INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN FLUIDS
Int. J. Numer. Meth. Fluids 2008; **57**:1249–1268
Published online 7 January 2008 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/fld.1718

Finite element approximation of the modified Boussinesq equations using a stabilized formulation

Ramon Codina[1,*,†], José M. González-Ondina[2], Gabriel Díaz-Hernández[2] and Javier Principe[1]

[1]*CIMNE, Universitat Politècnica de Catalunya, Jordi Girona 1-3, Edifici C1, 08034 Barcelona, Spain*
[2]*Instituto de Hidráulica Ambiental IH Cantabria, Universidad de Cantabria, Avda. de los Castros s/n, 39005 Santander, Spain*

SUMMARY

In this work, we present a finite element model to approximate the modified Boussinesq equations. The objective is to deal with the major problem associated with this system of equations, namely, the need to use stable velocity-depth interpolations, which can be overcome by the use of a stabilization technique. The one described in this paper is based on the splitting of the unknowns into their finite element component and the remainder, which we call the subgrid scale. We also discuss the treatment of high-order derivatives of the mathematical model and describe the time integration scheme. Copyright © 2008 John Wiley & Sons, Ltd.

Received 14 September 2007; Revised 14 November 2007; Accepted 18 November 2007

KEY WORDS: non-linear waves; Boussinesq equations; mixed interpolations; stabilized finite elements

1. INTRODUCTION

There are several mathematical models for flows in shallow domains. However, a feature they have in common is the mathematical structure of the coupling between the water elevation and the velocity, the unknowns of the problem. This coupling is already present in the simplest setting, modeling linear gravity waves in shallow domains, and is also present in more complex models, such as the Saint-Venant or the Boussinesq equations.

In this work, we present a finite element approximation of the modified Boussinesq equations introduced in [1]. We treat different aspects related to this problem, such as the way to deal with third-order derivatives, the linearization or the time integration. However, the main topic is the development of a formulation allowing to use *equal* interpolation for the water elevation and the

*Correspondence to: Ramon Codina, CIMNE, Universitat Politècnica de Catalunya, Jordi Girona 1-3, Edifici C1, 08034 Barcelona, Spain.
†E-mail: ramon.codina@upc.edu

velocity. In general, this is not possible, not even for the linear problem using the classical Galerkin method.

Our formulation is based on the variational multiscale approach in the format introduced in [2, 3]. The basic idea is to split the unknowns into a *resolvable* component, which can be reproduced by the discretization method (in our case finite elements) and the remainder, which we will call *subgrid scale* or *subscale*. Rather than solving exactly for the latter, the formulation results from a closed-form approximation for the subscales, which is designed in order to capture their *effect* on the discrete finite element solution. This leads to a formulation that allows the use of equal velocity-depth interpolations. In this sense, this work is an extension of [4].

Several attempts to approximate the modified Boussinesq equations can be found in the literature. An early finite difference approximation can be found in [5] and another popular finite difference model in [6]. Finite element approximations were introduced later, see, for example, [7–10]. In these references, high-frequency oscillations over the grid used to discretize the domain are mentioned, and methods to overcome them by *ad hoc* filtering techniques or by the addition of numerical viscosity are reported, for example, in [9] and references therein (see also [10]). In the finite difference context, different grids can be used for the approximation of velocity and water elevation (see [7], for example). Surprisingly, there seems to be no explicit association between the instability problems encountered and the lack of stability of the Galerkin method and, consequently, the problem has not been tackled using stabilized finite element methods. This is precisely the approach advocated in this work. Of course, there are several works dealing with stabilization for shallow water equations, but usually intended either to stabilize low viscosity flows or important source terms. For example, the Taylor–Galerkin method, originally presented in [11], was used for solving the shallow water equations in [12], whereas characteristic-based schemes were proposed in [13, 14]. A Galerkin/least-square formulation was presented in [15].

This paper is organized as follows. In Section 2 we state the initial and boundary value problem to be solved, both in its differential and in its weak form. The space discretization is presented in Section 3. The main contribution of this work is presented in Section 4, where a stabilized finite element method is proposed. After presenting the basis of the formulation, its application to the linearized non-dispersive model is studied in detail. The algorithmic parameters on which the formulation depends are designed on the basis of a Fourier analysis of the problem, similar to that proposed in [4, 16]. The formulation is then extended to the extended Boussinesq model, after some considerations on the application of this type of stabilization techniques to non-linear problems. In Section 5 we propose a finite difference scheme to integrate the equations in time based on a predictor–corrector algorithm. We then present the numerical results of two representative examples, merely intended to demonstrate the misbehavior of the Galerkin method and the improvement provided by the proposed formulation. Some concluding remarks close this work.

2. PROBLEM STATEMENT

2.1. Initial and boundary value problem

Let us consider the motion of a fluid in a shallow domain whose horizontal projection is $\Omega \subset \mathbb{R}^2$ and whose depth, measured when the fluid is at rest from a horizontal free surface to the bottom of the domain, is $H(\mathbf{x}), \mathbf{x} = (x_1, x_2) \in \Omega$. The vertical coordinate is taken $x_3 = 0$ at the free surface at rest, so that $x_3 = -H(\mathbf{x})$ is the equation for the bathymetry. Let $\eta(\mathbf{x}, t)$ be the free surface elevation

DOI: 10.1002/fld

of the fluid in motion and $\mathbf{u}(\mathbf{x},t)$ the velocity measured at $x_3=\beta H$, with the parameter β given, and with $t\in[0,T]$, the time interval of analysis.

Let a be the amplitude and λ the wavelength of a characteristic mode of a wave propagating in the domain of analysis. Let also H_0 be a characteristic depth of this domain, and define the dimensionless numbers

$$\varepsilon:=\frac{a}{H_0},\quad \mu:=\frac{H_0}{\lambda}$$

The Boussinesq wave theory is obtained by expanding the equations of motion for an inviscid incompressible fluid in terms of ε and μ, and retaining only the terms of order up to $\mathcal{O}(\varepsilon)$ and $\mathcal{O}(\mu^2)$, so that it requires $\varepsilon\ll 1$, $\mu\ll 1$ and $\varepsilon/\mu^2=\mathcal{O}(1)$.

The modified Boussinesq equations presented in [1] can be expressed as

$$\partial_t\eta+\nabla\cdot(H\mathbf{u})+\varepsilon\nabla\cdot(\eta\mathbf{u})+\mu^2\nabla\cdot\mathbf{J}_\eta=0 \tag{1}$$

$$\partial_t\mathbf{u}+g\nabla\eta+\varepsilon\mathbf{u}\cdot\nabla\mathbf{u}+\mu^2\mathbf{J}_u=\mathbf{0} \tag{2}$$

where g is the magnitude of the gravity acceleration and we have introduced the auxiliary fields

$$\mathbf{J}_\eta:=C_1H^3\mathbf{E}+C_3H^2\mathbf{E}^H \tag{3}$$

$$\mathbf{J}_u:=C_2H^2\partial_t\mathbf{E}+\beta H\partial_t\mathbf{E}^H \tag{4}$$

$$\mathbf{E}:=\nabla D,\quad D:=\nabla\cdot\mathbf{u} \tag{5}$$

$$\mathbf{E}^H:=\nabla D^H,\quad D^H:=\nabla\cdot(H\mathbf{u}) \tag{6}$$

where C_i, $i=1,2,3$, are constants defined in terms of β by

$$C_1=\frac{1}{2}\left(\beta^2-\frac{1}{3}\right),\quad C_2=\frac{\beta^2}{2},\quad C_3=\beta+\frac{1}{2} \tag{7}$$

The value for the parameter β suggested in [1] is $\beta=-0.531$.

The boundary conditions to be considered are of three types:

- *Inflow boundary*, Γ_I: The elevation is known, so that

  $$\eta=\bar{\eta}\quad\text{on }\Gamma_\mathrm{I}$$

  where the overbar denotes given boundary conditions. The velocity $\bar{\mathbf{u}}$ depends on the elevation $\bar{\eta}$, and can be computed from the linear theory, if wished. Imposing this velocity is required if one wants to consider ε arbitrary.
- *Reflecting boundary*, Γ_R: In this case, the normal component of the velocity must be zero. It can be shown that this implies that the normal component of $\mathbf{J}_\eta$ must vanish [6], so that

  $$\mathbf{n}\cdot\mathbf{u}=0\quad\text{and}\quad\mathbf{n}\cdot\mathbf{J}_\eta=0\quad\text{on }\Gamma_\mathrm{R}$$

- *Absorbing boundary*, $\Gamma_\mathrm{A}=\partial\Omega\setminus\Gamma_\mathrm{I}\setminus\Gamma_\mathrm{R}$: Following Wei and Kirby [6], where the ideas of Israeli and Orzag [17] are implemented, a possible way to deal with absorbing boundaries is to add a diffusion term to both Equations (1) and (2) close to the boundary where the waves

DOI: 10.1002/fld

need to be absorbed. The diffusion coefficient is considered to be of exponential type, varying from zero to a given value in a layer next to the absorbing boundary and with numerically tuned constants. The boundary condition is effectively applied when the fictitious diffusion term added is integrated by parts and the boundary term is dropped. Another possibility would be to use Berenger's perfectly matched layer [18].

Finally, initial conditions of the form $\eta(\mathbf{x},0)=\eta^0(\mathbf{x})$ and $\mathbf{u}(x,0)=\mathbf{u}^0(\mathbf{x})$ have to be appended to the problem defined by (1)–(6) and the boundary conditions just described.

