Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER Computers and Mathematics with Applications 52 (2006) 1243–1268
An International Journal
computers & mathematics with applications
www.elsevier.com/locate/camwa

Open Boundary Control Problem for Navier-Stokes Equations Including a Free Surface: Adjoint Sensitivity Analysis

I. YU. GEJADZE* AND G. J. M. COPELAND
Department of Civil Engineering
University of Strathclyde
John Anderson Building, 107 Rottenrow
Glasgow, G4 0NG, U.K.
<igor.gejadze><g.m.copeland>@strath.ac.uk

Abstract—This paper develops the adjoint sensitivities to the free-surface barotropic Navier-Stokes equations in order to allow for the assimilation of measurements of currents and free-surface elevations into an unsteady flow solution by open-boundary control. To calculate a variation in a surface variable, a mapping is used in the vertical to shift the problem into a fixed domain. A variation is evaluated in the transformed space from the Jacobian matrix of the mapping. This variation is then mapped back into the original space where it completes a tangent linear model. The adjoint equations are derived using the scalar product formulas redefined for a domain with variable bounds. The method is demonstrated by application to an unsteady fluid flow in a one-dimensional open channel in which horizontal and vertical components of velocity are represented as well as the elevation of the free surface (a 2D vertical section model). This requires the proper treatment of open boundaries in both the forward and adjoint models. A particular application is to the construction of a fully three-dimensional coastal ocean model that allows assimilation of tidal elevation and current data. However, the results are general and can be applied in a wider context.

Keywords—Navier-Stokes equations, Free surface, Open boundary, Optimal control, Adjoint equations, Sensitivity analysis, Ocean, Waves.

1. INTRODUCTION

The simulation of water circulation in coastal areas requires the application of either two-dimensional (2D) or three-dimensional (3D) computer flow models. These models calculate

*Author to whom all correspondence should be addressed.
The authors gratefully acknowledge support by The Engineering and Physical Sciences Research Council, Grants GR/R60881, GR/R60898.
The authors are grateful for the help of the following: Dr. S. Neill, Centre for Oceanography, University of Wales, Bangor, for the fsNSE solver used as a prototype code; Prof. V. Shutyaev, Institute of Numerical Mathematics, Russian Academy of Science, Moscow, for essential help in theoretical aspects of this work and valuable comments; Prof. I.M. Navon, Department of Mathematics and School of Computational Science and Information Technology, Florida State University, Tallahassee, for general guidance and a continuous support, advice and comments, which helped to improve the quality of the paper. We also acknowledge our colleagues on this collaborative research project at Imperial College of Science, Technology and Medicine, London: Prof. T. Goddard, Dr. C. Pain, Dr. G. Gorman, Dr. M. Piggot and Dr. Fangxin Fang.

0898-1221/06/$ - see front matter Typeset by AMS-TEX
doi:10.1016/j.camwa.2006.11.004

solutions for either the 2D shallow-water equations (SWE) [1], or for the layer-averaged 3D Navier-Stokes equations (NSE) [2], or for the 3D free-surface Navier-Stokes formulation (fsNSE) using sigma coordinates [3], or for the fully 3D nonhydrostatic fsNSE [4,5], etc. All of these calculations require that the solution is driven by an unsteady inflow Dirichlet boundary condition based on 'known' data. However, in a typical application it is unlikely that measurements are available at the boundary but it is much more likely that data is available at scattered locations within the model domain, measured by current meters and tide gauges, for example. Hence the boundary conditions are actually unknown and must be recovered by a process of adjustments until the model solution agrees with measured data at the internal points. This describes the inverse problem in which data is assimilated into a model solution and boundary conditions are recovered from internal data. The process of adjustment can be systematized by calculating appropriate sensitivities to guide a gradient descend algorithm. In the field of meteorology, the data-assimilation techniques based on optimal control methods appear in early 1970s [6,7]. A general sensitivity theory for nonlinear systems was formulated in [8]. Boundary-control problems for free-surface fluid flow were considered for the SWE [9,10], for a depth-integrated tidal model [11], and for the vorticity equation [12,13], etc. For the primitive hydrostatic equation (PHE) we can refer to [14], where the authors declare a *discrete* adjoint fsNSE option. The *discrete* (consistent) adjoint refers to the adjoint of the discretized model equations. It is quite likely that similar *discrete* adjoint codes have been generated for some other ocean models. We present here a *continuous* (inconsistent) adjoint model, which is an important tool for general qualitative analysis, for considering solvability issues, etc. Also, of course, it is a valid way to construct the adjoint solver for practical applications. This paper presents the adjoint formulation of the nonhydrostatic barotropic fsNSE in 2D vertical section (for the inviscid case this formulation first appeared in [15]). The method is general and can be extended to the 3D baroclinic fsNSE. The novelty of this work is in the complete treatment of the free surface in the adjoint problem and in the clarification of open-boundary conditions in both the forward and adjoint models.

Tidal flows in shallow coastal waters are often represented by solutions to the SWE. This is usually considered to be valid for well-mixed, barotropic conditions. It is clear that in cases where freshwater inflows cause baroclinic conditions a 3D representation is needed which must also include the salinity transport equations. However, 3D models offer advantages even for barotropic conditions because the effects of topographic steering are better represented. For example, flows over or around a submerged step or shoal may cause flow separation and enhanced mixing that can only be represented in 3D. Even flow past a change in coastline direction may result in different flows at the surface where there may be a tendency to separate, than near to the bed where the flow remains attached. The resulting lateral shears cause an increase in mixing that may be significant in a pollutant transport study.

The rapid improvement in reliability and availability of data from coastal waters is driving the need for data-assimilation methods effective in tidally dominated flows. Adjoint models developed for oceanographic (deep water) applications, see for example [16], do not determine sensitivities to variations in the free surface and therefore cannot directly assimilate water-elevation data (although it can be achieved by introducing the measured geopotential surface). The model presented here allows direct assimilation of unsteady water levels by the adjoint method applicable to shallow tidal flows (as well as to deep water flows). This is probably the first fsNSE *continuous* adjoint model that is free from the hydrostatic assumption in any form. Further applications for the method are to the creation of operational coastal flow models that assimilate measurements of current flows and water levels in order to calculate a 'nowcast' of flow conditions at all locations within the model area. This in turn can be used as an initial condition for a short-term forecast.

2. PROBLEM STATEMENT

Let us consider a 2-D free-surface flow in a channel, where the x-axis is directed along the channel, and the z-axis from the channel bed to the surface. Velocities $u = u(x,z,t)$, $w = w(x,z,t)$ are associated to x-, z-axes, respectively (see in Figure 1). The governing equations are as follows:

$$\mathbb{E}_1 := \frac{\partial h}{\partial t} + u(x,h(x,t),t)\frac{\partial h}{\partial x} - w(x,h(x,t),t) = 0, \tag{1}$$

$$\mathbb{E}_2 := \frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} = 0, \tag{2}$$

$$\mathbb{E}_3 := \frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + w\frac{\partial u}{\partial z} + \frac{\partial p}{\partial x} = \frac{\partial}{\partial x}\left(\mu_h \frac{\partial u}{\partial x}\right) + \frac{\partial}{\partial z}\left(\mu_v \frac{\partial u}{\partial z}\right), \tag{3}$$

$$\mathbb{E}_4 := \frac{\partial w}{\partial t} + u\frac{\partial w}{\partial x} + w\frac{\partial w}{\partial z} + \frac{\partial p}{\partial z} + g = \frac{\partial}{\partial x}\left(\mu_h \frac{\partial w}{\partial x}\right) + \frac{\partial}{\partial z}\left(\mu_v \frac{\partial w}{\partial z}\right), \tag{4}$$

$$x \in (0,L), \qquad z \in (H(x),h(x,t)), \qquad t \in (0,T),$$

where $h(x,t)$ is the elevation function describing a free surface, $H(x)$ is the channel-bed shape function, L is a position of the right boundary, μ_h and μ_v are effective viscosities in x and z directions, respectively, g is the acceleration due to gravity; $p = (p_0 - p_Z)/\rho$ is relative pressure, where p_0 is pressure, p_Z-atmospheric pressure at the sea level, and $\rho = \text{const}$ is density.

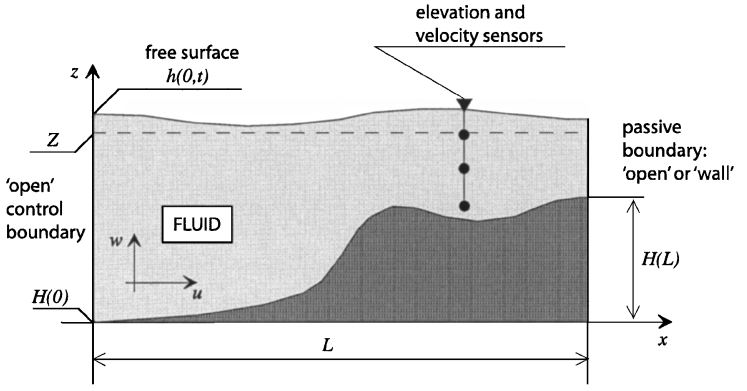


Figure 1. Schematic bathymetry.

The initial state is assumed to be known and we will consider it, for simplicity, as trivial

$$h(x,0) = Z, \qquad u(x,z,0) = 0, \qquad w(x,z,0) = 0, \tag{5}$$

where $Z = \text{const}$ is the elevation of the undisturbed fluid. For the channel bed ($z = H(x)$) in the viscous case, i.e., when $\mu_v + |\frac{\partial H(x)}{\partial x}|\mu_h \neq 0$, we must apply no-slip boundary conditions

$$u = 0, \qquad w = 0. \tag{6}$$

We also consider separately the inviscid flow, in which case we can apply free-slip condition, that is

$$u\frac{\partial H(x)}{\partial x} - w = 0. \tag{7}$$

By neglecting the surface tension and the viscous normal stress, we have the dynamic condition for the free surface ($z = h(x,t)$) as follows:

$$p = 0. \tag{8}$$

When considering boundary conditions for lateral boundaries one should distinguish between physical boundaries, where the known physical behavior of the state variables can be specified, and artificial or 'open' boundaries, introduced for computational purposes. The task of specifying boundary conditions on the open boundaries is not trivial and, in the case of systems which are predominantly hyperbolic, requires a characteristic analysis of the problem to be involved [17]. A fundamental point is that, in order to set up a well-posed open-boundary problem, one must specify the incoming characteristic variables, or in terms of primitive variables, prescribe as many boundary conditions as the number of incoming characteristics. Let us assume that the open control boundary is at $x = 0$ and denote by $U_+(0,z,t)$ the vector-function of incoming characteristic variables. The passive boundary is assumed at $x = L$. It can be either an open boundary or a real physical boundary (liquid/solid interface), i.e., a 'wall'. An open boundary should be assumed to contain only outgoing characteristics. The characteristic structure of lateral boundaries will be considered in detail in Section 7.

Let us denote as $\hat{h}_k := \hat{h}(x_k,t)$ elevation measurements given at some points $x_k \in (0,L)$; and as $\hat{u}_{l,m} := \hat{u}(x_l, Z_m, t)$ u-velocity measurements at some points $x_l \in (0,L)$ along the trajectories $Z_m := Z_m(x_l,t) \in (H(x), h(x,t))$, i.e., velocity sensors are allowed to move in the vertical direction. We formulate a boundary-control problem as follows: find $U_+(0,z,t)$ and $S = [h,p,u,w]^\top$ subjected to constraints (1)–(4) such that

$$
\begin{aligned}
&J\left(U_+(0,z,t)\right) = \inf_{(V_+)} J(V_+(0,z,t)),\\
&J\left(U_+(0,z,t)\right) = \frac{1}{2}\sum_k \int_0^T \left(h(x_k,t) - \hat{h}_k\right)^2 dt + \frac{1}{2}\sum_l \sum_m \int_0^T \left(u(x_l, Z_m(x_l,t)) - \hat{u}_{l,m}\right)^2 dt.
\end{aligned} \tag{9}
$$

The objective functional (9) can be modified to include covariances and penalty terms (regularization).

One of the key difficulties with the problem of interest consists of deriving the variation in the surface variable h. A classical approach is that the variation in a certain dependent variable is calculated assuming all the other dependent variables are functions of x, z, t, i.e., independent variables. For example, we can fix h, p, w and then find the variation in u using the common rules for partial differentiation. The difficulty arises when we calculate variation in h. To do so we have to fix all the other variables p, u, w. That cannot be done unless we fix h itself. Thus, we face a contradiction: in order to calculate the *variation* in $h(x,t)$, it has to be *fixed*. The situation can be resolved using either an explicit or implicit parameterization of the domain boundary [18], or using the concept of topological derivative [19], or using a mapping approach [20]. All these methods are widely used in the shape-optimization problems. In applications to the free-surface fluid flow problems we can mention [21], where the first approach (parameterization) is applied for solving steady potential flow. We use the approach based on a domain transformation. It is worth mentioning that compared to [20] we are dealing here with an unsteady problem, and we have an additional constraint to the mapping coefficients given by the kinematic condition (1). A similar difficulty would arise if $U_+(0,z,t)$ was actually dependent on $z \in (H(0), h(0,t))$. The value of the control variable, which is updated in the course of an iterative process $U_+^{i+1} = U_+^i + \delta U_+^i$ (i is the iteration index), generates a new control domain $z \in (H(0), h^{i+1}(0,t))$, as long as U_+^{i+1} remains defined in the old domain $y \in (H(0), h^i(0,t))$. We will see later that this difficulty can also be resolved using the domain-transformation approach.

