

Research Article

Estimation of Open Boundary Conditions for an Internal Tidal Model with Adjoint Method: A Comparative Study on Optimization Methods

Haibo Chen,^{1,2} Chunbao Miao,¹ and Xianqing Lv¹

¹Laboratory of Physical Oceanography, Ocean University of China, Qingdao 266100, China

²Research and Development Center for Offshore Oil Safety and Environmental Technology, China Offshore Environmental Services Ltd., Qingdao 266061, China

Correspondence should be addressed to Xianqing Lv; xqinglv@ouc.edu.cn

Received 16 October 2012; Revised 7 January 2013; Accepted 10 January 2013

Academic Editor: Fatih Yaman

Copyright © 2013 Haibo Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Based on an internal tidal model, the practical performances of the limited-memory BFGS (L-BFGS) method and two gradient descent (GD) methods (the normal one with Wolfe's line search and the simplified one) are investigated computationally through a series of ideal experiments in which the open boundary conditions (OBCs) are inverted by assimilating the interior observations with the adjoint method. In the case that the observations closer to the unknown boundary are included for assimilation, the L-BFGS method performs the best. As compared with the simplified GD method, the normal one really uses less iteration to reach a satisfactory solution, but its advantage over the simplified one is much smaller than expected. In the case that only the observations that are further from the unknown boundary are assimilated, the simplified GD method performs the best instead, whereas the performances of the other two methods are not satisfactory. The advanced L-BFGS algorithm and Wolfe's line search still need to be improved when applied to the practical cases. The simplified GD method, which is controllable and easy to implement, should be regarded seriously as a choice, especially when the classical advanced optimization techniques fail or perform poorly.

1. Introduction

In the recent years, with the rapid advances in ocean observing systems, increasingly more oceanic observation data have become available. At the same time, mainframe supercomputers have become more powerful, with improvements in memory size, the introduction of multiprocessor capabilities as well as enhanced processor speed. These advances have provided us with an opportunity to improve the precision of numerical simulations of the ocean with these available observation data, for both operational and research purposes. Indeed, this is an active topic of a fast-growing research field, data assimilation, which is an effective method for marine research, and it has become widely used in meteorological and oceanographic predictions in recent years. Among all data assimilation methods, four-dimensional variational (4D-Var) data assimilation is often considered as one of the most effective and powerful approaches developed over

the past two decades. It is an advanced data assimilation method that involves the adjoint technique. This method has the advantage of directly assimilating various observations distributed in time and space into the numerical model and can maintain dynamical and physical consistency with the model at the same time. Thus, 4D-Var has been widely applied to meteorological and oceanographic data assimilation, sensitivity studies, and parameter estimation. The issue of 4D-Var applied to the shallow water equations model has been the subject of numerous works (see, e.g., [1–29] and references therein).

In this paper, we are interested in solving a large-scale optimization problem minimizing a cost functional

$$\min_{\mathbf{p}} J(\mathbf{p}) = \min_{\mathbf{p}} \left\{ \frac{1}{2} \int_{t_i}^{t_f} [C(\mathbf{x}) - \hat{\mathbf{x}}]^T K_{\mathbf{x}} [C(\mathbf{x}) - \hat{\mathbf{x}}] dt \right\}, \quad (1)$$

where \mathbf{p} represents the control variables and is defined by the open boundary conditions (OBCs) for a regional internal tidal model, which are denoted by a series of Fourier coefficients in the present work, t is the time, $[t_i, t_f]$ denotes the assimilation window where t_i and t_f are the initial and final time, respectively, and $\mathbf{x}(t)$ and $\hat{\mathbf{x}}(t)$ are the time-dependent state variable in an n -dimensional Euclidean space \mathbf{R}^n and the time-dependent observation variable in an m -dimensional Euclidean space \mathbf{R}^m , respectively. The operator $C : \mathbf{R}^n \rightarrow \mathbf{R}^m$ represents a projection from the n -dimensional space (\mathbf{R}^n) of the model solution \mathbf{x} to the m -dimensional space (\mathbf{R}^m) of observations. Superscript T indicates transpose. The components of \mathbf{x} are values of the various model fields (elevation, velocity, temperature, salinity, etc., alone or in combination) at each of the model grid locations. The number of components of \mathbf{x} is denoted by n , and the number m represents the number of observations at any given time. In general, n is larger than m . In fact, the objective function J is the weighted sum of squares of distance between the model solution and available observations distributed in space and time. K_x is the weighting matrix and theoretically should be the inverse of observation error covariance matrix [1, 2]. However, determining the “exact” form for K_x is far from easy [3]. Since this work represents a preliminary study in the application of an adjoint assimilation model, we proceed under the best and even unrealistic scenario by assuming that there are no observation errors and that the model solutions will have a perfect fit to the observations such that $\min_{\mathbf{p}} J(\mathbf{p}) = 0$. Therefore, K_x is simplified to be unit matrix [30].

The control variables \mathbf{p} and the state vector \mathbf{x} satisfy the time-dependent model equations of the form

$$\frac{\partial \mathbf{x}}{\partial t} = F(\mathbf{x}, \mathbf{p}), \quad (2)$$

where $F : \mathbf{R}^n \rightarrow \mathbf{R}^n$ is model vector function of \mathbf{x} . In general, the model equations (2) are nonlinear and are assumed to be twice continuously differentiable.

The objective of the adjoint assimilation method is to find model control variables \mathbf{p} that minimize the cost functional J . Note that J is only a function of the control variables \mathbf{p} because \mathbf{x} is uniquely defined by the model equations (2) for any prescribed \mathbf{p} . The control variables can be initial conditions, boundary conditions or parameters to be estimated, or combination of them. In a regional ocean model, the open boundary conditions, which must be prescribed to complete the model description at non-land boundaries, are very important and have a critical impact on the modeling results. Therefore, for simplicity, only OBCs are taken as control variables in this paper. A number of studies have been concerned about the OBC estimation with adjoint method. For example, the work of Lardner et al. [4] discussed the optimal control of OBC in the channel using a 2D adjoint tidal model. Seiler [5] performed a series of identical twin assimilation experiments using the adjoint method to estimate lateral open boundary values of stream function and relative vorticity for a quasigeostrophic open ocean model. Zou et al. [2] also developed a sequential open boundary