2.2. *Variational problem*

Let us obtain now the weak or variational form of problem (1)–(2) with the boundary conditions just described. Let $\xi(\mathbf{x})$ and $\mathbf{v}(\mathbf{x})$ be the elevation and velocity test functions, respectively, belonging to the appropriate functional spaces. To account for the boundary conditions described, ξ must vanish on Γ_{I} and the normal component of $\mathbf{v}$ must vanish on Γ_{R}.

Multiplying (1) by ξ and (2) by $\mathbf{v}$ and integrating by parts, one obtains

$$\int_\Omega \xi\partial_t\eta\,\mathrm{d}\mathbf{x}-\int_\Omega \nabla\xi\cdot(H\mathbf{u})\,\mathrm{d}\mathbf{x}-\varepsilon\int_\Omega \nabla\xi\cdot(\eta\mathbf{u})\,\mathrm{d}\mathbf{x}-\mu^2\int_\Omega \nabla\xi\cdot\mathbf{J}_\eta\,\mathrm{d}\mathbf{x}=0 \tag{8}$$

$$\int_\Omega \mathbf{v}\cdot\partial_t\mathbf{u}\,\mathrm{d}\mathbf{x}+g\int_\Omega \mathbf{v}\cdot\nabla\eta\,\mathrm{d}\mathbf{x}+\varepsilon\int_\Omega \mathbf{v}\cdot(\mathbf{u}\cdot\nabla\mathbf{u})\,\mathrm{d}\mathbf{x}+\mu^2\int_\Omega \mathbf{v}\cdot\mathbf{J}_u\,\mathrm{d}\mathbf{x}=0 \tag{9}$$

which must hold for all test functions ξ and $\mathbf{v}$. The boundary terms that arise from integration by parts in (8) vanish, that is,

$$\int_{\partial\Omega}\xi\mathbf{n}\cdot\mathbf{u}(H+\varepsilon\eta)\,\mathrm{d}\mathbf{x}+\mu^2\int_{\partial\Omega}\xi\mathbf{n}\cdot\mathbf{J}_\eta\,\mathrm{d}\mathbf{x}=0$$

since $\xi=0$ on Γ_{I} and $\mathbf{n}\cdot\mathbf{u}=0$, $\mathbf{n}\cdot\mathbf{J}_\eta=0$ on Γ_{R}.

The auxiliary fields $\mathbf{J}_\eta$ and $\mathbf{J}_u$ defined in (3) and (4), respectively, involve second derivatives of the velocity. To cope with them, there are basically two options, either to project directly $\mathbf{E}$ and $\mathbf{E}^H$ in (5) and (6) or to project first D and D^H. Let us discuss both possibilities.

Projection of $\mathbf{E}$ *and* $\mathbf{E}^H$: The idea in this case is to consider (5) and (6) as additional equations to complete the problem in order to reduce the need for regularity of the finite element approximation. This approach is followed, for example, in [9]. Since $\mathbf{n}\cdot\mathbf{J}_\eta$ must vanish on Γ_{R} for all values of β, we must have $\mathbf{n}\cdot\mathbf{E}=\mathbf{n}\cdot\mathbf{E}^H=0$ on this boundary. Let $\mathbf{F}$ and $\mathbf{F}^H$ be the appropriate test functions for $\mathbf{E}$ and $\mathbf{E}^H$, respectively. The equations for fields $\mathbf{E}$ and $\mathbf{E}^H$ are

$$\int_\Omega \mathbf{E}\cdot\mathbf{F}\,\mathrm{d}\mathbf{x}=-\int_\Omega(\nabla\cdot\mathbf{F})(\nabla\cdot\mathbf{u})\,\mathrm{d}\mathbf{x}+\int_{\Gamma_{\mathrm{I}}}(\mathbf{n}\cdot\mathbf{F})(\nabla\cdot\mathbf{u})\,\mathrm{d}\mathbf{x} \tag{10}$$

$$\int_\Omega \mathbf{E}^H\cdot\mathbf{F}^H\,\mathrm{d}\mathbf{x}=-\int_\Omega(\nabla\cdot\mathbf{F}^H)(\nabla\cdot(H\mathbf{u}))\,\mathrm{d}\mathbf{x}+\int_{\Gamma_{\mathrm{I}}}(\mathbf{n}\cdot\mathbf{F}^H)(\nabla\cdot(H\mathbf{u}))\,\mathrm{d}\mathbf{x} \tag{11}$$

The boundary integral over Γ_{R} is zero, since the test functions $\mathbf{F}$ and $\mathbf{F}^H$ must vanish there. This is why only the integral over Γ_{I} appears in the previous expressions.

DOI: 10.1002/fld

This approach requires more regularity on the velocity than the original equations. In particular, $\nabla\cdot\mathbf{u}$ restricted to the boundary needs to make sense.

Projection of D and D^H: A second possibility is to project D and D^H in (5) and (6), respectively. If G and G^H are the corresponding test functions, the variational equations to be considered are

$$\int_\Omega \mathbf{E}\cdot\mathbf{F}\,\mathrm{d}\mathbf{x}=\int_\Omega \mathbf{F}\cdot\nabla D\,\mathrm{d}\mathbf{x},\quad \int_\Omega GD\,\mathrm{d}\mathbf{x}=\int_\Omega G\nabla\cdot\mathbf{u}\,\mathrm{d}\mathbf{x}$$

$$\int_\Omega \mathbf{E}^H\cdot\mathbf{F}^H\,\mathrm{d}\mathbf{x}=\int_\Omega \mathbf{F}^H\cdot\nabla D^H\,\mathrm{d}\mathbf{x},\quad \int_\Omega G^H D^H\,\mathrm{d}\mathbf{x}=\int_\Omega G^H\nabla\cdot(H\mathbf{u})\,\mathrm{d}\mathbf{x}$$

In this case, no boundary conditions are *explicitly* required, neither for the unknowns $\mathbf{E}$, $\mathbf{E}^H$, D and D^H nor for the corresponding test functions $\mathbf{F}$, $\mathbf{F}^H$, G and G^H.

Remark 1

The last term in (9) could be integrated by parts as follows:

$$\int_\Omega \mathbf{v}\cdot\mathbf{J}_u\,\mathrm{d}\mathbf{x}=-C_2\int_\Omega\nabla\cdot(H^2\mathbf{v})\nabla\cdot(\partial_t\mathbf{u})\,\mathrm{d}\mathbf{x}-\beta\int_\Omega\nabla\cdot(H\mathbf{v})\nabla\cdot(H\partial_t\mathbf{u})\,\mathrm{d}\mathbf{x}$$
$$+C_2\int_{\partial\Omega}\mathbf{n}\cdot(H^2\mathbf{v})\nabla\cdot(\partial_t\mathbf{u})\,\mathrm{d}\mathbf{x}+\beta\int_{\partial\Omega}\mathbf{n}\cdot(H\mathbf{v})\nabla\cdot(H\partial_t\mathbf{u})\,\mathrm{d}\mathbf{x}$$

The boundary integrals vanish on Γ_R, but not on Γ_I and therefore they would have to be evaluated over this boundary. This approach would avoid the need of using any projection to deal with $\mathbf{J}_u$, although it would be still needed for $\mathbf{J}_\eta$.

3. SPACE DISCRETIZATION USING THE GALERKIN METHOD

3.1. Discrete variational equations

Let $\{\Omega^e\}$ be a finite element partition of the domain Ω, with $e=1,\ldots,n_\mathrm{el}$, of size $h=\max_e h^e$, $h^e=\mathrm{diam}(\Omega^e)$. Let also V_h be a finite element space constructed from this partition using *continuous* Lagrangian interpolation within each element domain. Clearly, this space is a subspace of the space where the continuous unknowns (elevation and velocity components) must be defined. We intend to use equal interpolation for both, and therefore the problem consists in seeking $\eta_h(\cdot,t)\in V_h$ and $\mathbf{u}_h(\cdot,t)\in V_h^2$ satisfying the adequate boundary conditions and solution of the finite-dimensional time evolution problem

$$\int_\Omega \xi_h\partial_t\eta_h\,\mathrm{d}\mathbf{x}-\int_\Omega\nabla\xi_h\cdot(H\mathbf{u}_h)\,\mathrm{d}\mathbf{x}-\varepsilon\int_\Omega\nabla\xi_h\cdot(\eta_h\mathbf{u}_h)\,\mathrm{d}\mathbf{x}-\mu^2\int_\Omega\nabla\xi_h\cdot\mathbf{J}_{\eta,h}\,\mathrm{d}\mathbf{x}=0 \tag{12}$$

$$\int_\Omega \mathbf{v}_h\cdot\partial_t\mathbf{u}_h\,\mathrm{d}\mathbf{x}+g\int_\Omega\mathbf{v}_h\cdot\nabla\eta_h\,\mathrm{d}\mathbf{x}+\varepsilon\int_\Omega\mathbf{v}_h\cdot(\mathbf{u}_h\cdot\nabla\mathbf{u}_h)\,\mathrm{d}\mathbf{x}+\mu^2\int_\Omega\mathbf{v}_h\cdot\mathbf{J}_{u,h}\,\mathrm{d}\mathbf{x}=0 \tag{13}$$

which must hold for all test functions $\xi_h\in V_h$ and $\mathbf{v}_h\in V_h^2$ satisfying the corresponding homogeneous boundary conditions. Initial conditions have to be appended to this initial value problem.