We should also mention the two assumptions about the surface that are made in the given formulation. The first one is that the free surface can be defined by the graph of the elevation function $h(x,t)$, which must be a single-valued function of x. Thus, the chosen model cannot describe such phenomena as breaking waves, for example. Second, we use here a simplified form of the dynamic surface boundary condition, neglecting surface tension and the viscous normal stress (the general form could be considered, if necessary).

3. COORDINATE TRANSFORMATION

Let us denote $\mathbb{R}[x,z,t]$ the original coordinate system and introduce a new coordinate system $\mathbb{R}'[x',z',t']$ related to $\mathbb{R}$ by the transformation $\mathbb{Q}:\mathbb{R}\to\mathbb{R}'$ as follows:

$$t'=t,\qquad x'=x,\qquad z'=\frac{(z-H(x))}{(h(x,t)-H(x))},\tag{10}$$

assuming $h(x,t)>H(x)$, $\forall(x,t)$. To simplify presentation we will use notations $h:=h(x,t)$, $H:=H(x)$ until a special need arises. It can be seen that $\det(\mathbb{Q})=1/(h-H)>0$, i.e., the transformation is unique, and the inverse transformation $\mathbb{Q}':\mathbb{R}'\to\mathbb{R}$ exists. Now we can write for derivatives

$$\begin{aligned}
\frac{\partial\cdot}{\partial t}&=\frac{\partial\cdot}{\partial t'}-\frac{z'}{h-H}\frac{\partial h}{\partial t}\frac{\partial\cdot}{\partial z'},\\
\frac{\partial\cdot}{\partial x}&=\frac{\partial\cdot}{\partial x'}-\frac{1}{h-H}\left(\frac{\partial H}{\partial x}+z'\frac{\partial(h-H)}{\partial x}\right)\frac{\partial\cdot}{\partial z'},\\
\frac{\partial\cdot}{\partial z}&=\frac{1}{(h-H)}\frac{\partial\cdot}{\partial z'}.
\end{aligned}$$

It follows from these formulas that

$$\frac{\partial h}{\partial t}=\frac{\partial h}{\partial t'};\qquad \frac{\partial H}{\partial x}=\frac{\partial H}{\partial x'};\qquad \frac{\partial(h-H)}{\partial x}=\frac{\partial(h-H)}{\partial x'}.$$

Therefore, we can finally rewrite the expressions for derivatives as follows:

$$\frac{\partial\cdot}{\partial t}=\frac{\partial\cdot}{\partial t'}-\frac{e_1'}{h-H}\frac{\partial\cdot}{\partial z'},\qquad \frac{\partial\cdot}{\partial x}=\frac{\partial\cdot}{\partial x'}-\frac{e_2'}{(h-H)}\frac{\partial\cdot}{\partial z'},\qquad \frac{\partial\cdot}{\partial z}=\frac{1}{(h-H)}\frac{\partial\cdot}{\partial z'},\tag{11}$$

where

$$e_1'=z'\frac{\partial h}{\partial t'},\qquad e_2'=\frac{\partial H}{\partial x'}+z'\frac{\partial(h-H)}{\partial x'}.\tag{12}$$

Then we can write for $\mathbb{Q}'$

$$\frac{\partial\cdot}{\partial t'}=\frac{\partial\cdot}{\partial t}+e_1\frac{\partial\cdot}{\partial z},\qquad \frac{\partial\cdot}{\partial x'}=\frac{\partial\cdot}{\partial x}+e_2\frac{\partial\cdot}{\partial z},\qquad \frac{\partial\cdot}{\partial z'}=(h-H)\frac{\partial\cdot}{\partial z},\tag{13}$$

where

$$e_1=\frac{z-H}{h-H}\frac{\partial h}{\partial t},\qquad e_2=\frac{\partial H}{\partial x}+\frac{z-H}{h-H}\frac{\partial(h-H)}{\partial x}.\tag{14}$$

For deriving adjoint equations we have to define the scalar product for the domain with variable bounds. Let us denote

$$\langle\cdot,\cdot\rangle_R=\int_0^T\int_0^L\int_H^h(\cdot\,\cdot)\,dz\,dx\,dt,\qquad \langle\cdot,\cdot\rangle_{R'}=\int_0^T\int_0^L\int_0^1(\cdot\,\cdot)(h-H)\,dz'\,dx'\,dt'.$$

We consider now two arbitrary functions $\tilde{s}$, $s^*\in L_2$, and calculate

$$\left\langle s^*,\frac{\partial\tilde{s}}{\partial t}\right\rangle_R=\left\langle s'^*,\left(\frac{\partial\tilde{s}'}{\partial t'}-\frac{e_1'}{h-H}\frac{\partial\tilde{s}'}{\partial z'}\right)\right\rangle_{R'}.$$

Above, $L_2=L_2((0,T)\times(0,L)\times(H,h))$ is a space of square-integrable functions, and $\tilde{s}'=\tilde{s}'(x',z',t')$, $s'^*=s'^*(x',z',t')$. The right-hand side of the last expression is given in constant bounds and we evaluate it using the classical integration-by-parts rule. Next applying to the result the inverse transformation formulas (13),(14) we obtain

$$\left\langle s^*,\frac{\partial\tilde{s}}{\partial t}\right\rangle_R=\left\langle-\tilde{s},\frac{\partial s^*}{\partial t}\right\rangle_R+\int_H^h\int_0^L[\tilde{s}s^*]_0^\top\,dx\,dz-\int_0^T\int_0^L[e_1\tilde{s}s^*]_H^h\,dx\,dt.\tag{15}$$

As expected, the main part is the same as if the boundaries were constant, but the boundary set (i.e., those additional terms which arise in integration by parts) contains a term dependent on the coefficient e_1, which represents the domain-transformation rate. This term affects the adjoint boundary conditions and may appear in the sensitivity formulas. Using a similar approach we derive

$$\left\langle s^*, \frac{\partial \tilde{s}}{\partial x} \right\rangle_R = \left\langle -\tilde{s}, \frac{\partial s^*}{\partial x} \right\rangle_R + \int_0^T \int_H^h [\tilde{s} s^*]_0^L \, dz \, dt - \int_0^T \int_0^L [e_2 \tilde{s} s^*]_H^h \, dx \, dt, \tag{16}$$

$$\left\langle s^*, \frac{\partial \tilde{s}}{\partial z} \right\rangle_R = \left\langle -\tilde{s}, \frac{\partial s^*}{\partial z} \right\rangle_R + \int_0^T \int_0^L [\tilde{s} s^*]_H^h \, dx \, dt. \tag{17}$$

4. TRANSFORMED PROBLEM

Let us reformulate the problem in $\mathbb{R}'$ using the transformation $\mathbb{Q}$. All the variables in $\mathbb{R}'$ are marked by $'$. We start our investigation looking at the inviscid case. Some questions related to the presence of viscous terms will be discussed in Appendix A. For further consideration we need a generalized function $h(x,z,t) = h(x,t)$, $\forall z$. In this case equation (1) can be written in the form

$$\int_H^h \mathbb{E}_1 \, dz := \int_H^h 2 \left(\frac{\partial h}{\partial t} + u \frac{\partial h}{\partial x} - w \right) \delta(z-h) \, dz = 0. \tag{18}$$

The objective functional (9) can also be presented including the δ-function as follows:

$$\begin{aligned} J(U_+(0,z,t)) &= \frac{1}{2} \sum_k \int_0^T \int_0^L \left(h - \hat{h}_k \right)^2 \delta(x - x_k) \, dx \, dt \\ &+ \frac{1}{2} \sum_l \sum_m \int_0^T \int_0^L \int_H^h (u - \hat{u}_{l,m})^2 \, \delta(x - x_l) \, \delta(z - Z_m) \, dz \, dx \, dt. \end{aligned} \tag{19}$$

Applying (11),(12) to (18), (2)–(4), one gets

$$\int_0^1 \mathbb{E}_1' \, dz' := \int_0^1 2 \left(\frac{\partial h}{\partial t'} + u' \frac{\partial h}{\partial x'} - w' \right) \delta(z'-1) \, dz' = 0, \tag{20}$$

$$\mathbb{E}_2' := \frac{\partial u'}{\partial x'} - \frac{e_2'}{h-H} \frac{\partial u'}{\partial z'} + \frac{1}{h-H} \frac{\partial w'}{\partial z'} = 0, \tag{21}$$

$$\mathbb{E}_3' := \frac{\partial u'}{\partial t'} + u' \frac{\partial u'}{\partial x'} + \frac{w' - e_1' - u' e_2'}{h-H} \frac{\partial u'}{\partial z'} + \frac{\partial p'}{\partial x'} - \frac{e_2'}{h-H} \frac{\partial p'}{\partial z'} = 0, \tag{22}$$

$$\mathbb{E}_4' := \frac{\partial w'}{\partial t'} + u' \frac{\partial w'}{\partial x'} + \frac{w' - e_1' - u' e_2'}{h-H} \frac{\partial w'}{\partial z'} + \frac{1}{h-H} \frac{\partial p'}{\partial z'} + g = 0, \tag{23}$$

$$x' \in (0, L), \qquad z' \in (0,1), \qquad t' \in (0,T),$$

where $p'(x',z',t')$, $u'(x',z',t')$, and $w'(x',z',t')$ are new variables defined in the rectangular spatial domain $[0,L] \times [0,1]$ related to p, u, and w by mapping

$$p'(x',z',t') = p(x', H + z'(h-H), t'), \qquad \text{etc.} \tag{24}$$

Since h is the same in $\mathbb{R}$ and $\mathbb{R}'$, we do not use the notation h'. The initial and boundary conditions (inviscid case) are

$$\begin{gathered} h(x',0) = Z, \qquad u'(x',z',0) = 0, \qquad w'(x',z',0) = 0, \\ u'(x',0,t') \frac{\partial H(x')}{\partial x'} - w'(x',0,t') = 0, \\ p'(x',1,t') = 0, \\ x = 0 : U_+'(0,z',t') - \text{control input}, \qquad x = L : \text{'open' boundary or 'wall'}. \end{gathered}$$

We can see that the functions p', u', w' are defined for all $z' \in (0,1)$, i.e., h is no longer a part of the definitions. Therefore, we apply the classical approach to evaluate the variation in h, which is now explicitly incorporated into the model equations written in the transformed space. Also, in $\mathbb{R}'$ the control vector-function $U'_+(0,z',t')$ can be updated because it is always defined in the same domain $z' \in (0,1)$. The objective functional (19) in $\mathbb{R}'$ takes the form

$$\begin{aligned} J'(U'_+(0,z',t')) &= \frac{1}{2}\sum_k \int_0^T \int_0^L \left(h - \hat{h}_k\right)^2 \delta\left(x' - x'_k\right) dx'\, dt' \\ &+ \frac{1}{2}\sum_l \sum_m \int_0^T \int_0^L \int_0^1 \left(u' - \hat{u}_{l,m}\right)^2 \delta\left(x' - x'_l\right) \delta\left(z' - Z'_m\right) dz'\, dx'\, dt'. \end{aligned} \tag{25}$$

Apparently $J' = J$, and therefore, both J' and J have the same stationary points and we can solve the control problem in $\mathbb{R}'$.

5. TANGENT LINEAR MODEL

The forward model, TLM, and the adjoint model could all be formulated and resolved in $\mathbb{R}'$. This is often referred to as a problem stated in σ-coordinates [3]. In fact, in meteorology, continuous terrain-following coordinate transformations are generally favored. Recently, the classical transformation of Gal-Chen and Sommerville has been extended to a broad class of time-dependent vertical domains, which incorporate free-surface upper boundaries [22]. In practice, however, many solvers built for oceanographic applications work in the original space $\mathbb{R}$. In order to cope with existing solvers the TLM derived in $\mathbb{R}'$ and the related adjoint problem can be reformulated back to $\mathbb{R}$.

Let $Q(s)$ be the result of a mapping of an object $s \in \mathbb{R}$ into the transformed space $\mathbb{R}'$. For example, the object $s := \frac{\partial \cdot}{\partial t}$ is defined in $\mathbb{R}'$ by the right-hand side of the first formula in (11), i.e., we can formally write

$$Q\left(\frac{\partial \cdot}{\partial t}\right) := \frac{\partial \cdot}{\partial t'} - \frac{e'_1}{h - H}\frac{\partial \cdot}{\partial z'}.$$

Because $Q(s) \in \mathbb{R}'$, it can be used in those expressions which are written in $\mathbb{R}'$, showing the intention to perform the mapping of s at some stage. In the same way we denote $Q'(s')$ the result of the inverse mapping of an object $s' \in \mathbb{R}'$ into the original space $\mathbb{R}$. These notations are introduced to minimize the size of the formulas, i.e., for descriptive purposes only.

