level inherits all the properties of the underlying discretized primitive equation formulation. Hence, proper elements must be used that do not contain spurious surface-elevation (pressure) modes, as for the P_1^{NC}–P_1 pair. It is crucial to note that as a consequence of the orthogonality property of the linear P_1^{NC} basis functions [47], the important matrices (velocity mass and Coriolis) in (18) are 'naturally' diagonal. Such a desirable and unusual property of the FE method greatly enhances computational efficiency. This permits the *algebraic* substitution of *discrete* values of velocity from (18) into (19), and leads to a *discrete* Helmholtz problem for surface-elevation with a very sparse matrix, with an average of 13 nonzero elements per row. The substitution tremendously reduces the computational cost since a system of only N equations is solved, where N is the number of mesh vertices.

On the contrary, the GWCE formulation is formed at the mathematical model level rather than at the discrete level. Equation (21) is obtained *analytically* rather than *algebraically*. Through such a procedure, the troublesome gravity wave terms are modified so they do not contain spurious pressure modes in the discrete approximation and simple continuous linear elements can be used for all variables. The solution strategy is the same as above—solve for sea level then solve for velocity. For decoupling (20) and (21) a time-splitting procedure is usually employed where (21) is first solved for nodal elevations and then (20) is solved for the velocity field. In fact, the velocity terms in (21) are generally evaluated at time level n so that (20) and (21) are not coupled and can be solved sequentially. In (18) and (19) the decoupling is 'naturally' performed by an *algebraic* substitution and once the *discrete* Helmholtz problem is solved, the velocities are obtained from (18) without solving a linear system, since the velocity mass and Coriolis matrices are diagonal.

As far as data structure is concerned, the wave equation with linear elements and the P_1^{NC}–P_1 element are similar because both use a P_1 approximation for surface-elevation with the variable evaluated at the vertices in the grid. The number of velocity nodes is three times greater for the latter approximation because they are evaluated at the midside nodes rather than at the vertices. However, the velocities are obtained from (18) without solving a linear system for the P_1^{NC}–P_1 pair (because the velocity mass and Coriolis matrices are diagonal) contrary to the GWCE case. As a result, the efficiencies are comparable.

6.2. *Mass conservation*

The primitive equation and GWCE FE approximations have very different mass conservation properties. Formally, the wave equation formulation will conserve mass globally if the initial conditions conserve mass. Otherwise, the initial errors eventually decay over a period of time as shown in [28]. The effect of the G term is to correct the mass continuity at each time step and G can be considered an inverse relaxation time. For highly refined grids, the mass conservation can

DOI: 10.1002/fld

be very accurate; however, field scale problems generally involve rough data and this degree of accuracy may not be attainable.

Mass conservation properties for the GWCE have been examined by Kolar *et al*. [34], which arrived at two major conclusions: the discretization of the nonlinear terms in the momentum equation must be the same as the discretization in the GWCE, and the parameter G must be adjusted to minimize continuity errors while not permitting spurious oscillations. An inconsistent discretization can lead to spurious momentum sources and sinks that manifest themselves as continuity errors. In addition, a recommended value for G is $\tau_{\max} \leqslant G \leqslant 10\,\tau_{\max}$, where $\tau_{\max}$ is the maximum value of the bottom friction coefficient. These measures can reduce continuity errors but do not eliminate them. As a result, global mass conservation can be attained with suitable boundary conditions, but local mass balance is not satisfied even in a weighted residual sense.

On the other hand, the mixed primitive variable formulation with the P_1^{NC}–P_1 element conserves mass in the traditional FE sense both globally and in a weighted residual sense about the element vertices [50]. This approach does not satisfy continuity on an elemental basis such as a finite volume formulation unless the velocity field is corrected in some manner.

These mass conservation properties have important consequences for consistency, and hence the types of scalar transport models they can be coupled with. Because the discrete continuity equation and the scalar transport equation must be consistent, the P_1^{NC}–P_1 pair can be coupled to traditional Galerkin formulations that solve for scalar transport in a weighted residual sense (continuous Galerkin). These types of methods are more difficult to formulate when advection is significant. It is not clear which scalar transport schemes, if any, are consistent with the wave equation formulation. Generally, traditional continuous Galerkin methods are used. For smooth data and highly refined grids (the academic problem), inconsistency in the approximations is not apparent in the results. However, in practical applications with rough data, the approximation errors become obvious and the solution may not converge.