DOI: 10.1002/fld

The vector fields $\mathbf{J}_{\eta,h}$ and $\mathbf{J}_{u,h}$ are given by

$$\mathbf{J}_{\eta,h}=C_1H^3\mathbf{E}_h+C_3H^2\mathbf{E}_h^H \tag{14}$$

$$\mathbf{J}_{u,h}=C_2H^2\partial_t\mathbf{E}_h+\beta H\partial_t\mathbf{E}_h^H \tag{15}$$

where using option (10)–(11) to deal with derivatives of order higher than 2, $\mathbf{E}_h$ and $\mathbf{E}_h^H$ can in turn be obtained from the discrete variational equations

$$\int_\Omega \mathbf{E}_h\cdot\mathbf{F}_h\,\mathrm{d}\mathbf{x}=-\int_\Omega(\nabla\cdot\mathbf{F}_h)(\nabla\cdot\mathbf{u}_h)\,\mathrm{d}\mathbf{x}+\int_{\Gamma_\mathrm{I}}(\mathbf{n}\cdot\mathbf{F}_h)(\nabla\cdot\mathbf{u}_h)\,\mathrm{d}\mathbf{x} \tag{16}$$

$$\int_\Omega \mathbf{E}_h^H\cdot\mathbf{F}_h^H\,\mathrm{d}\mathbf{x}=-\int_\Omega(\nabla\cdot\mathbf{F}_h^H)(\nabla\cdot(H\mathbf{u}_h))\,\mathrm{d}\mathbf{x}+\int_{\Gamma_\mathrm{I}}(\mathbf{n}\cdot\mathbf{F}_h^H)(\nabla\cdot(H\mathbf{u}_h))\,\mathrm{d}\mathbf{x} \tag{17}$$

which must hold for all test functions $\mathbf{F}_h\in V_h^2$, $\mathbf{F}_h^H\in V_h^2$, again satisfying the adequate boundary conditions.

The standard Galerkin finite element approximation to problem (8)–(9) and (10)–(11) is (12)–(13) and (16)–(17). It is the main goal of this work to show that it is unstable and to devise a modification to enhance its stability properties. Before that, let us consider its matrix structure.

3.2. Matrix formulation

The discrete finite element problem (12)–(13) leads to the following system of ordinary differential equations:

$$M_\eta\dot{\eta}+K_{12}u+\varepsilon K_{11}(u)\eta+\mu^2(K_{13}E+K_{13}^HE^H)=0 \tag{18}$$

$$M_u\dot{u}+K_{21}\eta+\varepsilon K_{22}(u)u+\mu^2(K_{23}\dot{E}+K_{23}^H\dot{E}^H)=0 \tag{19}$$

where η denotes here the vector of nodal unknowns of the elevation function, u of the velocity and E and E^H of the vector fields $\mathbf{E}_h$ and $\mathbf{E}_h^H$. The dot denotes time differentiation and the identification of the different matrices in (18)–(19) with the terms from where they come in (12)–(13) is straightforward.

Similarly, Equations (16)–(17) can be expressed in the matrix form as

$$M_EE=K_{31}u,\quad E=M_E^{-1}K_{31}u \tag{20}$$

$$M_EE^H=K_{31}^Hu,\quad E^H=M_E^{-1}K_{31}^Hu \tag{21}$$

Therefore, inserting the expressions of E and E^H into (18)–(19) yields a system of ordinary differential equations with the matrix structure

$$M_\eta\dot{\eta}+K_{12}u+\varepsilon K_{11}(u)\eta+\mu^2K_{13}^*u=0 \tag{22}$$

$$M_u\dot{u}+K_{21}\eta+\varepsilon K_{22}(u)u+\mu^2K_{23}^*\dot{u}=0 \tag{23}$$

DOI: 10.1002/fld

with matrices K_{13}^* and K_{23}^* given by

$$K_{13}^* = K_{13} M_E^{-1} K_{31} + K_{13}^H M_E^{-1} K_{31}^H$$

$$K_{23}^* = K_{23} M_E^{-1} K_{31} + K_{23}^H M_E^{-1} K_{31}^H$$

Even though the stabilized formulation we will present later on will change the expression of the different matrices appearing in (22)–(23), the structure of the resulting system of ordinary differential equations will be the same. In Section 5 we will present a time integration scheme of predictor–corrector type, which will allow us to integrate in time taking into account the non-linearity of the system.

3.3. Numerical problems

Apart from the challenge of designing a time integration scheme capable to capture the frequencies involved in the physics modeled by the equations considered, there are two numerical difficulties associated with the Galerkin method. One of them is related to the least-squares projection involved in (22)–(23) and the other is the stability of the formulation.

Equation for the intermediate fields $\mathbf{E}$ and $\mathbf{E}^H$: Matrix M_E is not an M-matrix, since in general the off-diagonal components are greater than zero (recall that a matrix M is said to be an M-matrix if $M_{ij} \leqslant 0$ for $i \neq j$ and $M_{ij}^{-1} \geqslant 0$ for all i, j). Therefore, it is not guaranteed that the components of E have the same sign as the components of $K_{31}u$ (and similarly for E^H and $K_{31}^H u$). This is numerically reflected by the fact that the components of E (and E^H) may oscillate from one node of the finite element mesh to the other in a completely unphysical way when the components of $K_{31}u$ change abruptly from one node to its neighbors. This is the so-called *Gibbs phenomenon*. A way to avoid this problem is to use a diagonal expression for M_E, which can be obtained from a nodal numerical integration rule. In the case of linear elements, this coincides with the classical mass lumping technique.

Compatibility of the interpolation of η_h *and* $\mathbf{u}_h$: The question that arises once the discrete problem is set is whether it is stable or not. It turns out that for the equal-order interpolation for η_h and $\mathbf{u}_h$ considered the answer is that there is not enough stability from the numerical point of view. The explanation of this instability in the linear case can be found in [4]. The main idea is as follows. For the continuous problem, one may take as test functions in (8)–(9) $\xi = \nabla \cdot \mathbf{u}$ and $\mathbf{v} = \nabla \eta$. This yields control over the divergence of the velocity and the gradient of the elevation. However, for the discrete problem (12)–(13) *it is not possible to take* $\xi_h = \nabla \cdot \mathbf{u}_h$ *and* $\mathbf{v}_h = \nabla \eta_h$, since for continuous interpolations $\nabla \cdot \mathbf{u}_h$ and $\nabla \eta_h$ are obviously discontinuous and do not belong to the space of test functions for elevation and velocity, respectively.

4. STABILIZED FINITE ELEMENT METHOD

4.1. General framework

In this section, we present a stabilized finite element method aimed to overcome the instability problems of the standard Galerkin method. This formulation is an extension of [4] to account for convection and non-linearity.

DOI: 10.1002/fld

We start considering an abstract linear first-order partial differential equation of the form

$$\partial_t \mathbf{U} + \mathbf{A}_i \partial_i \mathbf{U} = \mathbf{F} \tag{24}$$

in which the linear non-dispersive problem can be recast. In (24), $\mathbf{A}_i$ are $n_{\text{unk}} \times n_{\text{unk}}$ matrices, $i = 1, 2, \ldots, d$, and $\mathbf{F}$ is a vector of n_{unk} components, n_{unk} being the number of scalar unknowns in $\mathbf{U}$ and d the space dimension (in our case, $d = 2$). It is understood that repeated subscripts sum over the number of space dimensions.

The stabilized finite element method we will present has its roots in the variational multiscale decomposition proposed in [3]. Let us split the unknown $\mathbf{U}$ as $\mathbf{U} = \mathbf{U}_h + \mathbf{U}'$, where $\mathbf{U}_h$ is the finite element solution we are looking for and $\mathbf{U}'$ the component of $\mathbf{U}$ that cannot be captured by the finite element mesh. We will call it *subgrid scale* or, simply, *subscale*. The idea is that an approximation for $\mathbf{U}'$ will lead to a problem for $\mathbf{U}_h$ with enhanced stability properties with respect to the standard Galerkin method.

Let us consider the problem posed for $\mathbf{U}(\cdot, t)$. The inner product in $\mathbf{L}^2(\Omega)$ is denoted by $(\cdot, \cdot)$. The weak form of the problem consists in finding $\mathbf{U}_h$ and $\mathbf{U}'$ such that

$$(\partial_t \mathbf{U}_h, \mathbf{V}_h) + (\partial_t \mathbf{U}', \mathbf{V}_h) + (\mathbf{A}_i \partial_i \mathbf{U}_h, \mathbf{V}_h) + (\mathbf{A}_i \partial_i \mathbf{U}', \mathbf{V}_h) = (\mathbf{F}, \mathbf{V}_h) \tag{25}$$

$$(\partial_t \mathbf{U}_h, \mathbf{V}') + (\partial_t \mathbf{U}', \mathbf{V}') + (\mathbf{A}_i \partial_i \mathbf{U}_h, \mathbf{V}') + (\mathbf{A}_i \partial_i \mathbf{U}', \mathbf{V}') = (\mathbf{F}, \mathbf{V}') \tag{26}$$

for all $\mathbf{V}_h$ in the finite element space and $\mathbf{V}'$ in the space of subscales.