Let us define $S = [h, S_\perp]^\top$, $S_\perp = [p, u, w]^\top$, a vector-function of state variables in $\mathbb{R}$, and $S' = [h, S'_\perp]^\top$, $S'_\perp = [p', u', w']^\top$, in the transformed space $\mathbb{R}'$ correspondingly. We will mark the variation in a certain object s as $\tilde{s}$, and the corresponding adjoint object as s^*. We assume that the TLM in $\mathbb{R}'$ can be presented in the form

$$\tilde{\mathbb{E}}' \cdot \tilde{S}' = 0.$$

Here $\tilde{\mathbb{E}}' = [\tilde{\mathbb{E}}'_1, \tilde{\mathbb{E}}'_\perp]^\top$, $\tilde{\mathbb{E}}'_\perp = [\tilde{\mathbb{E}}'_2, \tilde{\mathbb{E}}'_3, \tilde{\mathbb{E}}'_4]^\top$, is a linear matrix-operator acting on $\tilde{S}'$, such that

$$\tilde{\mathbb{E}}'_i = \left(\frac{\partial \mathbb{E}'_i}{\partial h}, \frac{\partial \mathbb{E}'_i}{\partial p'}, \frac{\partial \mathbb{E}'_i}{\partial u'}, \frac{\partial \mathbb{E}'_i}{\partial w'}\right); \qquad i = 1, \ldots, 4.$$

Then the adjoint matrix-operator $\mathbb{E}'^*$ in $\mathbb{R}'$ can be calculated from the following scalar product:

$$\left\langle (S'^*)^\top, \tilde{\mathbb{E}}' \cdot \tilde{S}' \right\rangle_{R'} = \left\langle \left(\tilde{S}'\right)^\top, \mathbb{E}'^* \cdot S'^* \right\rangle_{R'} + B', \tag{26}$$

where B' stands for the boundary set, which arises in integration by parts. We intend to use $Q'(\mathbb{E}'^*)$ as an adjoint matrix-operator in the original space.

Although we could proceed in that way, a more elegant (less laborious) approach can be suggested instead. Let $\mathbb{E}'^{*}_{\perp}$ be a submatrix of $\mathbb{E}'^{*}$ such that it consists of the columns of $\mathbb{E}'^{*}_{i}$, $i = 2,3,4$, and $B'_{\perp}$ be a subset of B' without those terms containing h^*. Let us assume now that $\mathbb{E}^{*}_{\perp}$ is the adjoint matrix-operator and $B_{\perp}$ are the boundary terms derived from a scalar product

$$\left\langle (S^{*}_{\perp})^{\top}, Q'\left(\tilde{\mathbb{E}}'_{\perp}\right)\cdot\tilde{S}\right\rangle_{R} = \left\langle \left(\tilde{S}\right)^{\top}, \mathbb{E}^{*}_{\perp}\cdot S^{*}_{\perp}\right\rangle_{R} + B_{\perp}. \tag{27}$$

It can be proved that if we use (15)–(17) to calculate (27), then

$$Q'\left(\mathbb{E}'^{*}_{\perp}\right) = E^{*}_{\perp}, \qquad Q'(B'_{\perp}) = B_{\perp}.$$

It can also be proved that

$$Q'\left(\tilde{\mathbb{E}}'_{i}\right) = \left(Q'\left(\frac{\partial \mathbb{E}'_{i}}{\partial h}\right), \frac{\partial \mathbb{E}_{i}}{\partial p}, \frac{\partial \mathbb{E}_{i}}{\partial u}, \frac{\partial \mathbb{E}_{i}}{\partial w}\right); \qquad i = 2,3,4. \tag{28}$$

Looking at (27) and (28) we may conclude that all the adjoint terms related to (2)–(4) could be calculated as follows: the variations $\tilde{p}$, $\tilde{u}$, $\tilde{w}$ are calculated in $\mathbb{R}$ from (2)–(4) and $\tilde{h}$ is calculated in $\mathbb{R}'$ from (21)–(23) and then is mapped back to $\mathbb{R}$; then the adjoint terms are finally calculated directly in $\mathbb{R}$. The adjoint terms related to (1) are calculated as follows: first, the variation in (20) is calculated, then the adjoint terms are calculated in $\mathbb{R}'$ and then are mapped to $\mathbb{R}$. Thus, the variation in (20) is

$$\int_{0}^{1} \tilde{\mathbb{E}}'_{1}\cdot\tilde{S}'\,dz' := \int_{0}^{1} 2\left(\frac{\partial\tilde{h}}{\partial t'} + u'\frac{\partial\tilde{h}}{\partial x'} + \frac{\partial h}{\partial x'}\tilde{u}' - \tilde{w}'\right)\delta(z'-1)\,dz' = 0, \tag{29}$$

and in $\mathbb{R}$ it is

$$\int_{H}^{h} 2\left(\frac{\partial\tilde{h}}{\partial t} + u\frac{\partial\tilde{h}}{\partial x} + \frac{\partial h}{\partial x}\tilde{u} - \tilde{w}\right)\delta(z-h)\,dz = 0. \tag{30}$$

As we mentioned, the last equation cannot be used for producing the adjoint terms and it is presented here just to complete the TLM formulation.

Following (28) we calculate first

$$\frac{\partial \mathbb{E}_{i}}{\partial p}, \frac{\partial \mathbb{E}_{i}}{\partial u}, \frac{\partial \mathbb{E}_{i}}{\partial w}; \qquad i = 2,3,4.$$

Thus we obtain the TLM equations as follows:

$$Q'\left(\tilde{\mathbb{E}}'_{2}\right)\cdot\tilde{S} := \frac{\partial\tilde{u}}{\partial x} + \frac{\partial\tilde{w}}{\partial z} + Q'\left(\frac{\partial\mathbb{E}'_{2}}{\partial h}\right)\cdot\tilde{h} = 0, \tag{31}$$

$$Q'\left(\tilde{\mathbb{E}}'_{3}\right)\cdot\tilde{S} := \frac{\partial\tilde{u}}{\partial t} + u\frac{\partial\tilde{u}}{\partial x} + w\frac{\partial\tilde{u}}{\partial z} + \frac{\partial\tilde{p}}{\partial x} + \frac{\partial u}{\partial x}\tilde{u} + \frac{\partial u}{\partial z}\tilde{w} + Q'\left(\frac{\partial\mathbb{E}'_{3}}{\partial h}\right)\cdot\tilde{h} = 0, \tag{32}$$

$$Q'\left(\tilde{\mathbb{E}}'_{4}\right)\cdot\tilde{S} := \frac{\partial\tilde{w}}{\partial t} + u\frac{\partial\tilde{w}}{\partial x} + w\frac{\partial\tilde{w}}{\partial z} + \frac{\partial\tilde{p}}{\partial z} + \frac{\partial w}{\partial x}\tilde{u} + \frac{\partial w}{\partial z}\tilde{w} + Q'\left(\frac{\partial\mathbb{E}'_{4}}{\partial h}\right)\cdot\tilde{h} = 0, \tag{33}$$

$$x\in(0,L), \qquad z\in(H,h), \qquad t\in(0,T).$$

Then we calculate the last terms in (31)–(33) as follows:

$$Q'\left(\frac{\partial\mathbb{E}'_{2}}{\partial h}\right) := (h-H)\left(-\frac{\partial u}{\partial z}Q'\left(\frac{d}{dh}\left(\frac{e'_{2}}{h-H}\right)\right) + \frac{\partial w}{\partial z}Q'\left(\frac{d}{dh}\left(\frac{1}{h-H}\right)\right)\right),$$

$$Q'\left(\frac{\partial\mathbb{E}'_{3}}{\partial h}\right) := (h-H)\left(\frac{\partial u}{\partial z}Q'\left(\frac{d}{dh}\left(\frac{w'-e'_{1}-u'e'_{2}}{h-H}\right)\right) - \frac{\partial p}{\partial z}Q'\left(\frac{d}{dh}\left(\frac{e'_{2}}{h-H}\right)\right)\right),$$

$$Q'\left(\frac{\partial\mathbb{E}'_{4}}{\partial h}\right) := (h-H)\left(\frac{\partial w}{\partial z}Q'\left(\frac{d}{dh}\left(\frac{w'-e'_{1}-u'e'_{2}}{h-H}\right)\right) + \frac{\partial p}{\partial z}Q'\left(\frac{d}{dh}\left(\frac{1}{h-H}\right)\right)\right).$$

Taking into account that

$$\frac{d}{dh}\left(\frac{1}{h-H}\right) = -\frac{1}{(h-H)^2},$$
$$\frac{d}{dh}\left(\frac{e_2'}{h-H}\right) = \frac{z'}{h-H}\frac{\partial\cdot}{\partial x'} - \frac{e_2'}{(h-H)^2},$$
$$\frac{d}{dh}\left(\frac{w'-e_1'-u'e_2'}{h-H}\right) = -\frac{z'}{h-H}\left(\frac{\partial\cdot}{\partial t'} + u'\frac{\partial\cdot}{\partial x'}\right) - \frac{w'-e_1'-u'e_2'}{(h-H)^2},$$

using the inverse transformation formulas (13),(14) and then collecting terms we have got eventually

$$Q'\left(\frac{\partial \mathbb{E}_i'}{\partial h}\right)\cdot\tilde{h} := a_{i,1}\frac{\partial\tilde{h}}{\partial t} + a_{i,2}\frac{\partial\tilde{h}}{\partial x} + a_{i,3}\frac{\partial\tilde{h}}{\partial z} + a_{i,4}\tilde{h}, \qquad i=2,3,4, \tag{34}$$

where

$$\begin{aligned}
a_{2,1} &= 0, \quad a_{2,2} = -\frac{z-H}{h-H}\frac{\partial u}{\partial z}, \quad a_{2,4} = -\frac{1}{h-H}\left(\frac{\partial w}{\partial z} - e_2\frac{\partial u}{\partial z}\right),\\
a_{3,1} &= -\frac{z-H}{h-H}\frac{\partial u}{\partial z}, \quad a_{3,2} = -\frac{z-H}{h-H}\left(u\frac{\partial u}{\partial z} + \frac{\partial p}{\partial z}\right),\\
a_{3,4} &= -\frac{1}{z-H}\left((w-e_1-ue_2)\frac{\partial u}{\partial z} - e_2\frac{\partial p}{\partial z}\right),\\
a_{4,1} &= -\frac{z-H}{h-H}\frac{\partial w}{\partial z}, \quad a_{4,2} = -\frac{z-H}{h-H}u\frac{\partial w}{\partial z},\\
a_{4,4} &= -\frac{1}{h-H}\left((w-e_1-ue_2)\frac{\partial w}{\partial z} + \frac{\partial p}{\partial z}\right),\\
a_{i,3} &= a_{i,1}e_1 + a_{i,2}e_2, \qquad i=2,3,4.
\end{aligned} \tag{35}$$

The initial and boundary conditions for the TLM in $\mathbb{R}$ (inviscid case) are as follows:

$$\begin{gathered}
\tilde{h}(x,0) = 0, \qquad \tilde{u}(x,z,0) = 0, \qquad \tilde{w}(x,z,0) = 0,\\
\tilde{u}(x,H,t) + \frac{\partial H(x)}{\partial x}\tilde{w}(x,H,t) = 0,\\
\tilde{p}(x,h,t) = 0,\\
x=0 : \tilde{U}_+(0,z,t) - \text{control input}; \qquad x=L : \text{'open' boundary or 'wall'}.
\end{gathered}$$

In order to derive adjoint equations in $\mathbb{R}$ we also need to calculate $\frac{\partial J'}{\partial u'}\cdot\tilde{u}'$ and $\frac{\partial J'}{\partial h}\cdot\tilde{h}$ and map the formulas to $\mathbb{R}$. It can be seen that

$$Q'\left(\frac{\partial J'}{\partial u'}\cdot\tilde{u}'\right) := \frac{\partial J}{\partial u}\cdot\tilde{u} = \sum_l\sum_m\int_0^T\int_0^L\int_H^h (u-\hat{u}_{l,m})\,\delta(x-x_l)\,\delta(z-Z_m)\tilde{u}\,dz\,dx\,dt. \tag{36}$$

For $\frac{\partial J'}{\partial h}\cdot\tilde{h}$ we have

$$\begin{aligned}
\frac{\partial J'}{\partial h}\cdot\tilde{h} &= \sum_k\int_0^T\int_0^L\left(h-\hat{h}_k\right)\delta(x'-x_k')\tilde{h}\,dx'\,dt'\\
&+\frac{1}{2}\sum_l\sum_m\int_0^T\int_0^L\left\{\int_0^1 (u'-\hat{u}_{l,m})^2\,\frac{\partial\delta(z'-Z_m')}{\partial h}\,dz'\right\}\delta(x'-x_l')\tilde{h}\,dx'\,dt'.
\end{aligned} \tag{37}$$

Then, using the differentiation rule for composite function we obtain

$$\frac{\partial\delta(z'-Z_m')}{\partial h} = -\frac{\partial\delta(z'-Z_m')}{\partial z'}\frac{\partial Z_m'}{\partial h}.$$

As long as $\frac{\partial Z'_m}{\partial h}$ does not depend on z', it can be exported from the integral with respect to z' and will appear later as a multiplier. Using the definition for derivative of δ-function we write as follows:

$$-\int_0^1 (u'-\hat{u}_{l,m})^2 \frac{\partial \delta(z'-Z'_m)}{\partial z'}\,dz' = \left[\frac{\partial}{\partial z'}\left((u'-\hat{u}_{l,m})^2\right)\right]_{z'=Z'_m} = \left[2(u'-\hat{u}_{l,m})\frac{\partial u'}{\partial z'}\right]_{z'=Z'_m}$$

Now we multiply the result by $\frac{\partial Z'_m}{\partial h}$ and substitute it into (37) instead of the expression in braces. Performing the inverse transformation we get eventually

$$\begin{aligned} Q'\left(\frac{\partial J'}{\partial h}\cdot\tilde{h}\right) &:= \sum_k \int_0^T \int_0^L \left(h-\hat{h}_k\right)\tilde{h}\delta(x-x_k)\,dx\,dt \\ +\sum_l \sum_m \int_0^T \int_0^L &\left[(u-\hat{u}_{l,m})\frac{\partial u}{\partial y}\right]_{z=Z_m}\left(\frac{\partial Z_m}{\partial h}-\frac{Z_m-H}{h-H}\right)\delta(x-x_l)\tilde{h}\,dx\,dt. \end{aligned} \tag{38}$$

This is interesting to note that when Z_m is a given function, then $\frac{\partial Z_m}{\partial h}=0$. However, if it is unknown, but depends explicitly on h, for example $Z_m = h$ (i.e., measurements are performed on the surface), then

$$\frac{\partial h}{\partial h}-\frac{h-H}{h-H}=1-1=0$$

and the second term in the right-hand side of (38) vanishes.