control scheme augmenting radiation conditions and applied it to idealized barotropic wind-driven ocean simulations. ten Brummelhuis et al. [6] used the data assimilation technique to estimate the OBCs for a shallow sea model of the entire European continental shelf. Heemink et al. [7] used the adjoint approach for a 3D shallow water flow system (TRIWAQ) to estimate the harmonic constants in the OBCs, the friction parameter, viscosity parameter, and water depth by assimilating tide gauge data as well as altimeter data. Shulman [8] proposed a data assimilation approach to specify open boundary conditions, and their method was then tested by Shulman et al. [9] with respect to the improvement of the model prediction skills of the M_2 tidal amplitudes and phases in the Yellow Sea. Taillandier et al. [10] used a four-dimensional variational data assimilation method to control boundary conditions in a nonstratified open coastal model. In the work of Zhang et al. [11], lateral tidal OBCs that forced tides in the internal region were estimated by assimilating predicted coastal tidal elevations into a 2D POM with the adjoint method. Gejadze and Copeland [12] studied the open boundary control problem for free-surface barotropic Navier-Stokes equations with adjoint data assimilation technology. By assimilating the tidal harmonic constants derived from T/P altimeter data with a 3D numerical barotropic adjoint tidal model constructed in [13], Zhang and Lu [14] optimized the OBCs and simulated the M_2 tide and tidal current in the Bohai and Yellow Seas. Based on a 2D tidal model, Cao et al. [15] and Guo et al. [16] studied the inversion of OBCs by assimilating the T/P altimeter data with adjoint method. Kazantsev [17] used a variational data assimilation technique to control boundary conditions on rigid boundaries for a shallow-water model. Recently, in the work of Chen et al. [18], an adjoint assimilation model for the simulation of internal tides was constructed and simply tested with a series of ideal experiments in which several prescribed spatial distributions of OBCs were successfully inverted by assimilating the model-generated pseudoobservations.

The previous minimization problem (1) requires large-scale optimization techniques. This problem attempts to find a solution of the model equations (2) that best fits, in the generalized least-squares sense, the observations distributed in space and time. In meteorology and oceanography, the fluid dynamics can be enforced through the use of a Lagrangian function constructed by appending the model equations to the cost function as constraints. To solve problem (1), many feasible large-scale optimization methods can be found in [31]. However, the number of studies discussing the performances of various optimization methods in the meteorological and oceanographic application is still relatively small. Wang et al. [19] proposed an adjoint truncated Newton algorithm for the large-scale unconstrained minimization in the meteorology application and applied this algorithm to a limited-area shallow-water equation model with model-generated data where the initial conditions served as control variables. Subsequently, Wang et al. [20] presented a new adjoint Newton algorithm for carrying out large-scale unconstrained optimization required in variational data assimilation and applied this algorithm to three 1D models and to one 2D limited-area shallow-water equation model

with both model-generated and First Global Geophysical Experiment data. Zou et al. [21] compared the performances of several limited-memory quasi-Newton and truncated Newton methods for unconstrained nonlinear optimization on some large-scale problems in oceanography and meteorology. Their results showed that among the tested limited-memory quasi-Newton methods, the limited-memory BFGS (L-BFGS) method of Liu and Nocedal [32] has the best overall performance for the problems examined, and the numerical performance of two truncated Newton methods, differing in the inner-loop solution for the search vector, is competitive with that of L-BFGS. More recently, Alekseev et al. [33] compared the performance of several advanced large-scale minimization algorithms, including the conjugate gradient method, quasi-Newton (BFGS), the limited-memory quasi-Newton (L-BFGS), Hessian Free Newton method, and a new hybrid algorithm, applied to the minimization of the cost functional in the solution of ill-posed inverse computational fluid dynamics problems related to parameter estimation. In the work of [14], the comparison of a simple gradient descent (GD) method with the L-BFGS algorithm was conducted within a 3D barotropic tidal model, and the results demonstrate that, compared with the steepest descent method, the L-BFGS really uses much fewer steps to reach a minimum, while the values of the minimized cost function and inverted Fourier coefficient are almost the same. In the work of [18], the L-BFGS version of Liu and Nocedal [32], combined with the adjoint assimilation technique, was applied to the optimization of the OBCs for an internal tidal model and was proved to be very efficient.

The main objective of this paper is to make a computational investigation on the performances of the L-BFGS method and two versions of GD method. All these methods do not require any evaluations of the Hessian matrices but gradient vectors and thus are computationally feasible. In the present work, the number of the control variables (OBCs) is often large. In this case, the limited-memory methods are suggested, and one of the most commonly used limited-memory methods is the L-BFGS method [31, 32]. The L-BFGS method is an adaptation of the standard BFGS method to large-scale optimization problems. This method has been proved to be globally convergent for convex objective functions and provides a fast rate of linear convergence as well as requiring minimal storage. As the competitor, we choose the GD method. The GD method is a fundamental and probably the earliest method for nonlinear optimization so that it attracts many researchers' attention since its proposition. A recent and interesting report on the GD method can be found in [31, 34]. The GD method has the negative gradient direction as the search direction and is one of the simplest minimization algorithms that require calculation of derivatives. The GD method is particularly useful when the dimension of the problem is very large [35]. Two versions of GD method are considered here, differing in the formulation of the step length. One normally uses the step length satisfying the strong Wolfe conditions [31, 36, 37] which have been widely used in line search (see [38, 39] and references therein). And the other is a simplified GD method without the line search which has been used in a relatively

small number of studies [13, 14, 22, 23]. In this paper, based on the model constructed by Chen et al. [18], the performances of the L-BFGS method and the two versions of GD method are compared through a series of ideal experiments in which the OBCs for a numerical internal tidal model are estimated by assimilating the interior observations with the adjoint method.

This paper is organized as follows. We start with a brief introduction of the two-layered internal tidal model in Section 2. For the details of the model, we refer interested readers to [18]. The optimization methods for comparison, including the L-BFGS method and two GD methods, are described briefly in Section 3. In Section 4, as illustrative numerical examples, a series of ideal experiments are carried out to invert the prescribed OBCs with the different methods examined here, and the experimental results are presented. Finally, we make a summary and draw some conclusions in Section 5.