Let us start introducing the approximations to compute $\mathbf{U}'$ and, as a consequence, leading to the stabilized finite element problem for $\mathbf{U}_h$. First of all, we consider that both $\mathbf{U}_h$ and $\mathbf{U}'$ are continuous across interelement boundaries and that $\mathbf{U}'$ varies in time much more slowly than $\mathbf{U}_h$, so that its time derivative can be neglected. This is defined in [16] as the assumption of *quasi-static* subscales. Problem (25)–(26) can be expressed as

$$(\partial_t \mathbf{U}_h, \mathbf{V}_h) + (\mathbf{A}_i \partial_i \mathbf{U}_h, \mathbf{V}_h) - (\mathbf{U}', \mathbf{A}_i^{\mathrm{t}} \partial_i \mathbf{V}_h) = (\mathbf{F}, \mathbf{V}_h) \tag{27}$$

$$P'(\mathbf{A}_i \partial_i \mathbf{U}') = P'(\mathbf{F} - (\partial_t \mathbf{U}_h + \mathbf{A}_i \partial_i \mathbf{U}_h)) =: \mathbf{R}_h \tag{28}$$

where P' is the L^2-projection onto the space of subscales.

Remark 2
It is also possible to consider the subscales to be time dependent, and therefore to integrate them in time. This approach is explained in [19, 20] for the incompressible Navier–Stokes equations, where it is shown that it improves considerably the time stability of the resulting formulation. However, our interest now is to improve the stability in spatial norms, and therefore we will not pursue this approach here.

Equation (28) needs to be approximated to obtain a closed-form expression for $\mathbf{U}'$ which, once inserted into (27), will lead to the stabilized finite element problem for $\mathbf{U}_h$. Observe that $\mathbf{R}_h$ in (28) can be considered as the residual of the finite element approximation projected onto the space of subscales.

The basic heuristic idea is to consider that since $\mathbf{U}'$ is the component of the unknown unresolved by the finite element space, its Fourier transform must be dominated by wave numbers of the form $(1/h)\mathbf{k}$, where $\mathbf{k} = (k_1, \ldots, k_d)$ is dimensionless and of order $\mathcal{O}(1)$ and h is the mesh size.

Let $\mathbf{M}$ be a symmetric and positive matrix that defines an inner product in the space of forcing terms, and let $|\cdot|_M$ be the norm with respect to this matrix, that is, $|\mathbf{F}|_M^2 = \mathbf{F}^{\mathrm{t}}\mathbf{M}\mathbf{F}$ (note that $\mathbf{F}^{\mathrm{t}}\mathbf{F}$

DOI: 10.1002/fld

in general is not even dimensionally meaningful). Likewise, let $\|\cdot\|_M$ be the L^2-norm of $|\cdot|_M$. The Fourier transform of a function f will be denoted as $\hat{f}$. Taking this Fourier transform of (28) yields

$$\mathscr{S}(\mathbf{k})\hat{\mathbf{U}}' \equiv -\mathrm{i}\frac{1}{h}k_j\mathbf{A}_j\hat{\mathbf{U}}' = \hat{\mathbf{R}}_h$$

and then taking the M-norm of this complex-valuated algebraic equation yields

$$\hat{\mathbf{U}}'^{\mathrm{t}}\mathscr{S}(\mathbf{k})^{\mathrm{t}}\mathbf{M}\mathscr{S}(\mathbf{k})\hat{\mathbf{U}}' = \hat{\mathbf{U}}'^{\mathrm{t}}\left(\frac{1}{h^2}k_ik_j\mathbf{A}_i^{\mathrm{t}}\mathbf{M}\mathbf{A}_j\right)\hat{\mathbf{U}}' = \hat{\mathbf{R}}_h^{\mathrm{t}}\mathbf{M}\hat{\mathbf{R}}_h \tag{29}$$

From this expression we obtain

$$\|\hat{\mathbf{R}}_h\|_M^2 = \int |\hat{\mathbf{R}}_h|_M^2 \mathrm{d}\mathbf{k} = \int |\mathscr{S}(\mathbf{k})\hat{\mathbf{U}}'|_M^2 \mathrm{d}\mathbf{k} \leqslant \int |\mathscr{S}(\mathbf{k})|_M^2 |\hat{\mathbf{U}}'|_M^2 \mathrm{d}\mathbf{k}$$

$$= \int |\mathscr{S}(\mathbf{k}^0)|_M^2 |\hat{\mathbf{U}}'|_M^2 \mathrm{d}\mathbf{k} = |\mathscr{S}(\mathbf{k}^0)|_M^2 \|\hat{\mathbf{U}}'\|_M^2$$

where $\mathbf{k}^0$ is a wave number whose existence is guaranteed by the mean value theorem and the integrals extend over the wave number space.

Let us assume now that the subscales are *approximated* as $\mathbf{U}' = \boldsymbol{\tau}\mathbf{R}_h$, where $\boldsymbol{\tau}$ is a symmetric and positive-definite matrix to be determined. A similar calculation to the previous one would yield $\|\hat{\mathbf{R}}_h\|_M^2 \leqslant |\boldsymbol{\tau}^{-1}|_M^2 \|\hat{\mathbf{U}}'\|_M^2$. This suggests to take $\boldsymbol{\tau}$ such that $|\boldsymbol{\tau}^{-1}|_M^2 = |\mathscr{S}(\mathbf{k}^0)|_M^2$, that is to say,

$$\sup_{|\mathbf{X}|_M=1} \mathbf{X}^{\mathrm{t}}\boldsymbol{\tau}^{-1}\mathbf{M}\boldsymbol{\tau}^{-1}\mathbf{X} = \sup_{|\mathbf{X}|_M=1} \mathbf{X}^{\mathrm{t}}\frac{1}{h^2}(k_i^0k_j^0\mathbf{A}_i^{\mathrm{t}}\mathbf{M}\mathbf{A}_j)\mathbf{X} \tag{30}$$

Of course $\mathbf{k}^0$ is unknown, and their components have to be understood as algorithmic constants. The hope is that the approximated subscale will bound the residual of the finite element solution as the exact subscale bounds the residual of the finite element component of the exact solution. This bound will not be exactly the same, but will have the same asymptotic behavior in terms of h and the coefficients of the equation to be solved.

A practical way to impose condition (30) is to compute the spectrum of matrices $\boldsymbol{\tau}^{-1}\mathbf{M}\boldsymbol{\tau}^{-1}$ and $h^{-2}(k_i^0k_j^0\mathbf{A}_i^{\mathrm{t}}\mathbf{M}\mathbf{A}_j)$ *with respect to matrix* $\mathbf{M}$ and impose that at least the spectral radius be the same. We will analyze different particular cases. In the simplest, the spectra of both matrices will be the same, in the second one only two of the three eigenvalues will coincide, and in the third one only the largest eigenvalue, i.e. the spectral radius, will be coincident.

The final problem for the finite element component of the unknown is obtained by inserting expression $\mathbf{U}' = \boldsymbol{\tau}\mathbf{R}_h$ into (27). Noting that the approximation for the subscale is *local* to each element and using the notation $\langle\cdot,\cdot\rangle_{\Omega^e}$ for the integral of the product of two functions on Ω^e, the final discrete variational equation is

$$(\partial_t\mathbf{U}_h, \mathbf{V}_h) + (\mathbf{A}_i\partial_i\mathbf{U}_h, \mathbf{V}_h) - (\mathbf{F}, \mathbf{V}_h)$$

$$+\sum_{e=1}^{n_{\mathrm{el}}} \langle \mathbf{A}_i^{\mathrm{t}}\partial_i\mathbf{V}_h, \boldsymbol{\tau}P'(\partial_t\mathbf{U}_h + \mathbf{A}_i\partial_i\mathbf{U}_h - \mathbf{F})\rangle_{\Omega^e} = 0 \tag{31}$$

DOI: 10.1002/fld

We will apply now this general framework to the problem we are considering. The first step is to design matrix $\boldsymbol{\tau}$. We will do this in the case of the linearized non-dispersive problem.

4.2. Application to the linearized non-dispersive problem

4.2.1. Stabilization parameters. To motivate the design of the stabilization parameters we propose, let us consider the non-dispersive problem and linearized using a constant velocity field $\mathbf{u}_0$, so that the differential equations of the problem are

$$\partial_t \eta + H\nabla\cdot\mathbf{u} + \varepsilon\mathbf{u}_0\cdot\nabla\eta = f_\eta \tag{32}$$

$$\partial_t \mathbf{u} + g\nabla\eta + \varepsilon\mathbf{u}_0\cdot\nabla\mathbf{u} = \mathbf{f}_u \tag{33}$$

In this case, the advection matrices $\mathbf{A}_i$, $i=1,2$, are given by

$$\mathbf{A}_1 = \begin{bmatrix} \varepsilon u_{0,1} & H & 0 \\ g & \varepsilon u_{0,1} & 0 \\ 0 & 0 & \varepsilon u_{0,1} \end{bmatrix}, \quad \mathbf{A}_2 = \begin{bmatrix} \varepsilon u_{0,2} & 0 & H \\ 0 & \varepsilon u_{0,2} & 0 \\ g & 0 & \varepsilon u_{0,2} \end{bmatrix} \tag{34}$$

and the scaling matrix $\mathbf{S}$ is given by

$$\mathbf{S} = \begin{bmatrix} \dfrac{g}{H} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{35}$$

In this particular case, the scaling matrix defines also an inner product in the force space, that is to say, the same matrix makes dimensionally well defined the matrix product $\mathbf{V}^{\mathrm{t}}\mathbf{S}\mathbf{F}$ and $\mathbf{F}^{\mathrm{t}}\mathbf{S}\mathbf{F}$. This dimensionality can be checked as follows:

$$\mathbf{V}^{\mathrm{t}}\mathbf{S}\mathbf{F} = \frac{g}{H}\xi f_\eta + v f_u$$

$$\left[\frac{g}{H}\xi f_\eta\right] = LT^{-2}L^{-1}LLT^{-1} = L^2T^{-3}$$

$$[v f_u] = LT^{-1}LT^{-2} = L^2T^{-3}$$

$$\mathbf{F}^{\mathrm{t}}\mathbf{S}\mathbf{F} = \frac{g}{H}f_\eta^2 + f_u^2$$

$$\left[\frac{g}{H}f_\eta^2\right] = LT^{-2}L^{-1}L^2T^{-2} = L^2T^{-4}$$

$$[f_u^2] = L^2T^{-4}$$

where the brackets $[\cdot]$ denote a dimensional group and L the length and T the time.