6. (PSEUDO)ADJOINT MODEL

Now we shall derive the adjoint equations and the boundary set. Let us construct the total variation in $\mathbb{R}'$ as follows:

$$\Lambda' = \left\langle h^*, \tilde{\mathbb{E}}'_1\cdot\tilde{S}'\right\rangle_{R'} + \left\langle (S'^*_\perp)^\top, \tilde{\mathbb{E}}'_\perp\cdot\tilde{S}'\right\rangle_{R'} + \frac{\partial J'}{\partial u'}\cdot\tilde{u}' + \frac{\partial J'}{\partial h}\cdot\tilde{h}. \tag{39}$$

As stated above, adjoint equations can be derived from Λ' and then mapped into the original space. This procedure is actually applied to the first term in the right-hand side of (39) only. For the second term, we follow a 'shortcut' discussed in Section 5 and use (27) to calculate $E^*_\perp$ directly in $\mathbb{R}$. The last two terms, which are mapped into $\mathbb{R}$ (see (36),(38)), are used to derive the right-hand sides for the corresponding adjoint equations. Thus, the first term in (39) produces the adjoint terms

$$-2\left(\frac{\partial h^*}{\partial t}+\frac{\partial uh^*}{\partial x}+e_2\frac{\partial u}{\partial z}h^*\right)\delta(z-h)\tilde{h} \tag{40}$$

$$+2\frac{\partial h}{\partial x}h^*\delta(z-h)\tilde{u} \tag{41}$$

$$-2h^*\delta(z-h)\tilde{w} \tag{42}$$

and a boundary set as follows:

$$\int_0^L \left[\tilde{h}h^*\right]_0^\top dx \tag{43}$$

$$+\int_0^T \left[\tilde{h}u(x,h,t)h^*\right]_0^L dt. \tag{44}$$

Next we apply formulas (15)–(17) to all the terms in (31)–(33), except the last ones (in $\tilde{S}$). Collecting results around $\tilde{p}$, $\tilde{u}$, $\tilde{w}$ including those in (41),(42) we compose the adjoint equations

$$-\frac{\partial u^*}{\partial x} - \frac{\partial w^*}{\partial z} = 0, \tag{45}$$

$$-\frac{\partial u^*}{\partial t} - u\frac{\partial u^*}{\partial x} - w\frac{\partial u^*}{\partial z} - \frac{\partial p^*}{\partial x} - \frac{\partial w}{\partial z}u^* + \frac{\partial w}{\partial x}w^* + 2\frac{\partial h}{\partial x}h^*\delta(z-h) = r_u, \tag{46}$$

$$-\frac{\partial w^*}{\partial t} - u\frac{\partial w^*}{\partial x} - w\frac{\partial w^*}{\partial z} - \frac{\partial p^*}{\partial z} + \frac{\partial u}{\partial z}u^* - \frac{\partial u}{\partial x}w^* - 2h^*\delta(z-h) = 0, \tag{47}$$

$$x \in (0, L), \qquad z \in (H, h), \qquad t \in (0, T),$$

where

$$r_u = -\sum_l \sum_m (u - \hat{u}_{l,m})\, \delta(x - x_l)\delta(z - Z_m).$$

Finally, we apply (15)–(17) to (34), which represents the last terms in (31)–(33). Collecting results around $\tilde{h}$ including those in (40), we get after integrating all with respect to z the adjoint equation for h^*

$$-\frac{\partial h^*}{\partial t} - \frac{\partial (u(x,h,t)h^*)}{\partial x} - \frac{\partial h}{\partial x}\frac{\partial u(x,h,t)}{\partial z}h^* - \mathbb{F} = r_h, \tag{48}$$
$$x \in (0, L), \qquad t \in (0, T),$$

where

$$\mathbb{F} = \sum_{i=2}^{4} \int_H^h \left(\frac{\partial a_{i,1} s_i^*}{\partial t} + \frac{\partial a_{i,2} s_i^*}{\partial x} + \frac{\partial a_{i,3} s_i^*}{\partial z} - a_{i,4} s_i^* \right) dz, \qquad s_i^* \in S^* = [h^*, p^*, u^*, w^*], \tag{49}$$

$$r_h = -\sum_k \left(h - \hat{h}_k\right) \delta(x - x_k)$$
$$- \sum_l \sum_m \left[(u - \hat{u}_{l,m}) \frac{\partial u}{\partial z} \right]_{z=Z_m} \left(\frac{\partial Z_m}{\partial h} - \frac{Z_m - H}{h - H} \right) \delta(x - x_l). \tag{50}$$

The boundary set, obtained as a result of calculating $\mathbb{E}_\perp^*$, is as follows:

$$B_\perp = \int_H^h \int_0^L \left[\tilde{u}u^* + \tilde{w}w^* + \sum_{i=2}^{4} \tilde{h} a_{i,1} s_i^* \right]_0^\top dx\, dz \tag{51}$$
$$+ \int_0^T \int_0^L \left[-e_2 \tilde{u} p^* + \tilde{w} p^* - e_2 \tilde{p} u^* + \tilde{p} w^* \right]_H^h dx\, dt$$
$$+ \int_0^T \int_0^L \left[-e_1 \tilde{u} u^* - e_2 \tilde{u} u u^* + \tilde{u} w u^* - e_1 \tilde{w} w^* - e_2 \tilde{w} u w^* + \tilde{w} w w^* \right]_H^h dx\, dt \tag{52}$$
$$+ \int_0^T \int_H^h \left[\tilde{u} p^* + \tilde{u} u u^* + \tilde{p} u^* + \tilde{w} u w^* + \sum_{i=2}^{4} \tilde{h} a_{i,2} s_i^* \right]_0^L dz\, dt. \tag{53}$$

Using Leibnitz's integration rule and taking into account (14) and (35), we can rewrite (49) in the equivalent form, that is convenient for numerical implementation

$$\mathbb{F} = \sum_{i=2}^{4} \left(\frac{\partial}{\partial t} \int_H^h a_{i,1} s_i^*\, dz + \frac{\partial}{\partial x} \int_H^h a_{i,2} s_i^*\, dz - \int_H^h a_{i,4} s_i^*\, dz \right). \tag{54}$$

The order of magnitude analysis made for coefficients $a_{i,j}$ in (34) shows that for weak dynamics typical for many problems in oceanography only those terms which contain $\frac{\partial p}{\partial z} \approx -g$ and p^* are significant.

We must emphasize here that equations (48), (45)–(47) represent in $\mathbb{R}$ the adjoint equations related to the transformed equations (20)–(23), rather than the adjoint model directly related to the original equations (1)–(4). The control problem we actually solve is the problem that involves the objective functional (25). Thus, the adjoint solution obtained in $\mathbb{R}$ must deliver a correct sensitivity in $\mathbb{R}'$. Let us note that the optimality system in $\mathbb{R}'$ may look different dependent on the transformation applied. The same is true for the corresponding 'adjoint' model in $\mathbb{R}$ (coefficients $a_{i,j}$). However, an optimal solution calculated in a chosen $\mathbb{R}'$ and mapped in $\mathbb{R}$ must coincide with the optimal solution of the original problem in $\mathbb{R}$, no matter what the transformation was used. Taking into account these considerations, the adjoint formulation presented here is referred to as a pseudo-adjoint to (1)–(4).

We shall derive the initial and boundary conditions for the adjoint problem. As usual, assuming zero terminal conditions for all adjoint variables, i.e.,

$$h^*(x,T)=0, \qquad p^*(x,z,T)=0, \qquad u^*(x,z,T)=0, \qquad w^*(x,z,T)=0, \tag{55}$$

we annul (43) and (51). Boundary terms related to the surface and the bed (52) can be rewritten in a more comprehensive form as follows:

$$\int_0^T \int_0^L \left[-p^*(\tilde{u}e_2-\tilde{w})-\tilde{p}(u^*e_2-w^*)-(\tilde{u}u^*+\tilde{w}w^*)(e_1+ue_2-w)\right]_H^h \, dx\,dt.$$

Looking at (14) and keeping in mind bed boundary conditions both for the forward model and for the TLM we note that

$$\begin{aligned}
[e_1+ue_2-w]_h &= \frac{\partial h}{\partial t}+u(x,h,t)\frac{\partial h}{\partial x}-w(x,h,t)\equiv 0,\\
[e_1+ue_2-w]_H &= u(x,H,t)\frac{\partial H}{\partial x}-w(x,H,t)\equiv 0,\\
[\tilde{u}e_2-\tilde{w}]_H &= \tilde{u}(x,H,t)\frac{\partial H}{\partial x}-\tilde{w}(x,H,t)\equiv 0.
\end{aligned}$$

Taking into account that $\tilde{p}(x,h,t)=0$ we annul (52) completely by choosing

$$p^*(x,h,t)=0, \tag{56}$$

$$u^*(x,H,t)\frac{\partial H}{\partial x}-w^*(x,H,t)=0. \tag{57}$$

The terms in (44),(53) could be used for deriving boundary conditions and gradients at the lateral boundaries. However, we use another approach, which is presented below.

7. BOUNDARY CONDITIONS FOR LATERAL BOUNDARIES

Let us consider a general 1D linear hyperbolic problem given as follows:

$$\frac{\partial U}{\partial t}+A\frac{\partial U}{\partial x}+BU=0, \qquad x\in(0,L), \quad t\in(0,T), \tag{58}$$

with an open boundary at $x=0$. Here $U(x,t)$ is a vector-function of state variables, A and B are coefficient matrices. This is a typical form in which the TLM of an original nonlinear hyperbolic problem can be cast. Let us also assume that there exists a decomposition $A=R\Lambda L$, such that $\Lambda=\mathrm{diag}(\lambda_i)$, $RL=I$, and $LR=I$. (Such a decomposition may exist only in some cases.) Then the previous equation can be presented in the form

$$\frac{\partial V}{\partial t}+\Lambda\frac{\partial V}{\partial x}+\left(LB-\frac{\partial L}{\partial t}-\Lambda\frac{\partial L}{\partial x}\right)RV=0, \tag{59}$$

where $V = LU$ are the characteristic variables. Obviously, one also has $U = RV$. Depending on the sign of eigenvalues λ_i we deal with either the incoming or the outgoing characteristics. A fundamental point is that, in order to set up a well-posed open-boundary problem, one must specify the incoming characteristic variables.

Let us write now the adjoint of (58) as follows:

$$-\frac{\partial U^*}{\partial t} - A^\top \frac{\partial U^*}{\partial x} + \left(B^\top - \frac{\partial A^\top}{\partial x}\right) U^* = 0. \tag{60}$$

In addition to (60) one must consider a boundary term

$$\int_0^T \left[U^\top A^\top U^*\right]_0^L \, dt, \tag{61}$$

which is used to derive boundary conditions for the adjoint problem. These conditions have to be chosen in such a way that (61) would eventually depend on those variables which we intend to control. Now we remember that, for the open-boundary problem to be well posed, only the incoming characteristic variables have to be specified. One can note that if $\lambda_i < 0$ for some indexes i, then $U(0,t)$ may not be considered as a boundary condition, because the problem becomes overdetermined and therefore ill-posed. In order to build the appropriate control we must turn to the characteristic form of (60),(61). In terms of the factors R, L, and Λ the adjoint equation (60) can be written in the form

$$-\frac{\partial V^*}{\partial t} - \Lambda \frac{\partial V^*}{\partial x} + \left(R^\top \left(LB - \frac{\partial L}{\partial t} - \Lambda \frac{\partial L}{\partial x}\right)^\top - \frac{\partial \Lambda}{\partial x}\right) V^* = 0, \tag{62}$$

where $V^* = R^\top U^*$ shall be called the adjoint characteristic variables. As above, one can write $U^* = L^\top V^*$. Comparing the transport part of (62) and (59) one can note that the eigenvalues of the two problems differ only by sign. It means that the incoming characteristics of the forward problem (59) are the outgoing characteristics for the adjoint problem (62), and vice versa. Now the kernel of (61) can be written as follows:

$$U^\top A^\top U^* = U^\top L^\top \Lambda R^\top U^* = (LU)^\top \Lambda \left(R^\top U^*\right) = V^\top \Lambda V^* = \sum_i V_i \lambda_i V_i^*. \tag{63}$$

In order to make (61) independent of the outgoing forward variables, we must put to zero the corresponding incoming adjoint characteristic variables. Thus, from (61),(63) we derive the boundary conditions for the adjoint problem

$$V_i^* = 0, \qquad \forall \lambda_i < 0, \quad x = 0. \tag{64}$$

It is interesting to note that when specifying zero incoming characteristic variables one obtains a radiation boundary, rather than a perfect nonreflecting boundary in terms of Hedstrom's definition [23], which is

$$\lambda_i R_i^\top \frac{\partial U^*}{\partial x} := \lambda_i \left(\frac{\partial V_i^*}{\partial x} - \frac{\partial R_i^\top}{\partial x} L_i^\top V_i^*\right) = 0. \tag{65}$$

This is obviously not equivalent to (64).