2. Internal Tidal Model

2.1. Forward Model. A two-layer version of the numerical internal tidal model described in [18] is considered in this paper. Assuming that the potential density ρ_k ($k = 1, 2$) in each layer is constant, the layer-averaged, nonlinear, time-dependent continuity and momentum equations of each layer subject to the hydrostatic approximations are derived from the primitive 3D governing equations. Here, the subscripts 1 and 2 refer to the upper and lower layers, respectively. Using spherical coordinates in the horizontal direction and isopycnic coordinates in the vertical one, we obtain the internal mode equations as follows:
upper layer:

$$\frac{\partial q_1}{\partial t} + \frac{1}{R \cos \phi} \frac{\partial (q_1 u_1)}{\partial \lambda} + \frac{1}{R \cos \phi} \frac{\partial (q_1 v_1 \cos \phi)}{\partial \phi} = 0, \quad (3a)$$

$$\begin{aligned} \frac{\partial u_1}{\partial t} + \frac{u_1}{R \cos \phi} \frac{\partial u_1}{\partial \lambda} + \frac{v_1}{R} \frac{\partial u_1}{\partial \phi} - \frac{u_1 v_1 \tan \phi}{R} \\ - f v_1 + \frac{g}{R \cos \phi} \frac{\partial}{\partial \lambda} \sum_{m=1}^l \left(\frac{q_m}{\rho_m} - h_m \right) \\ - A_h \Delta u_1 + \frac{A_v \bar{\rho}}{q_1 \bar{h}} (u_1 - u_2) = 0, \end{aligned} \quad (3b)$$

$$\begin{aligned} \frac{\partial v_1}{\partial t} + \frac{u_1}{R \cos \phi} \frac{\partial v_1}{\partial \lambda} + \frac{v_1}{R} \frac{\partial v_1}{\partial \phi} + \frac{u_1^2 \tan \phi}{R} \\ + f u_1 + \frac{g}{R} \frac{\partial}{\partial \phi} \sum_{m=1}^l \left(\frac{q_m}{\rho_m} - h_m \right) \\ - A_h \Delta v_1 + \frac{A_v \bar{\rho}}{q_1 \bar{h}} (v_1 - v_2) = 0; \end{aligned} \quad (3c)$$

lower layer:

$$\frac{\partial q_2}{\partial t} + \frac{1}{R \cos \phi} \frac{\partial (q_2 u_2)}{\partial \lambda} + \frac{1}{R \cos \phi} \frac{\partial (q_2 v_2 \cos \phi)}{\partial \phi} = 0, \quad (4a)$$

$$\begin{aligned} & \frac{\partial u_2}{\partial t} + \frac{u_2}{R \cos \phi} \frac{\partial u_2}{\partial \lambda} + \frac{v_2}{R} \frac{\partial u_2}{\partial \phi} - \frac{u_2 v_2 \tan \phi}{R} \\ & - f v_2 + \frac{g}{R \cos \phi} \frac{\partial}{\partial \lambda} \sum_{k=1}^2 \left(\frac{q_k}{\rho_2} - h_k \right) \end{aligned} \quad (4b)$$

$$\begin{aligned} & - A_h \Delta u_2 - \frac{A_v \bar{\rho}}{q_2 \bar{h}} (u_1 - u_2) \\ & + \frac{\kappa \rho_2 u_2}{q_2} \sqrt{(u_2)^2 + (v_2)^2} = 0, \end{aligned}$$

$$\begin{aligned} & \frac{\partial v_2}{\partial t} + \frac{u_2}{R \cos \phi} \frac{\partial v_2}{\partial \lambda} + \frac{v_2}{R} \frac{\partial v_2}{\partial \phi} + \frac{u_2^2 \tan \phi}{R} \\ & + f u_2 + \frac{g}{R} \frac{\partial}{\partial \phi} \sum_{k=1}^2 \left(\frac{q_k}{\rho_2} - h_k \right) \end{aligned} \quad (4c)$$

$$\begin{aligned} & - A_h \Delta v_2 - \frac{A_v \bar{\rho}}{q_2 \bar{h}} (v_1 - v_2) \\ & + \frac{\kappa \rho_2 v_2}{q_2} \sqrt{(u_2)^2 + (v_2)^2} = 0. \end{aligned}$$

Here, t is the time, λ and ϕ are the eastern longitude and northern latitude, respectively, $u(\lambda, \phi, t)$ and $v(\lambda, \phi, t)$ are horizontal velocities in λ and ϕ , respectively, $q(\lambda, \phi, t)$ is the time-varying layer mass, and

$$q_1 = \rho_1 (h_1 + \eta_1 - \eta_2), \quad q_2 = \rho_2 (h_2 + \eta_2), \quad (5)$$

where h_k and η_k are the undisturbed layer thickness and interface (surface for $k = 1$) elevation above the undisturbed level, respectively. R is the radius of the earth, g the gravitational acceleration, f the Coriolis parameter, and $f = 2\Omega \sin \phi$, where Ω represents the angular speed of Earth's rotation, A_h the horizontal eddy viscosity coefficient, Δ the Laplace operator, and

$$\Delta(u, v) = \frac{1}{R^2 \cos^2 \phi} \frac{\partial^2 (u, v)}{\partial \lambda^2} + \frac{1}{R^2 \cos \phi} \frac{\partial}{\partial \phi} \left[\cos \phi \frac{\partial (u, v)}{\partial \phi} \right]. \quad (6)$$

A_v , and κ are the interface and bottom friction coefficient, respectively, $\bar{\rho} = (\rho_1 + \rho_2)/2$, and $\bar{h} = (h_1 + h_2)/2$. In the forward model q , u and v are the main outputs and called the state variables in this paper.

By defining the barotropic currents

$$\begin{aligned} U &= \frac{(q_1 u_1 + q_2 u_2)}{Q}, \\ V &= \frac{(q_1 v_1 + q_2 v_2)}{Q}, \quad \text{where } Q = q_1 + q_2, \end{aligned} \quad (7)$$

we can derive a 2D external mode from the internal mode equations. For further details about the external mode, we refer the interested reader to the previous work of [18].

Usually, the model domain Σ is defined within a certain range of space and time, enclosed by the initial and boundary conditions. In the present model, a zero field is assigned to the model state variables as the initial conditions. The model is run for several tidal cycles, and the simulation results in the last cycle are taken as the model output. The boundary located in the wet grid is treated as the open boundary, and the boundary located in the dry grid (e.g., the island and the land) is treated as the closed boundary. The details of the boundary conditions are described as follows.

The Flather condition which was originally proposed by Flather [40] yielded good results [18, 41] when applied to the open boundaries in this model. In this paper, an adaptation of the Flather condition used by [18] is installed in the external mode of the forward model, which is given as follows:

$$\begin{aligned} \eta_1 &= \bar{\eta} \pm \sqrt{\left(1 - \frac{f^2}{\omega^2}\right) \frac{H}{g}} U, \\ \eta_1 &= \bar{\eta} \pm \sqrt{\left(1 - \frac{f^2}{\omega^2}\right) \frac{H}{g}} V, \end{aligned} \quad (8)$$

where $\bar{\eta}$ is external data beyond the model boundary representing the clamped surface elevation relating to the boundary barotropic tidal force and $H = h + \eta_1$ is the time-varying total water depth. The sign in (8) depends on the boundary location (positive for eastern and northern boundaries; negative for western and southern boundaries).