Our objective now is to design the matrix of stabilization parameters $\boldsymbol{\tau}$ so that (30) is satisfied. In fact, the optimal situation would be to choose $\boldsymbol{\tau}$ satisfying $\boldsymbol{\tau}^{-1}\mathbf{M}\boldsymbol{\tau}^{-1} = h^{-2}(k_i^0 k_j^0 \mathbf{A}_i^{\mathrm{t}}\mathbf{M}\mathbf{A}_j)$. However, we will also try to choose $\boldsymbol{\tau}$ as simple as possible. In particular, we shall see that it is possible to take $\boldsymbol{\tau}$ *diagonal* and satisfy condition (30), although the equality just mentioned *will not hold*.

DOI: 10.1002/fld

For the matrices given by (34) and (35), it is found that

$$\frac{1}{h^2}k_i^0k_j^0\mathbf{A}_i^{\mathrm{t}}\mathbf{S}\mathbf{A}_j$$

$$=\frac{1}{h^2}\begin{bmatrix} g^2|\mathbf{k}^0|^2+\varepsilon^2(\mathbf{k}^0\cdot\mathbf{u}_0)^2\dfrac{g}{H} & 2\varepsilon(\mathbf{k}_0\cdot\mathbf{u}_0)gk_1^0 & 2\varepsilon(\mathbf{k}_0\cdot\mathbf{u}_0)gk_2^0 \\ 2\varepsilon(\mathbf{k}_0\cdot\mathbf{u}_0)gk_1^0 & gH(k_1^0)^2+\varepsilon^2(\mathbf{k}^0\cdot\mathbf{u}_0)^2 & gHk_1^0k_2^0 \\ 2\varepsilon(\mathbf{k}_0\cdot\mathbf{u}_0)gk_2^0 & gHk_1^0k_2^0 & gH(k_2^0)^2+\varepsilon^2(\mathbf{k}^0\cdot\mathbf{u}_0)^2 \end{bmatrix}$$

When $\mathbf{u}_0=\mathbf{0}$ this matrix is singular. This is due to the fact that it does not contain the information on the boundary conditions that allows to invert the differential operator from where it comes.

Let us denote by $\mathrm{Spec}_M(\mathbf{B})$ the spectrum of a matrix $\mathbf{B}$ with respect to the inner product $\mathbf{M}$, that is, the set of eigenvalues λ of the generalized eigenvalue problem $\mathbf{BU}=\lambda\mathbf{MU}$. It turns out that

$$\mathrm{Spec}_S(k_i^0k_j^0\mathbf{A}_i^{\mathrm{t}}\mathbf{S}\mathbf{A}_j)=\{(\varepsilon(\mathbf{k}^0\cdot\mathbf{u}_0)+\sqrt{gH}|\mathbf{k}^0|)^2,\varepsilon^2(\mathbf{k}^0\cdot\mathbf{u}_0)^2,(\varepsilon(\mathbf{k}^0\cdot\mathbf{u}_0)-\sqrt{gH}|\mathbf{k}^0|)^2\}$$

Obviously, $\mathbf{k}^0$ is unknown, so that we may consider its norm and its angle with $\mathbf{u}_0$ algorithmic constants and write the previous expression as

$$\mathrm{Spec}_S(k_i^0k_j^0\mathbf{A}_i^{\mathrm{t}}\mathbf{S}\mathbf{A}_j)=\{(C_1\varepsilon|\mathbf{u}_0|+C_2\sqrt{gH})^2,C_1^2\varepsilon^2|\mathbf{u}_0|^2,(C_1\varepsilon|\mathbf{u}_0|-C_2\sqrt{gH})^2\}$$

The values of C_1 and C_2 (not to be confused with the constants of the model given by (7)) need to be determined from numerical experiments.

Let us now assume that $\boldsymbol{\tau}$ is diagonal. Since the two scalar equations for the velocity components have the same form, we may take it as $\boldsymbol{\tau}=\mathrm{diag}(\tau_\eta,\tau_u,\tau_u)$. Clearly, $\boldsymbol{\tau}^{-1}\mathbf{S}\boldsymbol{\tau}^{-1}$ will also be diagonal, and it is impossible that it behaves as $h^{-2}(k_i^0k_j^0\mathbf{A}_i^{\mathrm{t}}\mathbf{S}\mathbf{A}_j)$. However, since $\mathrm{Spec}_S(\boldsymbol{\tau}^{-1}\mathbf{S}\boldsymbol{\tau}^{-1})=\{\tau_\eta^{-2},\tau_u^{-2},\tau_u^{-2}\}$, a way to choose $\boldsymbol{\tau}$ satisfying (30) is to take $\tau_\eta=\tau_u\equiv\tau$ and

$$\boldsymbol{\tau}=\tau\mathbf{I},\quad \tau=\frac{h}{C_1\varepsilon|\mathbf{u}_0|+C_2\sqrt{gH}} \tag{36}$$

It is observed that

- In the 1D problem without convection ($\varepsilon=0$) it turns out that $\boldsymbol{\tau}^{-1}\mathbf{S}\boldsymbol{\tau}^{-1}=h^{-2}(k_i^0k_j^0\mathbf{A}_i^{\mathrm{t}}\mathbf{S}\mathbf{A}_j)$. This is the optimal situation.
- In the 2D problem without convection ($\varepsilon=0$), matrix $h^{-2}(k_i^0k_j^0\mathbf{A}_i^{\mathrm{t}}\mathbf{S}\mathbf{A}_j)$ has two eigenvalues equal to the diagonal entries of $\boldsymbol{\tau}^{-1}\mathbf{S}\boldsymbol{\tau}^{-1}$, and the third one is zero.
- In the 2D case with convection, the diagonal entries of $\boldsymbol{\tau}^{-1}\mathbf{S}\boldsymbol{\tau}^{-1}$ coincide with the largest eigenvalue of $h^{-2}(k_i^0k_j^0\mathbf{A}_i^{\mathrm{t}}\mathbf{S}\mathbf{A}_j)$.

In all the cases, condition (30) is satisfied.

4.2.2. Stabilized formulation. Let us consider again the linearized non-dispersive problem defined by Equations (32)–(33). As boundary conditions we will take $\eta=0$ on the whole $\partial\Omega$. In order

DOI: 10.1002/fld

to have a well-posed problem for all $\varepsilon>0$, we need to prescribe boundary conditions for $\mathbf{u}$ on the inflow part of the boundary, that is to say, where $\mathbf{n}\cdot\mathbf{u}_0<0$. We take $\mathbf{u}=\mathbf{0}$ there. With these boundary conditions, we will have that

$$\int_\Omega \eta(\mathbf{u}_0\cdot\nabla\eta)\,\mathrm{d}\mathbf{x}=\int_{\partial\Omega}\mathbf{n}\cdot\mathbf{u}_0\frac{\eta^2}{2}\,\mathrm{d}\mathbf{x}=0 \tag{37}$$

$$\int_\Omega \mathbf{u}\cdot(\mathbf{u}_0\cdot\nabla\mathbf{u})\,\mathrm{d}\mathbf{x}=\int_{\partial\Omega}\mathbf{n}\cdot\mathbf{u}_0\frac{|\mathbf{u}|^2}{2}\,\mathrm{d}\mathbf{x}\geqslant 0 \tag{38}$$

Using the matrix of stabilization parameters given by (36) and taking into account that for H and $\mathbf{u}_0$ constant the stabilization parameter τ will be the same for all the elements of the finite element mesh, the general stabilized formulation (31) becomes

$$\begin{aligned}0=&\frac{g}{H}(\partial_t\eta_h,\xi_h)-g(\mathbf{u}_h,\nabla\xi_h)-\frac{g}{H}(\varepsilon\mathbf{u}_0\eta_h,\nabla\xi_h)-\frac{g}{H}(f_\eta,\xi_h)\\&+(\partial_t\mathbf{u}_h,\mathbf{v}_h)+g(\nabla\eta_h,\mathbf{v}_h)+(\varepsilon\mathbf{u}_0\cdot\nabla\mathbf{u}_h,\mathbf{v}_h)-(\mathbf{f}_u,\mathbf{v}_h)\\&+\tau\frac{g}{H}(P'(\partial_t\eta_h+H\nabla\cdot\mathbf{u}_h+\varepsilon\mathbf{u}_0\cdot\nabla\eta_h-f_\eta),H\nabla\cdot\mathbf{v}_h+\varepsilon\mathbf{u}_0\cdot\nabla\xi_h)\\&+\tau(P'(\partial_t\mathbf{u}_h+g\nabla\eta_h+\varepsilon\mathbf{u}_0\cdot\nabla\mathbf{u}_h-\mathbf{f}_u),g\nabla\xi_h+\varepsilon\mathbf{u}_0\cdot\nabla\mathbf{v}_h)\end{aligned} \tag{39}$$

which must hold for all test functions ξ_h and $\mathbf{v}_h$ in the appropriate spaces. The terms in the first two rows of this variational equation correspond to the Galerkin contribution, whereas those multiplied by τ should provide stabilization. This single equation can obviously be expressed in the more usual form of two discrete variational equations.