Now, we should recognize that the incompressible (inelastic) fsNSE (1)–(4) are not a hyperbolic system. However, their eigensolutions have time-dependent behavior which is like that of hyperbolic systems. That is, the barotropic behavior is dominated by free-surface gravity waves. We will use this property to construct an approximate characteristic representation of open boundaries. The well-posedness of problem (1)–(4) for periodic boundary conditions is justified

in [25], and for the inviscid case in [24]. However, we have found no results that demonstrate the well-posedness of the open-boundary problem for (1)–(4). The nearest results are those related to the primitive hydrostatic equations (PHE) [26–28]. A rigorous characteristic analysis of the compressible fsNSE (which do constitute a hyperbolic system) is probably possible to conduct. A proper elaboration of this issue may constitute a subject for a future research.

Let us note that this discussion is related to the boundaries only. We can assume that at the boundaries we have physical conditions that could justify a simplified approach, but at the same time inside the domain we have a full-physics model. Let us suppose that in the vicinity of the control boundary the 2D fsNSE support a surface gravity wave that could be described by the SWE in terms of variables ϕ, q, which are uniquely defined through the variables of the 2D fsNSE problem as follows:

$$\phi = h - Z, \qquad q = \int_H^h u\, dz.$$

However, the variables of the original problem cannot be uniquely defined through the SWE variables, unless we associate with q a certain velocity profile $\breve{u}(z)$, such that

$$\int_H^h \breve{u}(z)\, dz = q. \tag{66}$$

A simple choice is

$$\breve{u}(z) = \bar{u}, \tag{67}$$

but one could use functions which actually depend on z. For example, it could be the exponential function to represent the profile associated with a short wave. However, we will confine ourselves to considering long waves. In this case (67) is valid. If equation (58) represents the 1D SWE, then

$$A = \begin{pmatrix} 0 & 1 \\ c^2 - \bar{u}^2 & 2\bar{u} \end{pmatrix}, \qquad U = \begin{pmatrix} \phi \\ q \end{pmatrix},$$

where

$$\bar{u} = \frac{q}{(h-H)}, \qquad c = \sqrt{g(h-H)}, \tag{68}$$

and A can be factorized in the form $A = R\Lambda L$, where

$$R = \frac{1}{2c}\begin{pmatrix} 1 & -1 \\ c+\bar{u} & c-\bar{u} \end{pmatrix}, \qquad \Lambda = \begin{pmatrix} c+\bar{u} & 0 \\ 0 & -c+\bar{u} \end{pmatrix}, \qquad L = \begin{pmatrix} c-\bar{u} & 1 \\ -(c+\bar{u}) & 1 \end{pmatrix}. \tag{69}$$

Now, assuming that $\breve{u}$ is transported along the characteristics related to the eigenvalues $\lambda_1 = \bar{u}+c$ and $\lambda_2 = \bar{u} - c$, we define the following characteristic variables:

$$V_1 = q + (c - \bar{u})\phi, \qquad V_2 = q - (c + \bar{u})\phi.$$

Thus, we have two characteristic variables at the open boundary: incoming V_1 and outgoing V_2. The control variable is therefore the incoming characteristic variable V_1, i.e., $U_+(0,t) = V_1(t)$. As long as it does not depend on z, the control can be exercised directly in $\mathbb{R}$. The value of the outgoing characteristic variable V_2 has to be calculated from its values in the interior. Since we are unable to derive exact equations for characteristic variables at the boundary, we will evaluate the values of V_2 simply by extrapolating from interior values, using the second-order scheme

$$V_2(0,t) = 2V_2(\Delta x, t) - V_2(2\Delta x, t).$$

One can note that the relationship between V_1, V_2 and primitive variables h, u of the fsNSE is nonlinear, unless $\bar{u}$, c in (68), and h in (66) are considered as coefficients which are dependent on x, t only. Let us assume that $\bar{u}$, c, and h are defined (for example, in the semi-implicit

numerical implementation these values are taken from the previous time step). Then h, u can be linearly expressed via V_1, V_2 (using R from (69)) as follows:

$$\begin{aligned} h(0,t) &= Z + \frac{1}{2}\left(\frac{V_1}{c} - \frac{V_2}{c}\right), \\ u(0,z,t) &= \frac{1}{2(h-H)}\left[\left(1+\frac{\bar{u}}{c}\right)V_1 + \left(1-\frac{\bar{u}}{c}\right)V_2\right]. \end{aligned} \tag{70}$$

The boundary values for w can be calculated using the irrotationality condition that yields

$$\frac{\partial w(0,z,t)}{\partial x} = 0. \tag{71}$$

This condition is used for control and passive boundaries in both forward and adjoint models. As usual for the incompressible NSE, no boundary condition for pressure has to be specified at lateral boundaries. Similar definitions for the adjoint characteristic variables are

$$\phi^* = h^*(h-H), \qquad q^* = \int_H^h u^*\,dz, \qquad \breve{u}^*(z) = \frac{q^*}{h-H},$$
$$V_1^* = \frac{1}{2c}\left(\phi^* + (c+\bar{u})q^*\right), \qquad V_2^* = \frac{1}{2c}\left(-\phi^* + (c-\bar{u})q^*\right),$$

and the variables of the fsNSE adjoint problem expressed in terms of V_i^* (using $L^\top$ from (69)) are

$$\begin{aligned} h^*(0,t) &= \frac{(c-\bar{u})}{h-H}V_1^* - \frac{(c+\bar{u})}{h-H}V_2^*, \\ u^*(0,z,t) &= \frac{V_1^* + V_2^*}{h-H}. \end{aligned} \tag{72}$$

Taking into account condition (64) one can finally obtain

$$h^*(0,t) = \frac{c-\bar{u}}{h-H}V_1^*, \qquad u^*(0,z,t) = \frac{1}{h-H}V_1^*, \tag{73}$$

where V_1^* is the outgoing adjoint characteristic variable to be estimated from its interior values. It follows from (61),(63) that the sensitivity on $V_1(0,t)$ is

$$\nabla J(t) := \frac{\partial J(V_1(0,t))}{\partial V_1(0,t)} = -(c+\bar{u})V_1^*(0,t). \tag{74}$$

The boundary conditions for a 'physical' boundary can also be derived from the characteristic representation. For example, if the boundary $x = L$ is a wall (fluid/structure interface), then we can reasonably assume that

$$u(L,z,t) = 0, \qquad q(L,t) = 0.$$

It follows from the last assumption that $V_1 = -V_2$. Thus, we obtain

$$h(L,t) = Z + \frac{V_1}{c}, \qquad u(L,z,t) = 0, \tag{75}$$

where V_1 is an outgoing characteristic variable at $x = L$ and can be calculated based on its interior values. Similarly, for the adjoint problem one has

$$h^*(L,t) = -\frac{2c}{h-H}V_2^*, \qquad u^*(L,z,t) = 0, \tag{76}$$

where V_2^* is an outgoing adjoint characteristic variable at $x = L$ and can be calculated based on its interior values.

Although the approach presented here is approximate, it works reasonably well so long as at the boundaries the SWE model remains compatible with the barotropic fsNSE model, i.e., for the long-wave situation, as will be demonstrated in Section 8. Indeed, it is exact in the long-wave limit, which is approached as the wave-length becomes much greater than the depth. Nevertheless, we show here a practical way of designing boundary conditions for lateral boundaries for the fsNSE model, either forward or adjoint. The method can be generalized by using the characteristics of the multilayer SWE. This will allow us to consider the baroclinic fsNSE and introduce nonuniform in depth u-velocity profiles. Boundary conditions of that type are presently under investigation.

8. NUMERICAL EXAMPLES

8.1. General Description of the Numerics

A trial numerical implementation has been made using a finite-difference semi-explicit solution of the problem with a fixed regular mesh similar to the well-known SOLA algorithm [29]. Equations are discretized on a staggered grid using a hybrid scheme for advection terms. The Poisson equation for pressure is formed and resolved by the SOR method in the case of the forward equations and the TLM, and by a direct solver based on banded LU-decomposition in the case of the adjoint equation. In solving the forward model, the solution h, p, u, w is saved in the memory. This data is recalled when the TLM or the adjoint problem is running. Because the forward, the TLM, and the adjoint equations differ only in source terms a single solver is used. All calculations presented here have been performed for inviscid flow. No approximation operator for measurements is used.

In order to underline a general applicability of the method we design a special test case when the bed function suddenly changes from being deep to shallow. This change happens over one space discretization step, i.e., $\frac{dH}{dx}$ is bounded, and transformation (10) remains well posed. In reality this describes a situation near the 'shelf' edge. This is a severe test for methods that include the hydrostatic approximation in some way. A particular bathymetry sketch is shown in Figure 2. Channel dimensions are: $L = 3000\,\mathrm{m}$, $Z = 100\,\mathrm{m}$. The number of grid nodes used in all calculations presented here is $N_x = 100$ and $N_z = 20$, and the discretization steps are $\Delta x = 30\,\mathrm{m}$, $\Delta z = 5\,\mathrm{m}$, and $\Delta t = 0.36\,\mathrm{s}$. The velocity measurements are defined at all grid nodes at a chosen location in the middle of the shallow region ($x = 2250\,\mathrm{m}$).

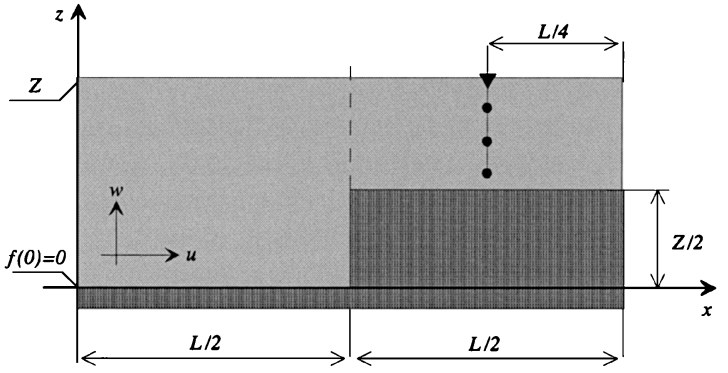


Figure 2. The bathymetry.

The numerical experiments carried out here are sometimes referred to as identical twin experiments. We consider the two forward solutions generated by two distinct control inputs $V_1^{(\mathrm{tr})}$ and $V_1^{(\mathrm{gs})}$, which may be interpreted as a 'true value' and a 'guess value'. We shall call $\delta V_1(t) = V_1^{(\mathrm{tr})}(t) - V_1^{(\mathrm{gs})}(t)$ the 'estimation error'. The difference between the two solutions at the measurement points yields the residuals r_u or r_h, which are used as driving sources for solving the adjoint problem to produce the sensitivity $\nabla J(t)$. When a need arises, we will use (u), (h), or (u,h) as subscripts to $\nabla J(t)$ in order to indicate that the sensitivities are obtained using either r_u, or r_h, or both. Now, the estimation error has to be eliminated in the course of an iterative minimization procedure, based on $\nabla J(t)$.

Let us consider a linear hyperbolic transport problem and assume that all the incoming characteristics (i.e., those which are directed from the control boundary toward the sensors) are measurable. Then the sensitivity has to coincide, within a scaling factor, with the estimation error. In this case one iteration of the steepest descend method should identify the boundary control. In reality the problem under consideration is not truly hyperbolic, and it is nonlinear. We

cannot perfectly get rid of numerical diffusion either. Nevertheless, comparing normalized values of the estimation error and the sensitivity function we can estimate the expected convergence rate. So, we consider normalized functions

$$\overline{\delta V}_1(t) := \frac{\delta V_1(t)}{\|\delta V_1(t)\|_{L_2}}, \qquad \overline{\nabla J}(t) := \frac{\nabla J(t)}{\|\nabla J(t)\|_{L_2}},$$

and a difference between them as follows:

$$\epsilon(t) = \overline{\delta V}_1(t) - \overline{\nabla J}(t).$$

In fact, the remainder ϵ may be considered as a new value of the estimation error after one optimization step, and $\|\epsilon(t)\|_{L_2}^{-1}$ as a convergence rate (in terms of the solution error). The situation is more complicated when the incoming characteristics are not measurable, i.e., measurements contain reflected signals. (The problem of 'ghost' signals is well known in inverse scattering. So far, no perfect algorithmic remedy is suggested to filter the incident signal.) Unlike diffusion that 'dissipates' information, reflections redistribute it over the spatial domain. Thus, it may happen that all sensors are located in the areas of low information content, i.e., we have information loss for an available set of sensors.

In order to generate the control variable $V_1(t)$ we use the expression as follows:

$$V_1(t) = \frac{Z}{2}(A + \delta A)\left(1 - \cos\left(\frac{\pi t}{P + \delta P}\right) + \delta\psi\right), \tag{77}$$

where A is the amplitude and P is the wave period. These parameters describe $V_1^{(\mathrm{tr})}(t)$. We use δA, δP, and $\delta\psi$ to produce a deviation from $V_1^{(\mathrm{tr})}(t)$. Then (77) gives us a 'guess value', i.e., $V_1^{(\mathrm{gs})}(t)$. Equation (77) defines a 'lifted' sinusoid function which generates simultaneously the two phenomena: a wave and a rise of the mean flow level. That produces a nonzero mean flow and causes vertical eddies to develop in the vicinity of the step.