For the internal mode, the relaxation conditions, which have been found to have a robust performance in a variety of situations [42, 43], are applied at open boundaries. In this condition, a relaxation region of several grids close to the boundary is defined. Within this region, total tidal values including the interface elevation and fluid velocity of both barotropic and baroclinic tides are first calculated using the standard discretization and then relaxed towards the barotropic values as follows:

$$x(i, j) = x_{Bt}(i, j) + [1 - \mu(i, j)] \cdot x_{Bc}(i, j), \quad (9)$$

to give the final values x . Here, x represents the tidal states η_k , u_k , or v_k , x_{Bt} and x_{Bc} represent the barotropic and baroclinic tides, respectively, and $\mu(i, j) \in [0, 1]$ is the relaxation factor in grid (i, j) . In our model, μ is chosen to be increased linearly from 0 to 1 while the grid is getting close to the open boundary.

Closed boundary conditions for both the internal and the external modes are zero flow normal to the coast; that is, $\vec{u} \cdot \vec{n} = 0$ for the grid points at closed boundary, where \vec{n} is the outward unit vector and $\vec{u} = (u, v)$ is the velocity vector.

2.2. Adjoint Model. The adjoint method is a powerful tool for parameter estimation. The basic idea of the adjoint method is quite simple: a model is defined by an algorithm and its control variables including initial conditions, boundary conditions, and empirical parameters. The cost function that

measures the data misfit between the model output and observations is minimized through optimizing the control variables. In detail, the cost function decreases along a certain search direction which can be calculated with a certain optimization algorithm according to the gradient of cost function with respect to the control variables, and this gradient is calculated by what has been known as the adjoint model. Based on the governing equations (3a)–(4c) of the forward model, its adjoint model can be constructed as follows. The details of the adjoint model in a generalized form (not limited to the two-layer case) can be found in the work of [18].

Concretely, the cost function is defined by

$$J(q, u, v; \mathbf{p}) = \frac{1}{2} \int_{\Sigma} \left[\sum_{k=1}^2 (u_k - \hat{u}_k)^2 + \sum_{k=1}^2 (v_k - \hat{v}_k)^2 \right] d\sigma, \quad (10)$$

where k is layer index, Σ denotes the model domain of both time and space, \mathbf{p} represents the generalized control variables and is defined by the open boundary conditions (OBCs) for a regional internal tidal model, which are denoted by a series of Fourier coefficients in the present work, u_k and v_k are the models simulated, and \hat{u}_k and \hat{v}_k are the observations. In (10), the variables u_k and v_k can be uniquely determined by the control \mathbf{p} via the model equations (3a)–(4c). Hence, the functional J is implicitly dependent on \mathbf{p} .

The Lagrangian function is defined by

$$\begin{aligned} L(q, u, v; q_a, u_a, v_a; \mathbf{p}) &= J(q, u, v; \mathbf{p}) \\ &+ \int_{\Sigma} [q_{a1} \cdot F_{(3a)} + u_{a1} q_1 \cdot F_{(3b)} \\ &\quad + v_{a1} q_1 \cdot F_{(3c)}] d\sigma \\ &+ \int_{\Sigma} [q_{a2} \cdot F_{(4a)} + u_{a2} q_2 \cdot F_{(4b)} \\ &\quad + v_{a2} q_2 \cdot F_{(4a)}] d\sigma, \end{aligned} \quad (11)$$

where q_{ak} , u_{ak} , and v_{ak} are called adjoint variables of the state variables q_k , u_k , and v_k , respectively. $F_{(3a)}$, $F_{(3b)}$, \dots , $F_{(4c)}$ denote the left-hand side of (3a), (3b), \dots , (4c), respectively.

According to the typical theory of Lagrangian multiplier method, we have the following first-order derivatives of Lagrangian function with respect to all the variables and parameters:

$$\frac{\partial L}{\partial q_{ak}} = 0, \quad \frac{\partial L}{\partial u_{ak}} = 0, \quad \frac{\partial L}{\partial v_{ak}} = 0; \quad (12)$$

$$\frac{\partial L}{\partial q_k} = 0, \quad \frac{\partial L}{\partial u_k} = 0, \quad \frac{\partial L}{\partial v_k} = 0; \quad (13)$$

$$\frac{\partial L}{\partial \mathbf{p}} = 0. \quad (14)$$

Equations (12) return the governing equations (3a)–(4c). The adjoint equations can be derived using (13). From (14), we

can obtain the gradients of the cost function with respect to control variables.

Similar to the forward model, the adjoint model also consists of the internal and external modes. Actually, the equations derived from (13) are considered as the internal mode, and the external mode can be derived from the internal mode in a similar way the external mode of the forward model is derived. The details of both the internal and external modes of the adjoint model can be found in [18].

2.3. Discretization. Several numerical methods have been widely used in the discretization of time-dependent 3D primitive equations. The time integration schemes of these methods can be fully explicit [44], semi-implicit [45, 46] or fully implicit [47]. For large-scale oceanic problems, the applications of 3D models are becoming a reality with the aid of modern computers. The fully explicit finite difference method is relatively simple to implement, except that its time step is strictly restricted by the Courant-Friedrich-Lewy (CFL) stability criterion [48]. At present, many existing ocean models are based on an Alternating Direction Implicit (ADI) method which was proposed for the approximate solution of the shallow-water equations in [49]. ADI method results in computational efficiency superior to fully explicit methods because their improved stability allows large time steps to be employed, and it is also easy for implementation (see [13, 14]). Since the model must simulate fields of both velocity and elevations in each isopycnic layer, a technique known as external-internal mode splitting has been used in several ocean models by Simons [50]. Complete details on the model discretization were given in [18].

2.4. Gradients. Among all the control variables in this model, the OBCs are the most important and have critical impacts for a regional tidal simulation. Solutions in model interior are uniquely determined by the tidal OBCs once the initial conditions and other parameters have been determined. The Flather condition (8) can be determined once the external data $\tilde{\eta}$ is described. Hence, the determination of the open boundary conditions for the present model is equivalent to the estimation of the tidal force $\tilde{\eta}$. Assume that at a certain open boundary grid point (i_B, j_B) , the M_2 tidal force $\tilde{\eta}$ at the n th time step is subject to

$$\tilde{\eta}_{i_B, j_B}^n = a_{i_B, j_B} \cos(\omega n \Delta t) + b_{i_B, j_B} \sin(\omega n \Delta t), \quad (15)$$

where ω denotes the frequency of M_2 constituent, Δt is the time step length, and a_{i_B, j_B} and b_{i_B, j_B} are the Fourier coefficients as well as tunable parameters in the model. The gradients of cost function with respect to a_{i_B, j_B} and b_{i_B, j_B} can be deduced from (14) which yields

$$\frac{\partial J}{\partial a_{i_B, j_B}} + \sum_n T_{i_B, j_B}^n \cos(\omega n \Delta t) = 0, \quad (16)$$

$$\frac{\partial J}{\partial b_{i_B, j_B}} + \sum_n T_{i_B, j_B}^n \sin(\omega n \Delta t) = 0.$$

Details on computing T_{i_B, j_B}^n were given in [18].