6.3. Advection

Approximation of the advection term is one of the long-standing problems with the wave equation formulation. The formulation is accurate for waves, but has serious stability problems when advection becomes significant (more than a few percent of the force balance). For instance, simulations of the effects of tropical cyclones in the Bay of Bengal required a time-step 10–100 times smaller than would be predicted from stability considerations [51]. In addition, small-scale filters for velocity were necessary to stabilize the solution. This particular approach was not stable for any river simulations that were attempted.

Some progress has been made in [34, 52] in resolving these issues. Using a form of the advection that is consistent with the conservative momentum equations improves both stability and mass conservation [34]. In addition, Dresback [52] employed a predictor method to help stabilize the advection. However, both these models use an implicit time-integration scheme which has been shown to be highly dissipative [27] and hence probably helps control the small-scale noise generated by the advective terms. Moreover, even with these improvements the Courant number still must remain small with respect to general stability constraints on the advective term.

On the other hand, the primitive equation formulations are not limited by the advection approximation that can be used. Depending on the specific problem and goals, both explicit methods with Courant number stability constraints [53] or semi-Lagrangian methods without these constraints [54] can be utilized. All things considered, the P_1^{NC}–P_1 element using the primitive equations would appear to be a replacement candidate for the wave equation with linear elements and would provide a means to mitigate some of the problems with the latter.

7. CONCLUDING REMARKS

Initial applications of the FE method to surface water flows were plagued by spurious pressure oscillations caused by the coupling in the gravity wave terms. The wave equation formulation

DOI: 10.1002/fld

provided a means to modify the primitive equation formulation and remove these modes. However, the wave equation approach has been shown to have problems with conservation properties [34], problems with advection instabilities [52], poor accuracy using implicit methods [43], and lack of consistency with scalar transport schemes. As shown in the references, these problems can be partially mitigated by the proper choice of the parameter G and consistent treatment of the advection term.

However, these problems have largely been resolved with many new ocean models using the primitive equation approach. We present one such FE approach using the P_1^{NC}–P_1 element pair and show how it solves the problems associated with the wave equation approach while not incurring any new problems. In particular, the element pair has orthogonal P_1^{NC} bases that result in a diagonal mass matrix for velocity such that the method retains the high efficiency of the GWCE. In addition, specification of advection schemes is flexible from Eulerian–Lagrangian methods to explicit momentum conserving approaches. Finally, the approach conserves mass in the traditional FE sense both globally and in a weighted residual sense. Hence, the equations are consistent with continuous Galerkin scalar transport methods.

Dispersion analysis presented here has shown that the P_1^{NC}–P_1 FE approach has comparable or better amplitude and phase accuracy as the GWCE. Relative accuracy is much better when using implicit methods in particular. The numerical simulations illustrate these results.

ACKNOWLEDGEMENTS

This work is supported by grants to D. Y. L. from the Natural Sciences and Engineering Research Council (NSERC) and ArcticNet.

REFERENCES

1. Cushman-Roisin B. *Introduction to Geophysical Fluid Dynamics*. Prentice-Hall: Englewood Cliffs, NJ, 1994.
2. Pedlosky J. *Geophysical Fluid Dynamics*. Springer: Berlin, 1987.
3. Walters RA, Carey GF. Analysis of spurious oscillation modes for the shallow water and Navier–Stokes equations. *Computers and Fluids* 1983; **11**:51–68.
4. Arakawa A, Lamb VR. Computational design of the basic dynamical processes of the UCLA general circulation model. *Methods in Computational Physics* 1977; **17**:173–265.
5. Brezzi F, Fortin M. *Mixed and Hybrid Finite Element Methods*. Springer Series in Computational Mathematics, vol. 15. Springer: Berlin, 1991.
6. Iskandarani M, Haidvogel D, Boyd J. A staggered spectral finite element model for the shallow-water equations. *International Journal for Numerical Methods in Fluids* 1995; **20**:393–414.
7. Le Roux DY, Rostand V, Pouliot B. Analysis of numerically induced oscillations in 2D finite-element shallow-water models. Part I: inertia-gravity waves. *SIAM Journal on Scientific Computing* 2007; **29**:331–360.
8. Batteen ML, Han YJ. On the computational noise of finite-difference schemes used in ocean models. *Tellus* 1981; **33**:387–396.
9. Adcroft AJ, Hill CN, Marshall JC. A new treatment of the Coriolis terms in C-grid models at both high and low resolutions. *Monthly Weather Review* 1999; **127**:1928–1936.
10. Walters RA, Carey GF. Numerical noise in ocean and estuarine models. *Advances in Water Resources* 1984; **7**:15–20.
11. Rostand V, Le Roux DY, Carey G. Kernel analysis of the discretized finite difference and finite element shallow water models. *SIAM Journal on Scientific Computing* 2008; **31**:531–556.
12. Kleptsova O, Pietrzak J, Stelling GS. On the accurate and stable reconstruction of tangential velocities in C-grid ocean models. *Ocean Modelling* 2009; **28**:118–126.
13. Rostand V, Le Roux DY. Raviart–Thomas and Brezzi–Douglas–Marini finite element approximations of the shallow-water equations. *International Journal for Numerical Methods in Fluids* 2008; **57**:951–976.
14. Le Roux DY. Dispersion relation analysis of the P_1^{NC}–P_1 finite-element pair in shallow-water models. *SIAM Journal on Scientific Computing* 2005; **27**:394–414.
15. Le Roux DY, Carey GF. Stability/dispersion analysis of the discontinuous Galerkin linearized shallow-water system. *International Journal for Numerical Methods in Fluids* 2005; **48**:325–347.
16. Randall DA. Geostrophic adjustment and the finite-difference shallow-water equations. *Monthly Weather Review* 1994; **122**:1371–1377.
17. Williams RT, Zienkiewicz OC. Improved finite-element forms for the shallow-water wave equations. *International Journal for Numerical Methods in Fluids* 1981; **1**:81–97.
18. King IP, Norton WR, Iceman KR. A finite element model for two-dimensional flow. In *Finite Elements in Flow Problems*, Oden JT (ed.). University of Alabama at Huntsville (UAH) Press, 1974; 133–137.