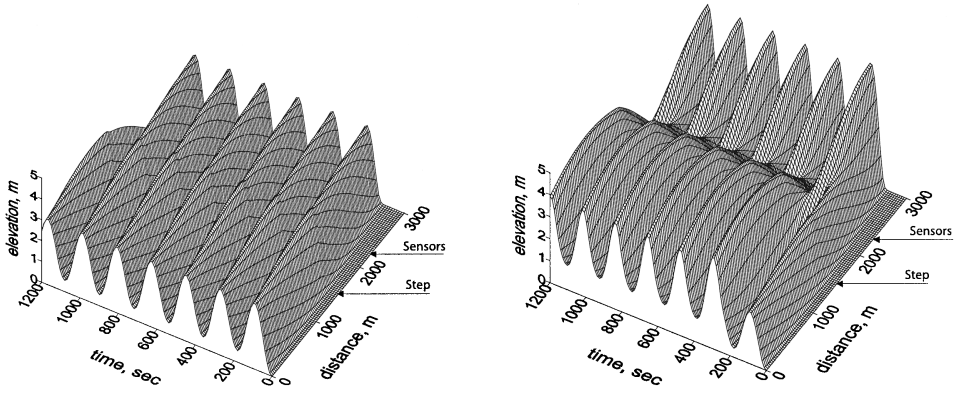
8.2. Results of the One-Step Identical Twin Experiment

We have performed seven numerical experiments all using the same channel configuration (Figure 2), but differing in the type of boundary conditions applied at the passive boundary, in the length of assimilation window T, in parameters used for (77) and (78) to generate 'true value' and 'guess value'. We refer to the 'guess value' as 'trivial' if it is equivalent to zero. A summary of these experiments is given in Table 1.

Table 1. Summary of numerical experiments.

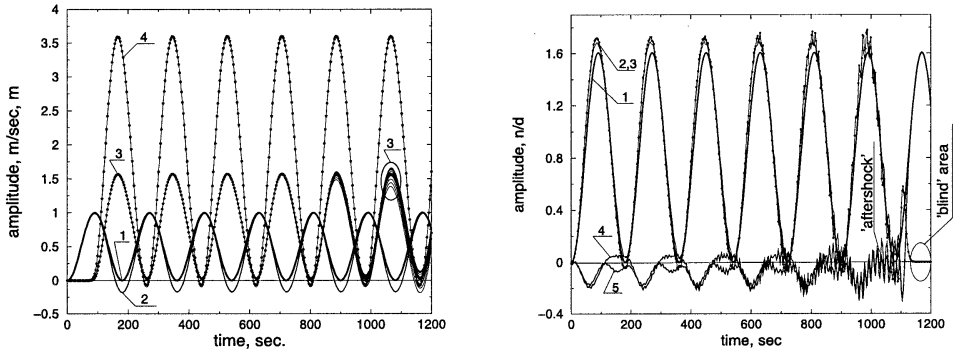
Name	Figures	B.C. at $x = L$	$T(h)$	Parameters in (77),(78): $A[\mathrm{ms}^{-1}]$, P[s]	
				True Value	Guess Value
A1	Figure 3a, Figure 4	open	1/3	$A = 1.0$, $P = 180$	trivial
A2	Figure 3b, Figures 5, 6	wall	1/3	$A = 0.75$, $P = 180$	trivial
B1	Figure 7	open	1	$A = 1.0$, $P = 432$	trivial
B2	Figure 8a	open	1	the same as B1	$\delta A = -0.1$, $\delta P = 0$, $\delta\psi = 0.3\pi$
B3	Figure 8b	open	1	the same as B1	$\delta A = 0.1$, $\delta P = 86$, $\delta\psi = 0.3\pi$
C1	Figure 9a, Figure 11	open	3	$A = 1.0$, $A_1 = 0.5$, $A_2 = -0.5$, $P = 2160$, $P_1 = 432$, $P_2 = 88$	trivial
C2	Figure 9b, Figure 11	open	3	the same as C1	$\delta A = 0.2$, $\delta A_1 = 0.1$, $\delta A_2 = -0.1$, $\delta P = -180$, $\delta P_1 = -36$, $\delta P_2 = -7.2$, $\delta\psi = 0$

We consider first waves of wave-length that is comparable to the length of the channel, within the observation window $T = 1200\,\mathrm{s}$. We model two types of passive boundary: 'open' boundary, Case A1, and 'wall', Case A2. The structure of the surface about the mean level Z for these two cases is shown in Figures 3a,b. In Case A1 we have a progressive wave in the shallow region (h is constant along the characteristics). Thus, the sensors located in the shallow region measure 'clean' incident data. In the deep region we have a mix of progressive and standing waves due to reflection from the step (h slightly changes along the characteristics). In Case A2 we have standing waves in both parts of the channel. One can clearly see the wave node at $x \approx 2400\,\mathrm{m}$ at which the surface has almost no periodic motion, but is lifted, i.e., $h > Z$. Therefore, the surface elevation does not carry information about the periodic component of $V_1^{(\mathrm{tr})}t$ at this location. In Figures 4a and 5a the value of $V_1^{(\mathrm{tr})}(t)$ is presented by line 1, the mean inlet velocity $\bar{u}(0,t)$ by line 2, the measured velocities $\hat{u}(x_1, z_k, t)$, $k = 1, \ldots, 10$, by line 3, and the measured elevations $\hat{h}(x_1, t)$ by line 4. Let us note the following: for progressive waves, Figure 4a, the velocity and the elevation are in phase, but for standing waves, Figure 5a, there is a phase shift $\pi/2$. For progressive waves the u-velocity profile at $x = x_1$ is uniform at first, and it remains uniform until the eddy near the step starts to produce effects downstream. This is shown in line 3, Figure 4a at $t \approx 1100\,\mathrm{s}$.



(a) Case A1. (b) Case A2.

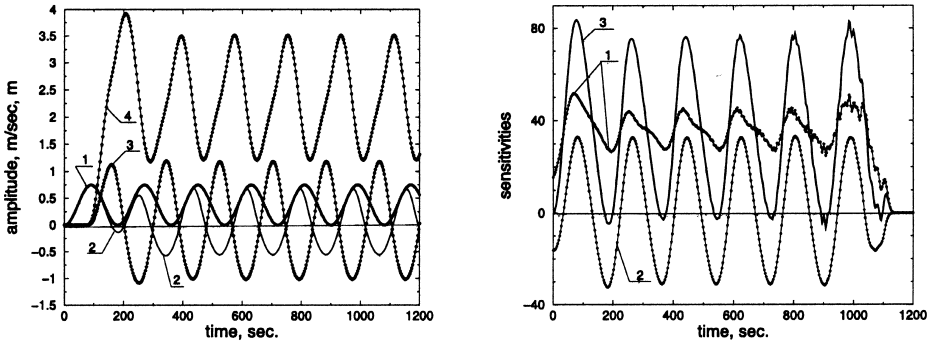
Figure 3. Structure of the free surface $h(x,t)$ for different boundary conditions at $x = L$.



(a) Control variable $V_1(t)$—line 1 and solutions from the forward model: mean flow velocity $\bar{u}(0,t)$—line 2; measurements $\hat{u}(x_1, z_k, t)$, $k = 1, \ldots, 10$—line 3; $\hat{h}(x_1, t)$—line 4.

(b) Estimation error $\overline{\delta V}_1(t)$—line 1 and solutions from the adjoint model: sensitivities $\overline{\nabla J}_h(t)$—line 2; $\overline{\nabla J}_u(t)$—line 3; and remainders $\epsilon_h(t)$—line 4; $\epsilon_u(t)$—line 5.

Figure 4. Case A1.



(a) Control variable $V_1(t)$—line 1 and solutions from the forward model: mean flow velocity $\bar{u}(0,t)$—line 2; measurements $\hat{u}(x_1,z_k,t)$, $k = 1,\ldots,10$—line 3; $\hat{h}(x_1,t)$—line 4.

(b) Solutions from the adjoint model: sensitivities $\overline{\nabla J}_h(t)$—line 1; $\overline{\nabla J}_u(t)$—line 2; and $\overline{\nabla J}_{h,u}(t)$—line 3.

Figure 5. Case A2.

A trivial 'guess value' $V_1^{(\mathrm{gs})}(t) \equiv 0$ is chosen first. Thus, the residuals coincide with the measurements and the estimation error $\delta V_1(t)$ coincides with $V_1^{(\mathrm{tr})}(t)$. Solving the adjoint equations we obtain sensitivities, which for the progressive wave, Case A1, are shown in Figure 4b. Here, line 1 shows $\overline{\delta V}_1(t)$; line 2—$\overline{\nabla J}_h(t)$; line 3—$\overline{\nabla J}_u(t)$; line 4—$\epsilon_h(t)$; and line 5—$\epsilon_u(t)$. The appearance of these sensitivities is quite typical. We can clearly see a 'blind' spot located at the end of the time domain (1130–1200 s), where there is no 'signal', i.e., the perturbation from the sensors has not yet reached the control boundary. Then follows an 'aftershock' area (800–1130 s), where we observe some initial oscillations related to the shock input. After those oscillations have passed beyond the domain, $\epsilon(t)$ becomes a regular function. For an inviscid problem this remainder is mostly due to the nonlinearity of the problem. For longer waves, the nonlinear contributions associated with $\frac{\partial h}{\partial t}$ and $\frac{\partial h}{\partial x}$ become smaller and $\|\epsilon\|_{L_2}$ decreases. It can be seen that the difference between $\overline{\nabla J}_h(t)$ and $\overline{\nabla J}_u(t)$ is very small. From this example we can draw the following conclusion: *if the sensors are located in a progressive wave area, then measurements of the elevation or the velocity profile performed at a single location provide necessary and sufficient information to identify the boundary control* $V_1(t)$.

Let us look now at Figure 5b, Case A2. We have a standing wave in the shallow part of the channel, i.e., the sensors do not measure only incident data. Here line 1 shows $\nabla J_h(t)$; line 2—$\nabla J_u(t)$. Neither of these two sensitivities actually resembles the error function. We have already mentioned that the reflections may cause loss (or redistribution) of information. Although the adjoint equations could be considered as a model that describes transport of information backward

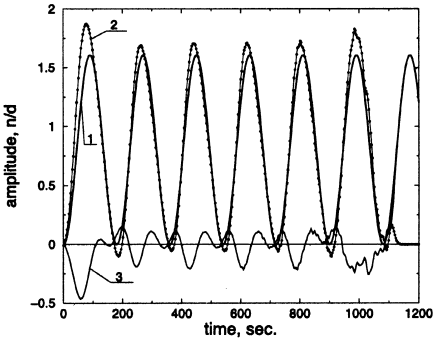
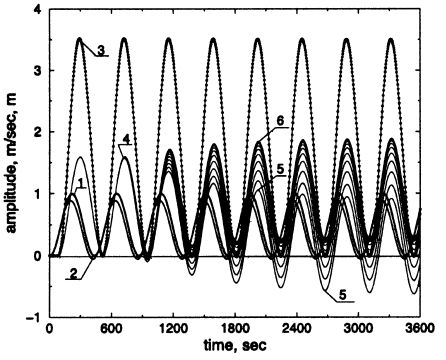


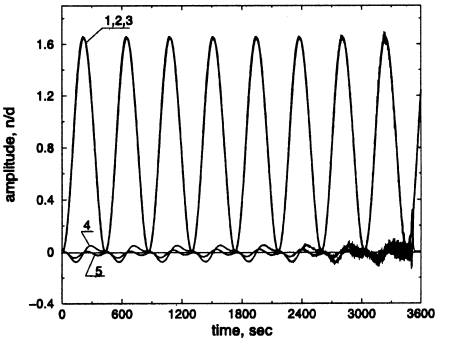
Figure 6, Case A2. Estimation error $\overline{\delta V}_1(t)$—line 1, and solutions from the adjoint model: sensitivity $\overline{\nabla J}_{h,u}(t)$—line 2; and remainder $\epsilon_{h,u}(t)$—line 3.

in time and space we can see that the adjoint run does not separate out incident and reflected components once they have become mixed. We have observed, however, that there always exists a weighted sum of $\nabla J_h(t)$ and $\nabla J_u(t)$, which does look like a good estimation of $\delta V_1(t)$. This function denoted as $\nabla J_{h,u}(t)$ is presented by line 3. The comparison is given in Figure 6. Here line 1 shows $\overline{\delta V}_1(t)$; line 2—$\overline{\nabla J}_{h,u}(t)$; and line 3—$\epsilon_{h,u}(t)$. Thus, another conclusion follows: *if the sensors are located in the area where standing waves occur, then co-located measurements of the elevation and the velocity profile are needed to provide necessary (but maybe not sufficient) information to identify the boundary control* $V_1(t)$. We have to mention that both conclusions have been drawn from our numerical experiments and should be rigorously validated.

Another group of numerical experiments has been carried out for longer waves and a longer observation window, $T = 1$ h, Cases B1–B3. The results of the forward modeling are presented in Figure 7a. Here, line 1 shows $V_1^{(\mathrm{tr})}(t)$; line 2—$\bar{u}(0,t)$; line 3—$\hat{h}(x_1,t)$; line 4—$\hat{u}(x_1,z_k,t)$, $k = 1,\ldots,10$. We can see that at the beginning the velocity profile is uniform in depth, but starting from $t \approx 700$–800 s the curve (given by line 4) splits and all the velocity sensors show different values. Here lines 5 and 6 show the velocities near the bed and near the surface. Thus one can observe an eddy: at some points there exists a reverse flow near the channel bed. For a trivial

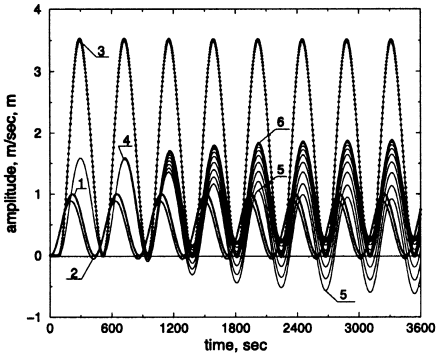


(a) Control variable $V_1(t)$—line 1 and solutions from the forward model: mean flow velocity $\bar{u}(0,t)$—line 2 and measurements $\hat{h}(x_1,t)$—line 3; $\hat{u}(x_1,z_k,t), k = 1,\ldots,10$—line 4; $\hat{u}(x_1,z_1,t)$—line 5; $\hat{u}(x_1,z_{10},t)$—line 6.