Having determined the adjoint variables (q_a , u_a and v_a) with the adjoint model, the gradients of cost function with respect to the OBCs (i.e., the Fourier coefficients a and b) can thus be calculated. Then, the OBCs are optimized with some minimization algorithms which will be described in the following section.

3. Optimization Methods

In general, numerical methods solving the minimization problem (1) have the following iterative formula:

$$\mathbf{p}_{k+1} = \mathbf{p}_k + \alpha_k \mathbf{d}_k, \quad (17)$$

where \mathbf{p}_k , α_k , and \mathbf{d}_k are current iterative point, a positive step length, and a search direction, respectively. For simplicity, $J(\mathbf{p}_k)$ and $\nabla J(\mathbf{p}_k)$ are denoted by J_k and \mathbf{g}_k , respectively. There are many formulas to define search direction \mathbf{d}_k in formula (17) [31]. In this paper, three optimization methods are compared when applied to the assimilation model. They are different in the selection of α_k or \mathbf{d}_k . The three methods are given as follows.

3.1. L-BFGS Method. The L-BFGS method is a well-known limited-memory quasi-Newton method for solving large-scale problems whose Hessian matrices cannot be computed at a reasonable cost or are not sparse [31]. It requires the search direction to be

$$\mathbf{d}_k = -\mathbf{H}_k \mathbf{g}_k, \quad (18)$$

where

$$\begin{aligned} \mathbf{H}_k = & (\mathbf{V}_{k-1}^T \cdots \mathbf{V}_{k-m}^T) \mathbf{H}_k^0 (\mathbf{V}_{k-m} \cdots \mathbf{V}_{k-1}) \\ & + \delta_{k-m} (\mathbf{V}_{k-1}^T \cdots \mathbf{V}_{k-m+1}^T) \mathbf{s}_{k-m} \mathbf{s}_{k-m}^T (\mathbf{V}_{k-m+1} \cdots \mathbf{V}_{k-1}) \\ & + \delta_{k-m+1} (\mathbf{V}_{k-1}^T \cdots \mathbf{V}_{k-m+2}^T) \\ & \times \mathbf{s}_{k-m+1} \mathbf{s}_{k-m+1}^T (\mathbf{V}_{k-m+2} \cdots \mathbf{V}_{k-1}) \\ & + \cdots \\ & + \delta_{k-1} \mathbf{s}_{k-1} \mathbf{s}_{k-1}^T. \end{aligned} \quad (19)$$

Here, $\delta_k = 1/\mathbf{y}_k^T \mathbf{s}_k$, $\mathbf{V}_k = \mathbf{I} - \delta_k \mathbf{y}_k \mathbf{s}_k^T$, $\mathbf{s}_k = \mathbf{p}_{k+1} - \mathbf{p}_k = \alpha_k \mathbf{d}_k$, $\mathbf{y}_k = \mathbf{g}_{k+1} - \mathbf{g}_k$, \mathbf{I} is the identity matrix, and \mathbf{H}_k^0 is the initial Hessian approximation. Many previous studies have shown that typically $3 \leq m \leq 7$, where $m > 7$ does not improve the performance of L-BFGS. In this paper, the L-BFGS version of Liu and Nocedal [32] is employed, and the Fortran codes are authored by Nocedal [51], while the number of corrections m used in the L-BFGS update is taken as 5 [33].

3.2. Gradient Descent Method with Wolfe's Line Search (GDM-W). This method requires the search direction to satisfy

$$\mathbf{d}_k = -\mathbf{g}_k, \quad (20)$$

while the step length α_k satisfies the following strong Wolfe conditions [36, 37]:

$$J(\mathbf{p}_k + \alpha_k \mathbf{d}_k) \leq J_k + c_1 \alpha_k \mathbf{d}_k^T \mathbf{g}_k, \quad (21a)$$

$$\left| \nabla J(\mathbf{p}_k + \alpha_k \mathbf{d}_k)^T \mathbf{d}_k \right| \leq c_2 \left| \mathbf{g}_k^T \mathbf{d}_k \right| \quad (21b)$$

with $0 < c_1 < c_2 < 1$. The line search routine enforcing the strong Wolfe conditions can also be found in [31].

3.3. Simplified Gradient Descent Method (GDM-S). This method requires the search direction to satisfy

$$\mathbf{d}_k = -\mathbf{g}_k, \quad (22)$$

while the step length α_k is simply defined by

$$\alpha_k = \frac{\alpha_0}{\|\mathbf{g}_k\|_2}, \quad (23)$$

where the denominator $\|\mathbf{g}_k\|_2$ denotes the L_2 norm of the gradient \mathbf{g}_k and α_0 is an empirical constant. This method has been used in a relatively small number of studies (see [13, 14, 23, 52]) showing that this method is indeed feasible and effective in practical application.

4. Numerical Experiments

4.1. Numerical Implementation. The performances of L-BFGS, GDM-W, and GDM-S are compared by a set of ideal experiments. The computing domain has 51×41 horizontal grids (with the horizontal resolution $10' \times 10'$) and two vertical isopycnic layers (Figures 1(a) and 1(b)). The maximum undisturbed water depth is 6000 m, and the undisturbed interface is placed at the depth of 2000 m. The 2D bottom topography is constructed by the superposition of an inclined plane and a Gaussian surface (Figure 1(b)). The angular frequency of M_2 tide is $1.4050789025 \times 10^{-4} \text{ s}^{-1}$, and the whole-time step is 496.863 s (1/90 of the period of M_2 tide) for both external and internal modes. The horizontal eddy viscosity coefficient is chosen to be $A_h = 1000$. The coefficients of bottom and interface frictions are taken as $\kappa = 0.0025$ and $A_v = 0.003$, respectively.