DOI: 10.1002/fld

19. Raviart PA, Thomas JM. A mixed finite element method for 2nd order elliptic problems. *Mathematical Aspects of the Finite Element Method*. Lecture Notes in Mathematics. Springer: Berlin, 1977.
20. Sigurdsson S. Treatment of the convective term in staggered finite element schemes for shallow water flow. In *Computational Methods in Water Resources IX, Volume 1*: *Numerical Methods in Water Resources*, Russel TF, Ewing RE, Brebia CA, Gray WG, Pinder GF (eds). Computational Mechanics Publications, Elsevier Applied Science, 1992; 291–298.
21. Williams RT. On the formulation of the finite-element prediction models. *Monthly Weather Review* 1981; **109**:463–466.
22. Hua BL, Thomasset F. A noise-free finite element scheme for the two-layer shallow water equations. *Tellus* 1984; **36A**:157–165.
23. Le Roux DY. A new triangular finite-element with optimum constraint ratio for compressible fluids. *SIAM Journal on Scientific Computing* 2001; **23**:66–80.
24. Hughes TJR, Franca LP, Balestra M. A new finite element formulation for computational fluid dynamics: V. Circumventing the Babuska–Brezzi condition: a stable Petrov–Galerkin formulation of the Stokes problem accomodating equal-order interpolations. *Computer Methods in Applied Mechanics and Engineering* 1986; **59**:85–99.
25. Cullen MJP, Hall CD. Forecasting and general circulation results from finite-element models. *Quarterly Journal of the Royal Meteorological Society* 1979; **105**:571–592.
26. Staniforth A, Mitchell HL. A semi-implicit finite-element barotropic model. *Monthly Weather Review* 1977; **105**:154–169.
27. Lynch DR, Gray WG. A wave equation model for finite element tidal computations. *Computers and Fluids* 1979; **7**:207–228.
28. Kinnmark IPE, Gray WG. A two-dimensional analysis of the wave equation model for finite element tidal computations. *International Journal for Numerical Methods in Engineering* 1984; **20**:369–383.
29. Luettich RA, Westerink JJ, Scheffner NW. An advanced three-dimensional circulation model of shelves, coasts, and estuaries. *Report 1*: *Theory and Methodology of ADCIRC-2DDI and ADCIRC-3DL*, *Technical Report DRP-92-6*, Department of the Army, Vicksburg, MS, 1992.
30. Dawson C, Westerink JJ, Feyen JC, Pothina D. Continuous, discontinuous and coupled discontinuous continuous Galerkin finite element methods for the shallow water equations. *International Journal for Numerical Methods in Fluids* 2006; **52**:63–88.
31. Westerink JJ, Luettich RA, Feyen JC, Atkinson JH, Dawson C, Roberts HJ, Powell MD, Dunion JP, Kubatko EJ, Pourtaheri H. A basin- to channel-scale unstructured grid hurricane storm surge model applied to Southern Louisiana. *Monthly Weather Review* 2008; **136**:833–864.
32. Le Bars Y, Lyard F, Jeandel C, Dardengo L. The AMANDES tidal model for the Amazon estuary and shelf. *Ocean Modelling* 2010; **31**:132–149.
33. Kinnmark IPE. In *The Shallow-water Wave Equations*: *Formulation*, *Analysis and Application*, Brebbia CA, Orszag SA (eds). Lecture Notes in Engineering, vol. 15. Springer: Berlin, 1986; 1–187.
34. Kolar RL, Westerink JJ, Cantekin ME, Blain CA. Aspects of nonlinear simulations using shallow-water models based on the wave continuity equation. *Computers and Fluids* 1994; **23**:523–538.