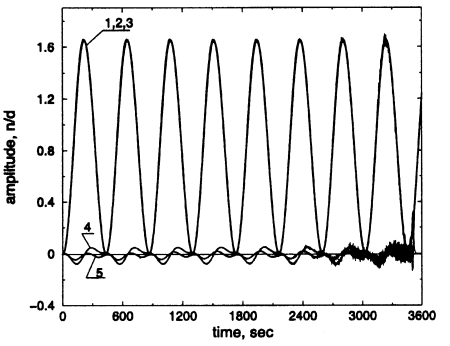


(b) Estimation error $\overline{\delta V}_1(t)$—line 1 and solutions from the adjoint model: sensitivities $\overline{\nabla J}_h(t)$—line 2; $\overline{\nabla J}_u(t)$—line 3 and remainders $\epsilon_h(t)$—line 4; $\epsilon_u(t)$—line 5.

Figure 7. Case B1.



Case B2.



(b) Case B3.

Figure 8. Estimation error $\overline{\delta V}_1(t)$—line 1 and solutions from the adjoint model: sensitivities $\overline{\nabla J}_h(t)$—line 2; $\overline{\nabla J}_u(t)$—line 3 and remainders $\epsilon_h(t)$—line 4; $\epsilon_u(t)$—line 5.

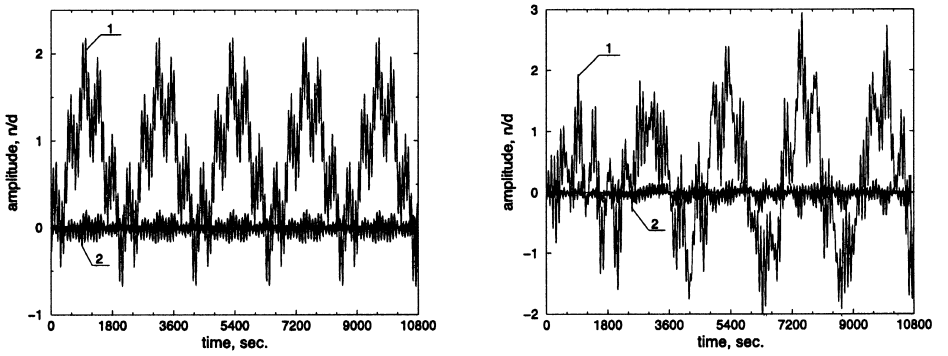
'guess value', Case B1, the sensitivities are presented in Figure 7b. Here, line 1 shows $\overline{\delta V}_1(t)$; line 2—$\overline{\nabla J}_h(t)$; line 3—$\overline{\nabla J}_u(t)$; line 4—$\epsilon_h(t)$; and line 5—$\epsilon_u(t)$. The same functions are shown in Figure 8, Cases B2 and B3, which instead show calculated sensitivities based on nontrivial 'guess values'.

We can see that the remainders $\epsilon_{(\cdot)}(t)$ are essentially smaller than those presented in Figures 4b and 6. This is because we consider longer waves and the nonlinearities associated with $\frac{\partial h}{\partial t}$ and $\frac{\partial h}{\partial x}$ are reduced. However, those generated by the step remain. The convergence rate for these examples is estimated as $\|\epsilon\|_{L_2}^{-1} \approx 20$–$30$.

In the next example, Cases C1 and C2, we introduce $V_1^{(\mathrm{tr})}$, which is given by (77) plus the term

$$Z(A_k + \delta A_k) \sin\left(\frac{\pi t}{P_k + \delta P_k} + \delta\psi\right), \qquad k = 1, 2. \tag{78}$$

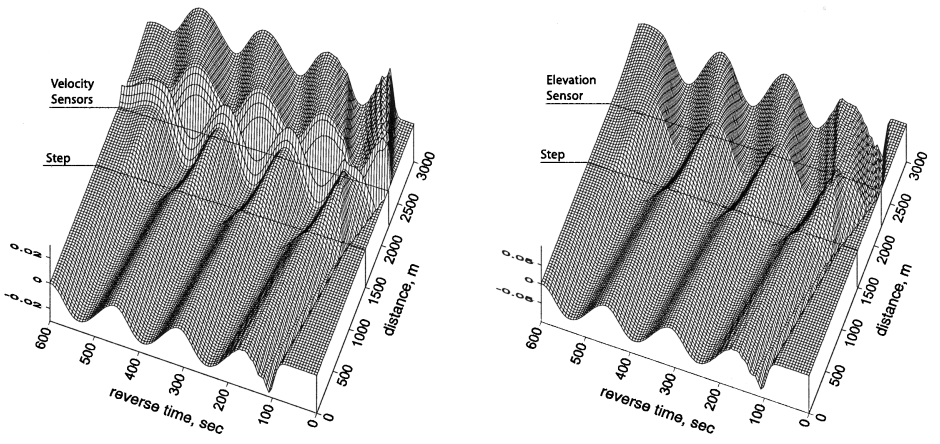
That is we now have three components. We model $V_1^{(\mathrm{tr})}(t)$ over the assimilation window $T = 3\,\mathrm{h}$ assuming an 'open' passive boundary. Both the elevation and the velocity data are used simultaneously. We calculate the sensitivity for a trivial 'guess value', Figure 9a, Case C1, and for a nontrivial 'guess value', Figure 9b, Case C2. Here, line 1 shows $\delta V_1(t)$ and line 2 shows $\epsilon_{h,u}(t)$. The experiments show no sign of a cumulative error. Therefore, the duration of the assimilation experiment is only limited by machine memory (we must keep the forward solution). The convergence rate for these examples is estimated as $\|\epsilon\|_{L_2}^{-1} \approx 15$, that is a little bit worse than in the previous examples. This is mostly due to the presence of a wave component with a period $P_2 = 88\,\mathrm{sec}$. In the present case we have a well-developed nonuniform in-depth u-velocity profile in both parts of the channel. This numerical experiment shows that the method works well to produce sensitivities and that no accumulation of error occurs even for complex flows and multicomponent data.



(a) Case C1. (b) Case C2.

Figure 9. Multicomponent variable control. Estimation error $\overline{\delta V}_1(t)$—line 1 and solution from the adjoint model: remainder $\epsilon_{h,u}(t)$—line 2.

It is worth looking at the patterns of the adjoint pressure generated by r_u and r_h, which are shown in Figures 10a,b, respectively. The adjoint pressure field is calculated for a trivial 'guess value'. The field values are taken at $z' = 1/2$. When it is driven by r_u, the adjoint pressure is a discontinuous function of x at the location of sensors. It can be seen that the wave and the antiwave are going in antipodal directions. For a straight channel p^* is a perfectly antisymmetrical function of x, $\forall t$. It is not quite so for the chosen bathymetry, because we have standing adjoint waves in the shallow part. We can also clearly see the difference in the adjoint wave celerity for deep and shallow regions with a transition near the step at $x = 1500\,\mathrm{m}$. In contrast, when driven by r_h adjoint fields are continuous, with discontinuity only in $\frac{\partial \cdot}{\partial x}$. Thus, the adjoint solution based on r_h can be obtained on a coarser grid and may require less computational work.



(a) (b)

Figure 10. Adjoint pressure field p^* at $z' = 1/2$ generated by r_u (a), and by r_h (b).

8.3. Verification of the Correctness of the Adjoint Model

To check that the adjoint model is correct and it is coded correctly we use the 'gradient test', which is an already-established method to test adjoint models [30]. Let

$$J(V_1 + \alpha\Psi) = J(V_1) + \alpha\Psi^\top \nabla J + O\left(\alpha^2\right)$$

be a Taylor expansion of the objective functional (9). Here α is a small scalar and $\Psi(t)$ is a vector-function of unit length. Rewriting the above formula we can define a function of α as

$$\phi(\alpha) := \frac{J(V_1 + \alpha\Psi) - J(V_1)}{\alpha\Psi^\top \nabla J} = 1 + O\left(\alpha^2\right). \tag{79}$$

When α decreases, the term $O(\alpha^2)$ vanishes and $\phi(\alpha) \to 1$. In a numerical implementation, however, this behavior is no longer valid as α approaches the machine zero. Considering $J(V_1)$ at the point $V_1 = V_1^{(\mathrm{gs})}$ and assuming that

$$\Psi(t) = \frac{\nabla J}{\|\nabla J\|_{L_2}}$$

we can rewrite (79) as follows:

$$\phi(\alpha) = \frac{J\left(V_1^{(\mathrm{gs})} + \alpha\nabla J/\|\nabla J\|_{L_2}\right) - J\left(V_1^{(\mathrm{gs})}\right)}{\alpha\|\nabla J\|_{L_2}}. \tag{80}$$

We use in (80) $J(\cdot)$ which is based on the numerical solution of the forward model, and ∇J, which is based on the numerical solution of the adjoint problem. For a consistent adjoint model (adjoint of the discretized model equations) discretization errors eliminate each other, and if everything is coded correctly, then one can achieve $\phi(\alpha)$ being as close to unity as the machine accuracy allows. Obviously, for an inconsistent adjoint model the value $\phi(\alpha)$ must be grid-and-case dependent, and serves actually as a consistency criterion between the two models. It shows whether the adjoint model is derived correctly, coded correctly, and, if both are true, the relative accuracy of the gradients. Now we consider the two cases which correspond to those shown in Figure 9. The results of the gradient test are presented in Figure 11. The solid line here corresponds to Case C1, and the dotted line to Case C2. One can see that as α decreases $\Phi(\alpha) := \log(|\phi(\alpha) - 1|)$ also decreases because of vanishing $O(\alpha^2)$. Then we have a flat shelf area

when $\Phi(\alpha)$ is insensitive to any change in α. The level of this shelf shows the relative accuracy of the gradients achievable for the inconsistent adjoint model. Then it starts to grow again and this happens because we approach the limits of the machine accuracy. In contrast, a typical result for a consistent adjoint model would be perfectly V-shaped without a shelf area, with a minimum value about 10^{-8}–10^{-10}. In conclusion we should mention that the level of relative accuracy of the gradients about $\Phi(\alpha) \approx 10^{-2} - 10^{-3}$ is sufficient to build a fast-converging minimization procedure. In fact, many optimization techniques, such as quasi-Newton methods, require only approximate gradients.

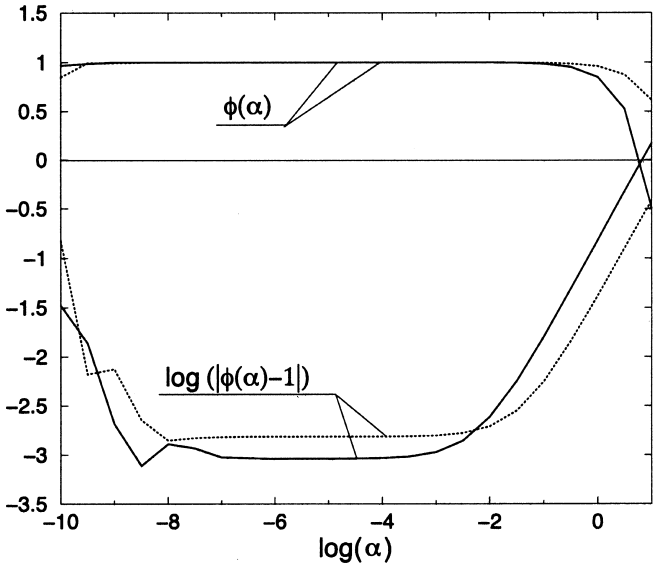


Figure 11. Results from the gradient test.

9. CONCLUSION

In this paper we present the *continuous (inconsistent) adjoint* model for barotropic Navier-Stokes equations including the free surface. The procedure for deriving the adjoint equations in the present case is not trivial. The main difficulty consists of deriving the variation on the surface elevation function h. We use a coordinate transformation in the vertical to shift the problem into a domain with fixed boundaries and calculate the variation in h from the Jacobian matrix of the transformation. This variation is then mapped back into the original coordinate system, where the variations in the other flow parameters are calculated in the usual way. Then we use scalar product formulas redefined for the domain with variable bounds to derive the adjoint equations and the surface and the bed boundary conditions. The adjoint equations derived in this way can be used in essentially nonhydrostatic cases such as sharp bed slopes and short waves. Although we consider a particular case when the boundary shape $h(x,t)$ is a one-valued function of x, the whole procedure can be considered as a method of deriving the adjoint equations for a general fluid flow problem with moving boundaries (at least for those cases where the domain remains connected). In the present case the transformation is defined in advance; otherwise such a transformation could be dynamically generated. In turn, the adjoint problem can be formulated in terms of the general transformation coefficients.

Another important issue is how to specify 'open' lateral boundaries for the fsNSE. This requires a rigorous characteristic analysis of the equations, even though they are not truly hyperbolic. We assume instead that in the vicinity of lateral boundaries our equations actually describe a shallow-water wave, and so they are locally hyperbolic. Therefore, we have used the characteristics of the SWE to derive boundary conditions and the sensitivity. Of course, within the domain the

full physics related to the fsNSE is properly represented. The numerical results validate this approach. Thus, the 'open' boundary introduced in this manner allows consideration of long waves and a uniform in-depth u-velocity profile at the boundary. As we already mentioned, the method can be generalized by coupling the fsNSE at the open boundary with a multilayer SWE model. This will enable us to consider the baroclinic case and a nonuniform in depth u-velocity profile.