4.3. Stabilized finite element method for the general problem

4.3.1. Stabilization of non-linear problems. The difficulty to extend stabilized finite element methods to non-linear problems is to define which is the *linear* operator that has to be applied to the test functions in the stabilization terms. This depends on the linearization technique employed. To explain this, let us consider an abstract *stationary* problem of the form

$$A(u)=f$$

with $u\in V$, and suppose that we solve it iteratively. In this equation, $A(u)$ is a non-linear operator and f a given forcing term. Given a guess for the solution, that we still denote by u, we compute a correction δu from the scheme

$$L(u)\delta u=f-A(u)$$

where $L(u)$ is the iteration operator of the iterative process. If $\langle\cdot,\cdot\rangle$ denotes the pairing between V and its dual, the weak form of the problem can be expressed in the abstract form as

$$\langle v,L(u)\delta u\rangle=\langle v,f-A(u)\rangle\qquad\forall v\in V$$

Let us consider the test function v split as $\bar{v}+v'$, for a certain decomposition $V=\bar{V}\oplus V'$ that can be associated with resolvable and unresolvable scales of a numerical approximation. Similarly,

DOI: 10.1002/fld

consider that $\delta u = \delta\bar{u} + \delta u'$, and assume that $\delta u'$ is approximated by

$$\delta u' = \tau_L P'(f - A(u) - L(u)\delta\bar{u})$$

where τ_L is a matrix that approximates $(P'L(u))^{-1}$ and P' is the projection onto V'. The equation to be solved projected onto $\bar{V}$ reads

$$\langle \bar{v}, L(u)\delta\bar{u}\rangle + \langle L^*(u)\bar{v}, \tau_L P'(f - A(u) - L(u)\delta\bar{u})\rangle = \langle \bar{v}, f - A(u)\rangle$$

where $L^*(u)$ is the adjoint of $L(u)$. At convergence, $\delta\bar{u} = 0$, so that the problem finally solved is

$$\langle \bar{v}, A(u)\rangle + \langle -L^*(u)\bar{v}, \tau_L P' A(u)\rangle = \langle \bar{v}, f\rangle + \langle -L^*(u)\bar{v}, \tau_L P' f\rangle \tag{40}$$

It is observed that the terms that could be considered associated with stabilization depend on operator $L(u)$. An obvious choice would be to use the tangent operator, that is to say, $L(u) = A'(u)$, where $A'(u)$ is the (Fréchet) derivative of $A(u)$. In the particular case $A(u) = B(u)u$, it is possible to use a fixed point iteration and take $L(u) = B(u)$. This is what we do next.

4.3.2. Application to the modified Boussinesq equations. The design of the matrix of stabilization parameters presented has been motivated in a linearized version of the problem as the last one described, assuming given the velocity in $\nabla\cdot(\eta\mathbf{u})$ of (1) and also the advective velocity in $\mathbf{u}\cdot\nabla\mathbf{u}$. Thus, our stabilization method will be of form (40) with $L(u)$ the operator obtained with a given velocity in the first component and in the advective term of the second component of this vector operator. The resulting formulation consists of finding η_h and $\mathbf{u}_h$ in the appropriate finite element spaces such that

$$\begin{aligned}
&\frac{g}{H_0}(\partial_t\eta_h, \xi_h) - \frac{g}{H_0}(H\mathbf{u}_h, \nabla\xi_h) - \frac{g}{H_0}\varepsilon(\eta_h\mathbf{u}_h, \nabla\xi_h) - \frac{g}{H_0}\mu^2(\mathbf{J}_{\eta,h}, \nabla\xi_h)\\
&\quad + (\partial_t\mathbf{u}_h, \mathbf{v}_h) + g(\nabla\eta_h, \mathbf{v}_h) + \varepsilon(\mathbf{u}_h\cdot\nabla\mathbf{u}_h, \mathbf{v}_h) + \mu^2(\mathbf{J}_{u,h}, \mathbf{v}_h)\\
&\quad + \frac{g}{H_0}\sum_{e=1}^{n_{\mathrm{el}}}\tau^e\langle P'(\partial_t\eta_h + \nabla\cdot(H\mathbf{u}_h) + \varepsilon\nabla\cdot(\eta_h\mathbf{u}_h) + \mu^2\nabla\cdot\mathbf{J}_{\eta,h}),\\
&\qquad \nabla\cdot(H\mathbf{v}_h) + \varepsilon\nabla\cdot(\xi_h\mathbf{u}_h)\rangle_{\Omega^e}\\
&\quad + \sum_{e=1}^{n_{\mathrm{el}}}\tau^e\langle P'(\partial_t\mathbf{u}_h + g\nabla\eta_h + \varepsilon\mathbf{u}_h\cdot\nabla\mathbf{u}_h + \mu^2\mathbf{J}_{u,h}), g\nabla\xi_h + \varepsilon\mathbf{u}_h\cdot\nabla\mathbf{v}_h\rangle_{\Omega^e} = 0
\end{aligned} \tag{41}$$

for all test functions ξ_h and $\mathbf{v}_h$. Here, H_0 is a characteristic depth only needed to scale the equations.

Problem (41), with τ given by (36) and the auxiliary equations (14)–(15) and (16)–(17), is the stabilized finite element method we propose for the spatial discretization of the modified Boussinesq equations.

DOI: 10.1002/fld

Remark 3

1. The stabilization parameter τ given by (36) needs to be evaluated now with a characteristic element velocity, for example, the mean over each element. Instead of the mesh size h, the element size h^e has to be used if the mesh is not uniform.
2. It has to be noted that if the space of subscales is taken orthogonal to the finite element space, that is, $P'=P^{\perp}$, then $P'(\partial_t \eta_h)=0$, $P'(\partial_t \mathbf{u}_h)=0$. The mass matrix of the linear system will *not* be modified with respect to the Galerkin method.
3. We have not included the dispersive operator applied to the test function. The hope is that its influence on the stability of the scheme is small. The formulation is in any case consistent, in the sense that the exact solution satisfies also Equation (41) for all test functions ξ_h and $\mathbf{v}_h$.

5. TIME INTEGRATION SCHEME

Problem (41) is a system of non-linear ordinary differential equations that, expressed in a matrix form, reads

$$M\dot{x}=F(x,\dot{x})=A(x)+B\dot{x} \tag{42}$$

where x is the vector of unknowns (elevations and velocities), $\dot{x}$ its time derivative, M is a mass matrix, $A(x)$ a non-linear vector function and B a constant matrix. The linear dependence of $F(x,\dot{x})$ with $\dot{x}$ comes from the additional fields $\mathbf{J}_\eta$ and $\mathbf{J}_u$ defined in (3) and (4), respectively.

The objective now is to present a time integration scheme for (42) taking into account its non-linearity. Let us consider a partition of the time interval $[0,T]$ into time steps of, for simplicity, equal size δt. We denote with superscript n the approximation of a function at time $t^n=n\delta t$.

As basic time integration algorithm, we consider the fourth-order Adams–Moulton method. Once the solution is known until time step n, the solution at time step $n+1$ is computed from

$$x^{n+1}=x^n+\frac{\delta t}{24}M^{-1}(9F^{n+1}+19F^n-5F^{n-1}+F^{n-2})+O(\delta t^4)$$

Once the $O(\delta t^4)$ terms are neglected, this is a non-linear algebraic equation due to the non-linear dependence of F on x. In order to deal with this non-linearity, we may use the following iterative version: given a guess $x^{n+1,i-1}$ for x^{n+1} at iteration $i-1$, the ith iterate is computed as

$$x^{n+1,i}=x^n+\frac{\delta t}{24}M^{-1}(9F^{n+1,i-1}+19F^{n,i-1}-5F^{n-1,i-1}+F^{n-2,i-1})+O(\delta t^4)$$

where

$$F^{k,i-1}=A(x^k)+B\dot{x}^{k,i-1},\quad k=n-2,n-1,n,n+1$$

The approximations used for the time derivative are

$$\dot{x}^{n+1,i-1}=\frac{1}{6\delta t}(11x^{n+1,i-1}-18x^n+9x^{n-1}-2x^{n-2})+O(\delta t^3)$$

$$\dot{x}^{n,i-1}=\frac{1}{6\delta t}(2x^{n+1,i-1}-3x^n-6x^{n-1}+x^{n-2})+O(\delta t^3)$$

DOI: 10.1002/fld

$$\dot{x}^{n-1,i-1}=\frac{1}{6\delta t}(-x^{n+1,i-1}+6x^{n}-3x^{n-1}-2x^{n-2})+O(\delta t^3)$$

$$\dot{x}^{n-2,i-1}=\frac{1}{6\delta t}(2x^{n+1,i-1}-9x^{n}+18x^{n-1}-11x^{n-2})+O(\delta t^3)$$

that are obtained from the classical Taylor expansion around the time level of interest ($n+1, n, n-1$ and $n-2$, respectively). We need to estimate $x^{n+1,0}$ only using x^n, x^{n-1} and x^{n-2}. This can be done by the explicit third-order Adams–Bashforth scheme, which reads

$$x^{n+1,0}=x^{n}+\frac{\delta t}{12}M^{-1}(23F^{n}-16F^{n-1}+5F^{n-2})+O(\delta t^3)$$

where

$$F^{k}=A(x^{k})+B\dot{x}^{k},\quad k=n-2,n-1,n$$

$$\dot{x}^{n}=\frac{1}{2\delta t}(3x^{n}-4x^{n-1}+x^{n-2})+O(\delta t^2)$$

$$\dot{x}^{n-1}=\frac{1}{2\delta t}(x^{n}-x^{n-1})+O(\delta t^2)$$

$$\dot{x}^{n-2}=-\frac{1}{2\delta t}(x^{n}-4x^{n-1}+3x^{n-2})+O(\delta t^2)$$

This completes the definition of a time integration scheme well suited to capture the oscillating flow phenomena described by the equations approximated. For similar schemes, see [6, 8, 10, 21].