35. Kolar RL, Westerink JJ. A look back at 20 years of GWC-based shallow-water models. In *Proceedings of the XIII International Conference on Computational Methods in Water Resources*, Calgary, Canada, Bentley LR *et al.* (eds), vol. 2. 2000; 899–906.
36. Comblen R, Lambrechts J, Remacle JF, Legat V. Practical evaluation of five partly discontinuous finite element pairs for the non-conservative shallow water equations. *International Journal for Numerical Methods in Fluids* 2010; **63**:701–724.
37. Le Roux DY, Pouliot B. Analysis of numerically induced oscillations in two-dimensional finite-element shallow-water models. Part II: free planetary waves. *SIAM Journal on Scientific Computing* 2008; **30**:1971–1991.
38. Hanert E, Walters RA, Le Roux DY, Pietrzak J. A tale of two elements: P_1^{NC}–P_1 and RT_0. *Ocean Modelling* 2009; **28**:24–33.
39. Le Roux DY, Hanert E, Rostand V, Pouliot B. Impact of mass lumping on gravity and Rossby waves in 2D finite-element shallow-water models. *International Journal for Numerical Methods in Fluids* 2009; **59**:767–790.
40. Foreman MGG. An analysis of the wave equation model for finite element tidal computations. *Journal of Computational Physics* 1983; **52**:290–312.
41. Kinnmark IPE, Gray WG. Stability and accuracy of spatial approximations for wave equation tidal models. *Journal of Computational Physics* 1985; **60**:447–466.
42. Foreman MGG. A two-dimensional dispersion analysis of selected methods for solving the linearized shallow water equations. *Journal of Computational Physics* 1984; **56**:287–323.
43. Gray WG, Lynch DR. Time-stepping schemes for finite element tidal model computations. *Advances in Water Resources* 1977; **1**:83–95.
44. Atkinson JH, Westerink JJ, Luettich RA. Two-dimensional dispersion analyses of finite element approximations to the shallow water equations. *International Journal for Numerical Methods in Fluids* 2004; **45**:715–749.
45. LeBlond PH, Mysak LA. *Waves in the Ocean*. Elsevier: Amsterdam, 1978; 602.
46. Walters RA, Barragy EJ. Comparison of *h* and *p* finite element approximations of the shallow water equations. *International Journal for Numerical Methods in Fluids* 1997; **24**:61–79.

DOI: 10.1002/fld

47. Thomasset F. *Implementation of Finite Element Methods for Navier–Stokes Equations*. Springer: Berlin, 1981.
48. Porter B. *Stability Criteria for Linear Dynamical Systems*. Academic Press: New York, 1968.
49. Leendertse JJ. Aspects of a computational model for long period water-wave propagation. *Technical Report RM-5294-PR*, Sancta Monica, Rand Memorandum, 1967.
50. Hanert E, Le Roux DY, Legat V, Deleersnijder E. Advection schemes for unstructured grid ocean modelling. *Ocean Modelling* 2004; **7**:39–58.
51. Henry RF, Duncalf DS, Walters RS, Osborne MJ, Murty TS. A study of tides and storm surges in offshore waters of the Meghna estuary using a finite element model. *Mausam* 1997; **48**:519–530.
52. Dresback KM, Kolar RL, Dietrich JC. A 2D implicit time-marching algorithm for shallow water models based on the generalized wave continuity equation. *International Journal for Numerical Methods in Fluids* 2004; **45**:253–274.
53. Hanert E, Le Roux DY, Legat V, Deleersnijder E. An efficient Eulerian finite-element method for the shallow-water equations. *Ocean Modelling* 2005; **10**:115–136.
54. Walters RA, Lane EM, Henry RF. Semi-Lagrangian methods for a finite element coastal ocean model. *Ocean Modelling* 2007; **19**:112–124.

DOI: 10.1002/fld