As shown in Figure 1(a), all four boundaries are open and installed with the passive Flather condition allowing the outgoing wave to radiate out of the computing domain through the open boundaries. In this work, for simplicity and no loss of generality, only the eastern boundary is treated as the source of tidal force driving the forward model (meaning that both Fourier coefficients a and b are equal to zero on the other three boundaries), and only the estimation of Fourier coefficient a is studied. To be specific, on the eastern boundary, the Fourier coefficient b is always treated as the known parameter (equal to 0 in this work), and the Fourier coefficient a to be estimated here is assumed to be spatially

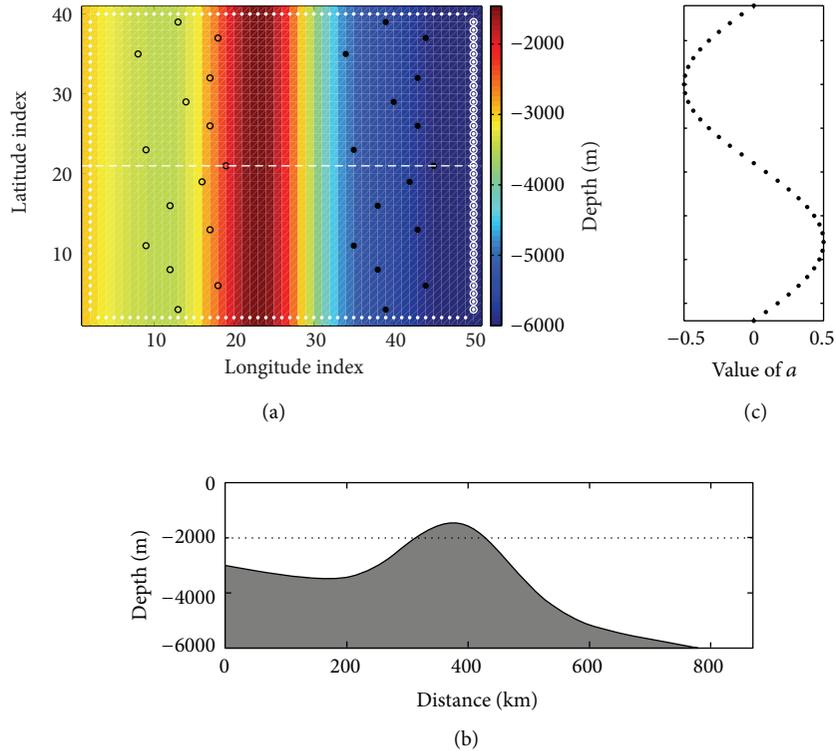


FIGURE 1: Experiment design. (a) Plane view of the computing domain and observation locations. The black dots and the black open circles represent the locations of the observations O1 and O2, respectively. The white dots and the white open circles denote the open boundaries and the open boundary points where the tidal force is installed, respectively. The white dashed line indicates the position of the longitudinal section which will be shown in (b). (b) Bottom topography along the dashed line shown in (a) and the location of the undisturbed interface shown as the dashed line. (c) Prescribed spatial distribution of OBCs (Fourier coefficient a specially here) along the eastern boundary.

distributed and is prescribed with a spatial distribution which is constructed by the trigonometric function (see Figure 1(c)).

Some grids are randomly picked out from the 51×41 horizontal grids and treated as the observation positions (see Figure 1(a)). The forward model is run with the prescribed OBCs, and the model-generated results of the surface currents (i.e., currents in the upper layer) at these observation points are taken as the pseudo-observations (referred to as observations for brevity hereafter). Then, the OBCs can be optimized with the following steps.

Step 1. An initial value is chosen empirically and assigned to the unknown control variables.

Step 2. Perform the simulation by running the forward model, and the simulation results, mainly the state variables especially including the current velocity at observation points, are obtained. The value of the cost function is worked out in this step.

Step 3. The difference between simulation results and observations serves as the external force driving the adjoint model. Also, the adjoint variables in a period of M_2 tide are calculated through backward integration of the adjoint equations.

Step 4. Having known both the state and adjoint variables from Steps 2 and 3, respectively, the gradients with respect to control variables to be inverted are calculated by formula (16).

Step 5. Update the unknown control variables with a certain optimization algorithm.

Step 6. If some stopping criterion is satisfied, then stop and return the optimized control variables, otherwise; replace the initial value with the new control variables and go to Step 2.

With the procedure previous repeated, the OBCs will be optimized continuously, and the difference between simulated values and observations will be diminished. Meanwhile, the difference between the prescribed and the inverted OBCs will also be decreased. The initial value of a is taken as zero in each experiment. For all methods examined here, the total number of iterations is allowed to be 100 at most, and we report the cost function values, the gradient L_2 norms, and the inversion results obtained by each method. Note that the L-BFGS iteration may stop before it reaches the maximum number due to the termination criterion installed in the L-BFGS routine authored by Nocedal [51]. The number of corrections m used in the L-BFGS update is taken as 5. For the

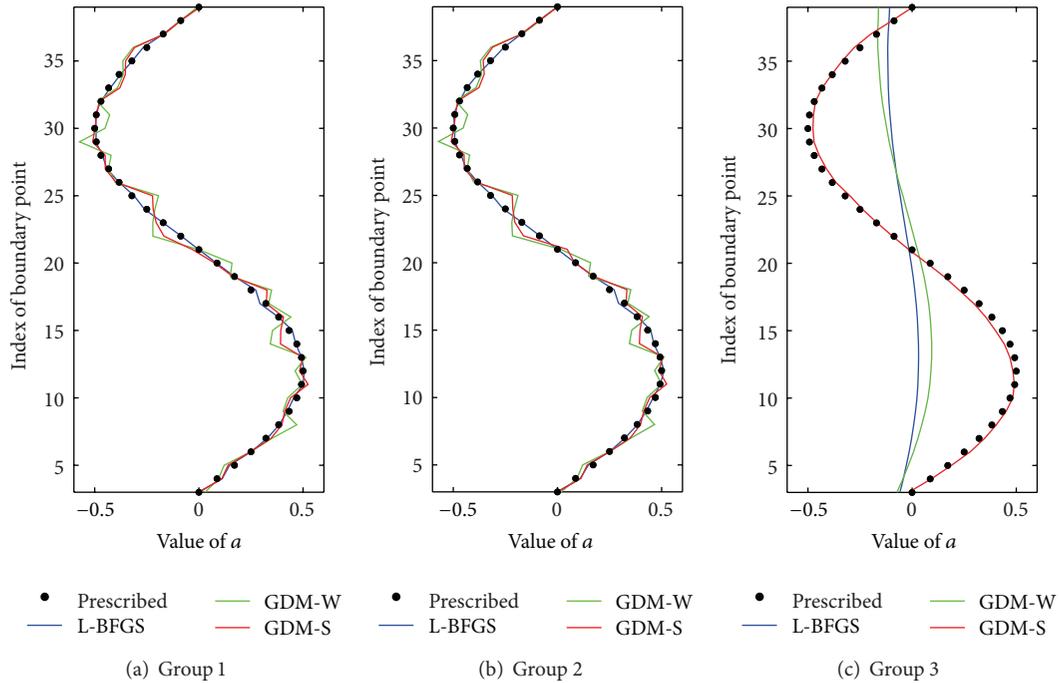


FIGURE 2: Inverted spatial distribution of OBCs (Fourier coefficient a specially here) compared with the prescribed one.