The results of numerical experiments have validated the theory developed in this paper. The adjoint solver is verified by performing the 'gradient test'. In a case of a consistent adjoint model, the gradient test shows merely the correctness of coding, see [30]. In the case of inconsistent adjoint model, however, it shows:

(a) the correctness of derivation the adjoint equations, adjoint boundary conditions, and the sensitivity formulas;
(b) the level of consistency between the continuous forward and the discretized forward;
(c) the level of consistency between the continuous adjoint and the discretized adjoint;
(d) the correctness of coding.

Thus, all these steps are validated for the model presented here.

APPENDIX

VISCOUS TERMS

We use the same procedure as before for deriving the adjoint to viscous terms in the right-hand side of (3),(4). As the structure of these terms is similar for both equations we consider only the u-momentum equation. Let us denote

$$D_x(u) = \frac{\partial}{\partial x}\left(\mu_h \frac{\partial u}{\partial x}\right), \qquad D_z(u) = \frac{\partial}{\partial z}\left(\mu_v \frac{\partial u}{\partial z}\right).$$

First, we take the variations in $D_x(u)$ and $D_z(u)$ in respect to u in $\mathbb{R}$. Then applying (15)–(17) we obtain self-adjoint terms

$$\frac{\partial}{\partial x}\left(\mu_h \frac{\partial u^*}{\partial x}\right) + \frac{\partial}{\partial z}\left(\mu_v \frac{\partial u^*}{\partial z}\right) \tag{A.1}$$

to be placed in the right-hand side of equation (46) and the following boundary set:

$$\int_0^T \int_H^h \left[\mu_h \left(u^* \frac{\partial \tilde{u}}{\partial x} - \tilde{u} \frac{\partial u^*}{\partial x}\right)\right]_0^L dz\, dt \tag{A.2}$$

$$+ \int_0^T \int_0^L \left[u^* \left(\mu_v \frac{\partial \tilde{u}}{\partial z} - e_2 \mu_h \frac{\partial \tilde{u}}{\partial x}\right) - \tilde{u}\left(\mu_v \frac{\partial u^*}{\partial z} - e_2 \mu_h \frac{\partial u^*}{\partial x}\right)\right]_H^h dx\, dt. \tag{A.3}$$

In the viscous case, we use no-slip boundary conditions for the channel bed, which are $u(x, H(x), t) = 0$, $w(x, H(x), t) = 0$. These conditions yield similar conditions for the TLM, i.e., $\tilde{u}(x, H(x), t) = 0$, $\tilde{w}(x, H(x), t) = 0$. Thus, we annul both (A.3) and (52) taken at $z = H(x)$ by assuming

$$u^*(x, H(x), t) = 0, \qquad w^*(x, H(x), t) = 0. \tag{A.4}$$

Since we neglect the surface normal viscous stress by assuming that $\mu_{\vec{n}} = 0$, we also neglect (A.3) taken at $y = h$. Next, for simplicity we will follow the assumption that

$$[\mu_h]_{x=0,L} = 0. \tag{A.5}$$

Thus, the term (A.2) also vanishes. Although this assumption is somewhat artificial, it is perfectly consistent with our treatment of lateral boundaries, where we have to assume a local hyperbolicity.

In order to find the variations in $D_x(u)$ and $D_z(u)$ in respect to h we must transform the viscous terms to $\mathbb{R}'$. Let us denote $e_3' = e_2'/(h-H)$. Thus, we obtain

$$D_x(u) \geq D_1' + D_2' + D_3' + D_4',$$

where

$$D_1' = \frac{\partial}{\partial x'}\left(\mu_h' \frac{\partial u'}{\partial x'}\right), \qquad D_2' = -\frac{\partial}{\partial x'}\left(\mu_h' e_3' \frac{\partial u'}{\partial z'}\right),$$
$$D_3' = -e_3' \frac{\partial}{\partial z'}\left(\mu_h' \frac{\partial u'}{\partial x'}\right), \qquad D_4' = e_3' \frac{\partial}{\partial z'}\left(\mu_h' e_3' \frac{\partial u'}{\partial z'}\right).$$

The derivatives of these terms in respect to h are

$$\frac{\partial D_1'}{\partial h} = 0, \qquad \frac{\partial D_2'}{\partial h} = -\frac{\partial}{\partial x'}\left(\mu_h' \frac{\partial u'}{\partial z'} \frac{\partial e_3'}{\partial h}\right), \qquad \frac{\partial D_3'}{\partial h} = -\frac{\partial}{\partial z'}\left(\mu_h' \frac{\partial u'}{\partial x'}\right) \frac{\partial e_3'}{\partial h},$$
$$\frac{\partial D_4'}{\partial h} = \frac{\partial}{\partial z'}\left(\mu_h' e_3' \frac{\partial u'}{\partial z'}\right) \frac{\partial e_3'}{\partial h} + e_3' \frac{\partial}{\partial z'}\left(\mu_h' \frac{\partial u'}{\partial z'} \frac{\partial e_3'}{\partial h}\right).$$

Adding these derivatives and then using the inverse transformation we obtain

$$\frac{\partial D_x(u)}{\partial h} = -\frac{\partial}{\partial x}\left(\mu_h \frac{\partial u}{\partial z} F_0\right) - \frac{\partial}{\partial z}\left(\mu_h \frac{\partial u}{\partial x}\right) F_0, \tag{A.6}$$

where

$$F_0 =: (h-H)\left(\frac{\partial e_3'}{\partial h}\right)' = \frac{z-H}{h-H}\left(\frac{\partial \cdot}{\partial x} + e_2 \frac{\partial \cdot}{\partial z}\right) - \frac{e_2 \cdot}{h-H}. \tag{A.7}$$

Assuming (A.5) holds, i.e., no boundary set can appear, we calculate

$$\begin{aligned}\left\langle u^*, \frac{\partial D_x(u)}{\partial h} \cdot \tilde{h} \right\rangle_R &= \left\langle F_0 \cdot \tilde{h}, F_1(u, u^*) \right\rangle_R \\ &= \left\langle \tilde{h}, \frac{\partial}{\partial x}\left(\frac{z-H}{h-H} F_1(u, u^*)\right) \right. \\ &\quad \left. + \frac{\partial}{\partial z}\left(e_2 \frac{z-H}{h-H} F_1(u, u^*)\right) - \frac{e_2}{h-H} F_1(u, u^*) \right\rangle_R,\end{aligned}$$

where

$$F_1(u, u^*) = -\mu_h \frac{\partial u}{\partial z}\frac{\partial u^*}{\partial x} - \frac{\partial}{\partial z}\left(\mu_h \frac{\partial u}{\partial x}\right) u^*. \tag{A.8}$$

After integrating all with respect to z and using Leibniz's integration rule as in deriving (54) we finally obtain what is the contribution of $D_x(u)$ to the right-hand side of (48),

$$D_x^*(u, u^*) = \frac{\partial}{\partial x}\int_H^h \left(\frac{z-H}{h-H} F_1(u, u^*)\right) dz - \frac{1}{h-H}\int_H^h e_2 F_1(u, u^*)\, dz. \tag{A.9}$$

Following similar steps as above we derive $D_z^*(u)$ as follows:

$$D_z^*(u, u^*) = -\frac{2}{h-H}\int_H^h \frac{\partial}{\partial z}\left(\mu_v \frac{\partial u}{\partial z}\right) u^*\, dz. \tag{A.10}$$

Thus, in presence of viscous terms the right-hand side of (48) must contain

$$D_x^*(u, u^*) + D_x^*(w, w^*) + D_z^*(u, u^*) + D_z^*(w, w^*). \tag{A.11}$$

REFERENCES

1. S.J. Couch and G.J.M. Copeland, Tidal straining, mixing and Lagrangian flow residuals around headlands, *J. of Marine Environmental Engineering* **7** (2), 25–45, (2003).
2. B. Lin and R.A. Falconer, Three-dimensional layer-integrated modelling of estuarine flows with flooding and drying, *Estuarine, Coastal and Shelf Science* **44**, 737–751, (1997).
3. T. Ezer and G.L. Mellor, Simulations of the Atlantic Ocean with a free surface sigma coordinate ocean model, *J. Geophys. Res.* **102** (15), 647–657, (1997).
4. V. Casulli and P. Zanolli, Semi-implicit numerical modelling of non-hydrostatic free-surface flows for environmental problems, *Mathl. Comput. Modelling* **36**, 1131–1149, (2002).
5. M.M. Namin, B. Lin and R.A. Falconer, An implicit numerical algorithm for solving non-hydrostatic free-surface flow problems, *Int. J. Numer. Meth. Fluids* **35**, 341–356, (2001).
6. Y. Sasaki, Some basic formalism in numerical variational analysis, *Monthly Weather Rev.* **98**, 875–883, (1970).
7. G.I. Marchuk, *Numerical Solution of the Problems of the Dynamics of the Atmosphere and Oceans*, Gidrometeoizdat, Leningrad, (1974).
8. D.G. Cacuci, Sensitivity theory for non-linear systems, *J. Math. Phys.* **22** (12), 2794–2812, (1981).
9. X. Zou, I.M. Navon, M. Berger, M.K. Phua, T. Schlick and F.X. Le-Dimet, Numerical experience with limited-memory, quasi-Newton methods for large-scale unconstrained nonlinear minimization, *SIAM J. Optimization* **3**, 582–608, (1993).
10. B.F. Sanders and N.D. Katopodes, Adjoint sensitivity analysis for shallow-water wave control, *J. Eng. Mech., ASCE*, 909–919, (2000).
11. R.W. Lardner, Optimal control of open boundary conditions for a numerical tidal model, *Comput. Meth. Appl. Meth. Eng.* **102**, 367–387, (1992).
12. J. Zou, W.W. Hsieh and I.M. Navon, Sequential open-boundary control by data assimilation in a limited-area model, *Month. Weath. Review* **123** (9), 2899–2909, (1995).
13. Z. Li, I.M. Navon, M.Y. Hussaini and F.X. Le Dimet, Optimal control of cylinder wakes via suction and blowing, *Computers and Fluids* **32** (2), 149–171, (2003).
14. H.G. Arango, A.M. Moore, A.J. Miller, B.D. Cornuelle, E. Di Lorenzo and D.J. Neilson, The ROMS tangent linear and adjoint models: A comprehensive ocean prediction and analysis system, `http://marine.rutgers.edu/po/index.php?model=roms`, (2003).
15. I.Yu. Gejadze and G.J.M. Copeland, Adjoint sensitivity analysis for fluid flow with free surface, *Int. J. Numer. Meth. Fluids* **47**, 1027–1034, (2005).
16. A.M. Moore, H.G. Arango, A.J. Miller, B.D. Cornuelle, E. Di Lorenzo and D.J. Neilson, A comprehensive ocean prediction and analysis system based on the tangent linear and adjoint components of a regional ocean model, *Ocean Modelling* **7**, 227–258, (2004).
17. E. Blayo and L. Debreu, Revisiting open boundary conditions from the point of view of characteristic variables, *Ocean Modelling* **9** (3), 231–252, (2005).
18. B. Mohammadi and O. Pironneau, Shape optimization in fluid mechanics, *Annu. Rev. Fluid Mech.* **36** (11), 1–25, (2004).
19. J. Sokolowski and A. Zochowski, On the topological derivative in shape optimization, *SIAM J. Control Optim.* **37**, 1251–1272, (1999).
20. A. Jameson, *Aerodynamic Shape Optimization Using the Adjoint Method*, Lecture Series at the Von-Karman Institute, Brussels, Belgium, (2003).
21. E.H. Van Brummelen and A. Segal, Numerical solution of steady free-surface flows by the adjoint optimal shape design method, *Int. J. Numer. Meth. Fluids* **41**, 3–27, (2003).
22. N.P. Wedi and P.K. Smolarkiewicz, Extending Gal-Chen and Somerville terrain-following coordinate transformation on time-dependent curvilinear boundaries, *J. Comp. Phys.* **193**, 1–20, (2003).
23. G.W. Hedstrom, Nonreflecting boundary conditions for nonlinear hyperbolic system, *J. Comp. Phys.* **30**, 222–237, (1979).
24. V.I. Sedenko, Solvability of initial-boundary value problems for the Euler equations of flows on an ideal incompressible nonhomogeneous fluid and ideal barotropic fluid that are bounded by free surfaces, English translation, *Russian Acad. Sci. Sb. Math.* **83** (2), 347–368, (1995).
25. V.A. Solonnikov, Solvability of the problem of dynamics of viscous incompressible flow bounded by a free surface, *Continuous Media Dynamics* **23**, 123–128, (1975).
26. J. Oliger and A. Sandstrom, Theoretical and practical aspects of some initial boundary value problems in fluid mechanics, *SIAM J. Appl. Math.* **35**, 419–446, (1978).
27. J.L. Lions, R. Temam and S. Wang, On the equations of the large-scale ocean, *Nonlinearity* **5**, 1007–1053, (1992).
28. R. Temam and J. Tribbia, Open boundary conditions for the primitive and Boussinesq equations, *J. Atmos. Sci.* **60**, 2647–2660, (2003).
29. C. Hirt, B. Nichols and N. Romero, SOLA—A numerical solution algorithm for transient fluid flows, Technical Report LA-5852, Los-Alamos National Lab, Los Alamos, NM, (1975).
30. I.M. Navon, X. Zou, J. Derber and J. Sela, Variational data assimilation with an adiabatic version of the NMC spectral model, *Mon. Weather Rev.* **120**, 1433–1446, (1992).