It is important to note that, due to the predictor–corrector type of the scheme, convergence problems may be encountered if the time step size is large. A Fourier analysis similar to the one presented in Section 4.2.1 reveals that the critical time step of an explicit time integration scheme must behave as τ defined in (36). Numerical experiments indicate that it is convenient to take δt of the same magnitude as τ in order to avoid convergence problems.

6. NUMERICAL EXAMPLES

In order to test the proposed scheme, it has been implemented into the MANOLO software developed at IH Cantabria, which is an extensive rework of the model presented in [10] with the numerical formulation presented in this paper. MANOLO allows the definition of the computational domain by means of unstructured triangular meshes. Linear interpolation at each element is used for all variables.

This model has been tested in a large variety of situations and it has been found that it shows a good agreement with analytical and experimental data. For a discussion of the quality of the results of this model (using a slightly different stabilization scheme), see [22]. The examples in this paper are more oriented to the comparison of the improvements of the stabilized formulation over the non-stabilized one (where the stabilization parameter τ is considered to be zero).

In order to find the appropriate values of the parameters C_1 and C_2 (see (36)), several tests have been carried out. As a result of that, we have found that $C_1=C_2=300$ give the best performance, avoiding instabilities without adding any appreciable diffusion.

DOI: 10.1002/fld

6.1. Gaussian hump

In this example, the behavior of fluid inside a square basin of perfectly reflective walls is studied. The length of the basin sides is 6 m, and its bottom is located at $z=-0.5$ m (here and below we denote $(x_1, x_2, x_3) \equiv (x, y, z)$). The initial condition for the free surface is given by the formula

$$\eta^0(x, y) = 0.045\, \mathrm{e}^{-2[(x-3)^2+(y-3)^2]}$$

which represents a Gaussian hump centered at the middle of the basin (see Figure 1). The velocity field is zero at $t=0$.

This problem admits analytical solution for the linear, dispersive case and, for this reason, it is commonly used for testing numerical models. In this case, we are interested in the solution of the non-linear, dispersive equations which, to our knowledge, has no analytical solution, so the comparison is done against the FUNWAVE finite difference model [6]. Figure 2 shows that there is a very good agreement between both models.

Owing to the spatial nature of the instabilities originated by the use of equal interpolation for the water elevation and the velocity without stabilization terms, these instabilities are more noticeable in spatial slices of the free surface than in time series over a single point.

Figure 3 shows the comparison of the results of the model with and without the stabilization mechanism. It can be seen that the latter shows spurious wiggles. These wiggles increase its amplitude as time advances and, eventually, interfere with the time integration scheme, which is not able to converge. On the other hand, the stabilized formulation shows a perfectly smooth free surface, as expected.

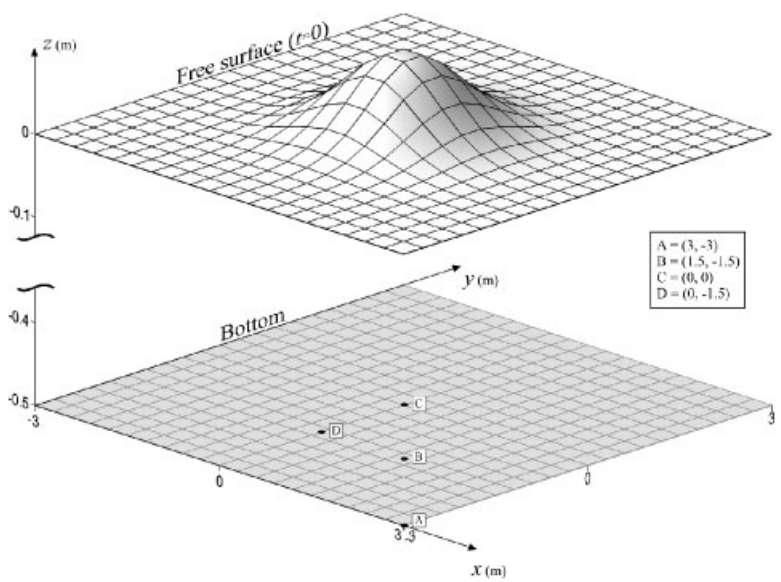

Figure 1. Gaussian hump setup.

DOI: 10.1002/fld

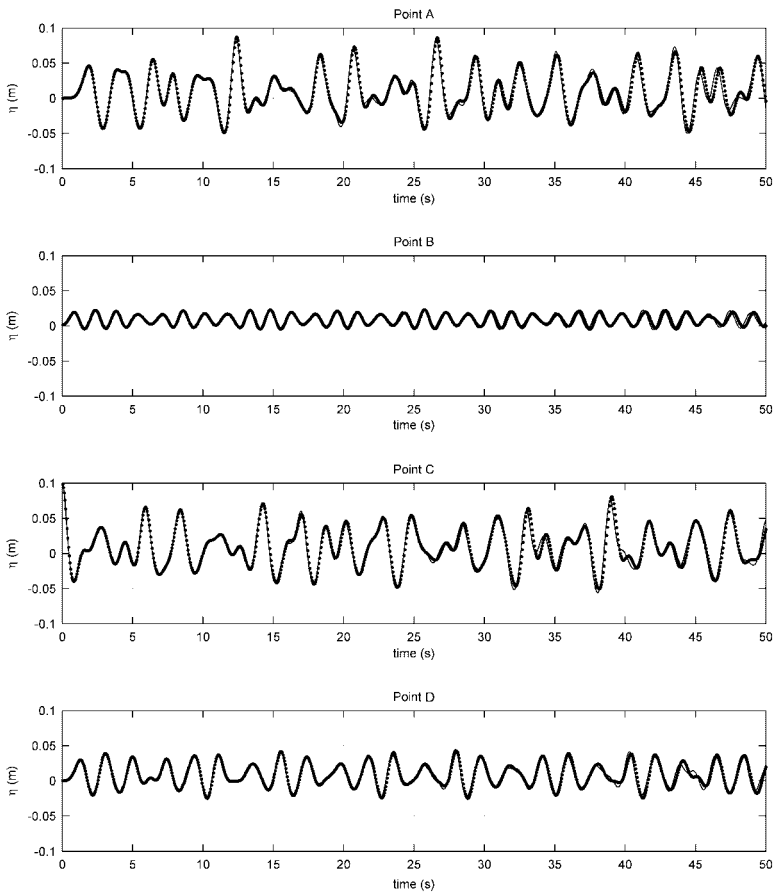

Figure 2. Free surface evolution: continuous line, finite differences model (FUNWAVE); dotted line, proposed scheme (MANOLO).

6.2. Real case: Lastres Harbor

As mentioned in the Introduction, some authors try to overcome the problems associated with the instability of the equations by means of artificial diffusion or *ad hoc* filters. We have found that, although this approach usually gives reasonable results for small meshes and short simulation periods, it fails for long periods of time and big meshes.

Our stabilized formulation allows to compute on large spatial domains and over long periods of time, which is an important issue when dealing with real cases. As an example of this we present a simulation for Lastres harbor, a small harbor in the north of Spain. The mesh used, which consists of 20 456 nodes and 40 492 elements, is shown in Figure 4.

In the wave generation area, a wave is generated by adding or removing water as necessary at each time step. The sponge layer uses a simple Newtonian cooling scheme to dissipate energy, consisting of adding a force proportional to the velocity and in the opposite direction (see [6, 10, 22]).

DOI: 10.1002/fld

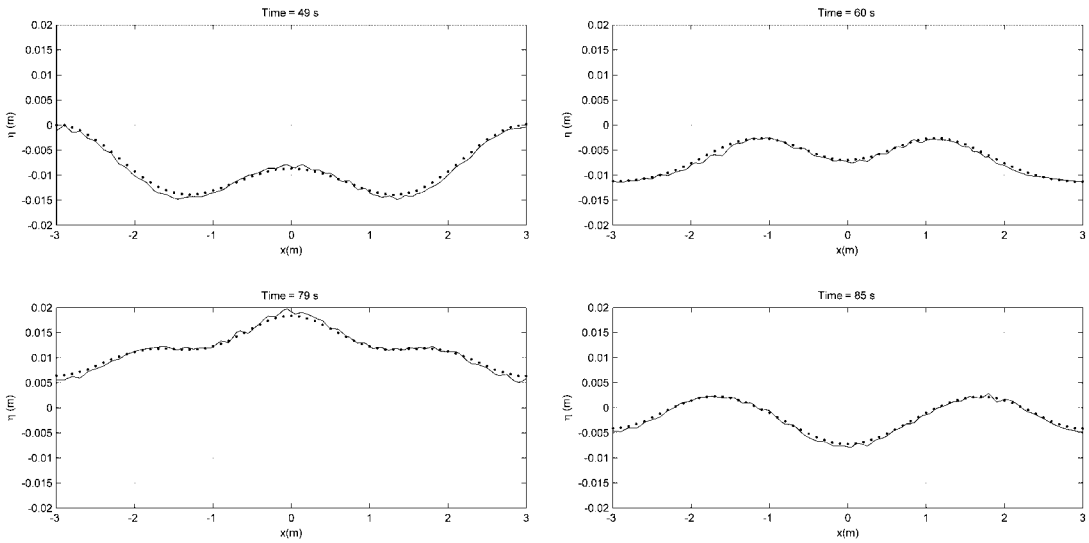

Figure 3. Free surface along the line $y=0$. Continuous line is the model with the non-stabilized equations. Dotted line is the stabilized one.