GDM-W, Wolfe's line search is carried out with parameters $c_1 = 10^{-4}$ and $c_2 = 0.9$. For the GDM-S, the constant α_0 is exponentially chosen to be 0.1.

4.2. Results. As shown in Figure 1(a), all observations can be divided into two groups: O1 which are closer to the eastern boundary and O2 which are further from the eastern boundary. Accordingly, the experiments are divided into three groups: Group 1 (both O1 and O2 are used), Group 2 (only O1 is used), Group 3 (only O2 is used). After assimilation, the inversion error (in the root mean square sense) and correlation coefficient between the prescribed and inverted OBCs are reported in Table 1, and the inverted OBCs are compared with the prescribed ones in Figure 2. All experiment results will be examined group by group as follows.

First, let us examine the performances of the three methods in the experiments of Group 1 in which observations of both O1 and O2 are assimilated. The results of Group 1 shown in Figure 2(a) and Table 1 indicate that with sufficient observations assimilated, the inversion results obtained by using all three methods are satisfactory in terms of the inverted OBC curve as well as the inversion error and the correlation coefficient. And especially, the solution obtained with the L-BFGS method is the closest to the exact solution (the prescribed OBCs). The variation of the cost function normalized by its initial value, that of the L_2 norm of the gradient of the cost function with respect to the OBCs, and that of the inversion error are plotted in Figures 3(a), 3(b), and 3(c), respectively, as a function of the iteration number. The optimization procedure clearly shows that among the three methods examined, the L-BFGS performs the best

and its convergence rate is much faster than those of the GD methods, which is consistent with the classical theories about the convergence rate of the quasi-Newton methods and GD method [31]. Now, let us pay more attention to the results obtained with GDM-W and GDM-S. The comparison between the two versions of GD method is rather interesting. At the beginning stage of the iteration, the GDM-W has a faster convergence rate than the GDM-S. After more than 20 iteration steps, the result of GDM-S begins to be better than that of the GDM-W. Then, after a neighborhood of the exact solution is reached, a slow convergence of GDM-S occurs in the following iterations, also with oscillations in both of the cost function and gradient norm which may be caused by the constant α_0 making the step length relatively overlarge when the solution is close to the exact one. In contrast, the GDM-W keeps the cost function and the gradient norm declining on the whole due to its step length satisfying the strong Wolfe conditions (21a) and (21b). It is worth noting that in the latter part of the assimilation procedure (after about 30 iteration steps when the oscillations in the cost function and in the gradient norm for the GDM-S appear (see Figures 3(a) and 3(b)), the convergence rate of the GDM-W is a little faster than that of the GDM-S. This leads to the final values of both the cost function and gradient norm obtained with the GDM-W slightly smaller than those obtained with the GDM-S. On the other hand, however, the comparison between the inversion errors for the GDM-W and GDM-S presents a different pattern. As shown in Figure 3(c), during the beginning several iterations, the descent rate for both GD methods is very rapid and the inversion error for the GDM-W has a faster descent rate than that for the GDM-S. This state continues until the 25th iteration step when the inversion

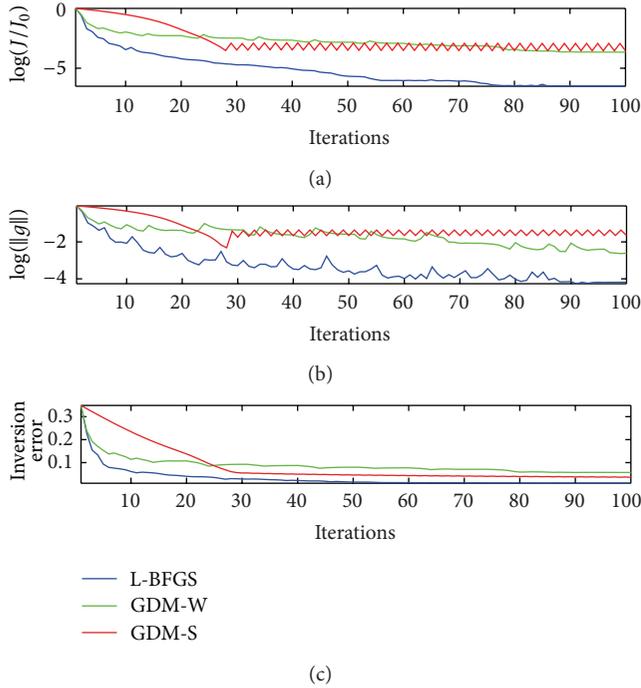


FIGURE 3: Optimization history for experiments of Group 1, about (a) the cost function normalized by its initial value J_0 , (b) the L_2 norm of the gradient of the cost function with respect to the OBCs, and (c) the inversion error (in the root mean square sense) between the inverted and prescribed OBCs.

error for the GDM-W begins to exceed that for the GDM-S, and these two methods almost begin to have the same slow descent rate of the inversion error. Finally, the inversion result obtained with the GDM-S is better than that obtained with the GDM-W (also see Figure 2(a) and Table 1).

For all experiments in Group 2, only the observations of O1 are assimilated. For this group, the variation of the cost function normalized by its initial value, that of the L_2 norm of the gradient of the cost function with respect to the OBCs, and that of the inversion error are plotted in Figures 4(a), 4(b), and 4(c), respectively, as a function of the iteration number. As we can see, the assimilation procedure of Group 2 shown in Figure 4 follows a similar pattern as that of Group 1 shown in Figure 3, although the number of observations for Group 2 is only half of that for Group 1. This suggests that removing the observations that are far from the unknown boundary may have very little effect upon the experiment results, whereas the observations that are closer to the eastern open boundary play a much more important role in the estimation of the OBCs in our model. Probably, this may be the main reason why we can obtain the satisfactory results in Group 1.