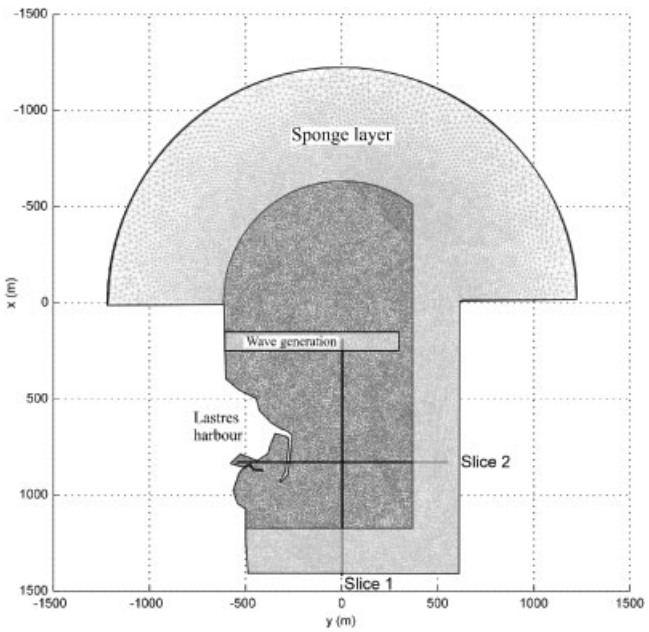

Figure 4. Lastres Harbor mesh.

DOI: 10.1002/fld

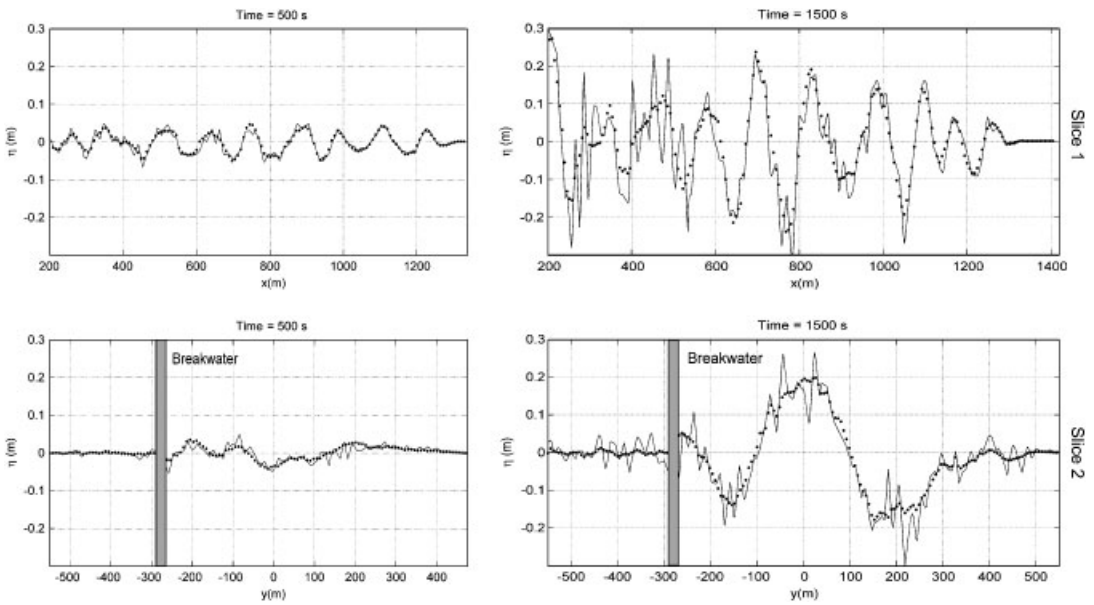

Figure 5. Free surface for Lastres Harbor along two different lines and at times $t = 500$ and 1500 s. Continuous line is the model with the non-stabilized equations. Dotted line is the stabilized one.

Figure 5 shows two slices of the free surface along the lines marked as Slice 1 and Slice 2 in Figure 4 for two different times, $t = 500$ and 1500 s. It is clear that, in the non-stabilized model, the amplitude of the instabilities increases as time advances. From a practical point of view, this results in values of the agitation inside the harbor predicted by the non-stabilized model much higher than the measured ones. On the other hand, the maximum values of agitation predicted by the stabilized model coincide with those measured in the harbor.

In addition, the model with the stabilized formulation allows long simulation times, which is crucial when studying physical processes that take long times to develop, like resonance inside harbors.

7. CONCLUSIONS

In this paper, we have proposed a stabilized finite element method to solve the modified Boussinesq equations. The lack of stability of the Galerkin method for this problem is inherited from the linear wave equation expressed in the mixed form. Here, we have proposed a subgrid scale stabilized method with a closed-form expression for the subscales derived from a Fourier analysis.

The implementation of resulting numerical method is simple, and fits naturally in classical finite element codes. The element structure of the arrays to be assembled does not change. Likewise, standard linearization procedures and time integration schemes can be adapted with no difficulty.

From the experimental point of view, the formulation certainly displays the behavior it was designed for. The 'noise' observed in the Galerkin approximation is removed and smooth discrete solutions are found.

DOI: 10.1002/fld

REFERENCES

1. Nwogu O. Alternative form of Boussinesq equations for near shore wave propagation. *Journal of Waterway, Port, Coastal and Ocean Engineering* (ASCE) 1993; **119**:618–638.
2. Hughes TJR. Multiscale phenomena: Green's function, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized formulations. *Computer Methods in Applied Mechanics and Engineering* 1995; **127**:387–401.
3. Hughes TJR, Feijóo GR, Mazzei L, Quincy JB. The variational multiscale method—a paradigm for computational mechanics. *Computer Methods in Applied Mechanics and Engineering* 1998; **166**:3–24.
4. Codina R. Finite element approximation of the hyperbolic wave equation in mixed form. *Computer Methods in Applied Mechanics and Engineering* 2007; DOI: 10.1016/j.cma.2007.11.006.
5. Abbot MB, Petersen HM, Skovgaard O. On the numerical modelling of short waves in shallow water. *Journal of Hydraulic Research* 1978; **16**:173–203.
6. Wei G, Kirby JT. Time dependent numerical code for extended Boussinesq equations. *Journal of Waterway, Port, Coastal and Ocean Engineering* (ASCE) 1995; **121**:251–261.
7. Langtangen HP, Pedersen G. Computational models for weakly dispersive nonlinear water waves. *Computer Methods in Applied Mechanics and Engineering* 1998; **160**:337–358.
8. Li YS, Liu S-X, Yu Y-X, Lai G-Z. Numerical modeling of the Boussinesq equations by finite element method. *Coastal Engineering* 1999; **37**:97–122.
9. Walkley M, Berzins M. A finite element method for the two-dimensional extended Boussinesq equations. *International Journal for Numerical Methods in Fluids* 2002; **39**:865–885.
10. Woo S-B, Liu PL-F. Finite element model for modified Boussinesq equations. I: model development. *Journal of Waterway, Port, Coastal and Ocean Engineering* (ASCE) 2004; **130**:1–16.
11. Donea J. A Taylor-Galerkin method for convection transport problems. *International Journal for Numerical Methods in Engineering* 1984; **20**:101–119.
12. Peraire J, Zienkiewicz OC, Morgan K. Shallow water problems: a general explicit formulation. *International Journal for Numerical Methods in Engineering* 1986; **22**:547–574.
13. Miglio E, Quarteroni A, Saleri F. Finite element approximation of quasi-3d shallow water equations. *Computer Methods in Applied Mechanics and Engineering* 1999; **174**:355–369.
14. Zienkiewicz OC, Ortiz P. A split-characteristic based finite element model for the shallow water equations. *International Journal for Numerical Methods in Fluids* 1995; **20**:1061–1080.
15. Hauke G. A symmetric formulation for computing transient shallow water flows. *Computer Methods in Applied Mechanics and Engineering* 1998; **163**:111–122.
16. Codina R. Stabilized finite element approximation of transient incompressible flows using orthogonal subscales. *Computer Methods in Applied Mechanics and Engineering* 2002; **191**:4295–4321.
17. Israeli M, Orzag SA. Approximation of radiation boundary conditions. *Journal of Computational Physics* 1981; **42**:115–135.
18. Berenger JP. A perfectly matched layer for the absorption of electromagnetic waves. *Journal of Computational Physics* 1994; **114**:185–200.
19. Codina R, Principe J, Guasch O, Badia S. Time dependent subscales in the stabilized finite element approximation of incompressible flow problems. *Computer Methods in Applied Mechanics and Engineering* 2007; **196**: 2413–2430.
20. Codina R, Principe J. Dynamic subscales in the finite element approximation of thermally coupled incompressible flows. *International Journal for Numerical Methods in Fluids* 2007; **54**:707–730.
21. Woo S-B, Liu PL-F. Finite element model for modified Boussinesq equations. II: application to nonlinear harbor oscillations. *Journal of Waterway, Port, Coastal and Ocean Engineering* (ASCE) 2004; **130**:17–28.
22. Losada IJ, González-Ondina JM, Díaz-Hernandez G, González EM. Numerical modeling of nonlinear resonance of semi-enclosed water bodies: description and experimental validation. *Coastal Engineering* 2008; **55**:21–34.