To sum up, from the previous results, we have seen that the L-BFGS method can hold its advantage over the other two versions of GD method in terms of both the optimization procedure and the final results providing the observations that are closer to the unknown boundary are included for assimilation, and the L-BFGS method is recommended in this

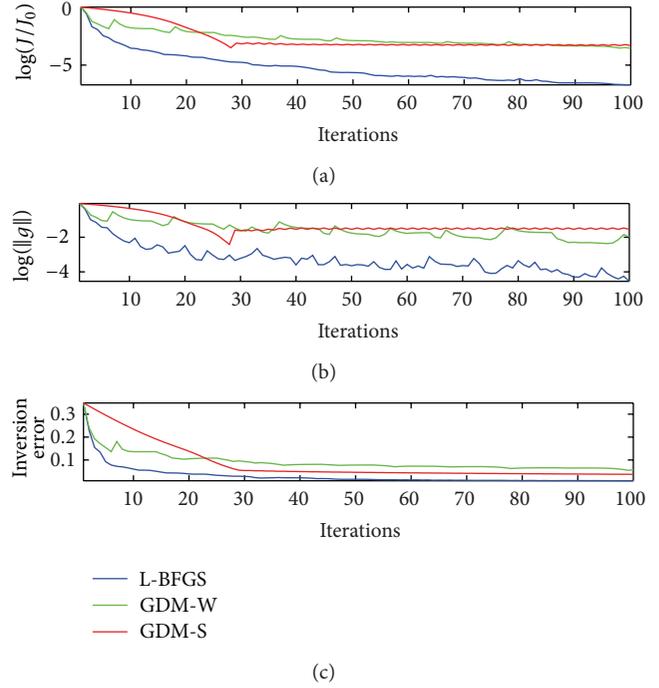


FIGURE 4: As in Figure 3 but for the experiments of Group 2.

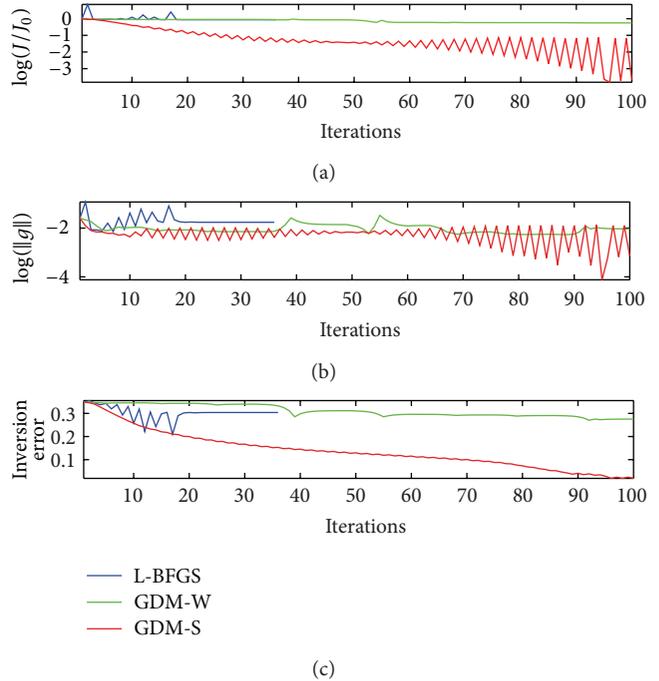


FIGURE 5: As in Figure 3 but for the experiments of Group 3.

case. Meanwhile, both results obtained with the two versions of GD method are also satisfactory. The comparison between the results obtained with the GDM-W and GDM-S indicates that the Wolfe's line search is effective in finding a reasonable step length that can achieve adequate reductions in the cost function and can guarantee a rapid rate of convergence at

TABLE 1: Inversion error (in the root mean square sense) and correlation coefficient between the inverted and prescribed OBCs.

Group	Inversion error			Correlation coefficient		
	L-BFGS	GDM-W	GDM-S	L-BFGS	GDM-W	GDM-S
1	0.0105	0.0571	0.0361	0.9995	0.9865	0.9947
2	0.0094	0.0566	0.0372	0.9996	0.9868	0.9943
3	0.3038	0.2753	0.0193	0.8935	0.8482	0.9986

the beginning of iteration. This allows the GDM-W to gain an advantage over the GDM-S in terms of the final values of the cost function and the gradient norm. Nevertheless, their difference is smaller than expected, and finally, the GDM-S excels the GDM-W in the inversion result which is an essential criterion testing whether the whole model is valid. Therefore, the GD method simplified with a constant step length is competitive to that with the Wolfe's line search in this case.

It is very interesting to find that when we come to the results of Group 3 which are dramatically different from those of Group 2 although the same number of observations is used in these two groups. In Group 3, only the observations O2 that are further from the unknown boundary are assimilated. For this group, the variation of the cost function normalized by its initial value, that of the L_2 norm of the gradient of the cost function with respect to the OBCs, and that of the inversion error are plotted in Figures 5(a), 5(b) and 5(c), respectively, as a function of the iteration number. From the optimization procedure shown in Figure 5, we can clearly see that, in this case, the L-BFGS method stops at the 36th iteration step and fails to converge, and the GDM-W has an unacceptable slow convergence rate, both leading to a relatively large discrepancy between the inverted and prescribed OBCs (see Figure 2(c) and Table 1), whereas the simple GDM-S performs the best and, more importantly, can still yield satisfactory inversion result (see Figure 2(c) and Table 1). The performance of GDM-S shown in Figure 5 is more interesting. After about 10 steps of iteration in the beginning, the cost function and the gradient norm begin to vary with oscillations which become increasingly larger as the iteration goes on, but meanwhile, their minimum values are getting smaller. At last when the maximum of 100 iterations is reached, the cost function and the gradient norm are reduced by more than 3 orders of magnitude and more than 1.5 orders of magnitude, respectively, compared with their initial values, which are comparable to the results of GDM-S in Groups 1 and 2. On the other hand, the variation of the inversion error for GDM-S also contains oscillations, but compared with the oscillations in the cost function and gradient norm (see Figures 5(a) and 5(b)), the ones in the inversion error are much smaller, and therefore the variation curve looks much smoother (see Figure 5(c)). This may be caused by the high complexity of the cost function in the control variable space if the observations to be assimilated are far from the unknown boundary, and this high complexity may also be a possible reason for the failure of the L-BFGS method as well as for the inefficiency of Wolfe's line search in Group 3. The results of Group 3 indicate that the far distance between the observations and the unknown boundary has almost

little effect on the validity and feasibility of the GDM-S, and therefore the simplified GD method is recommended in this case, instead of the L-BFGS method.

5. Summary and Conclusions

In this paper, based on the numerical internal tidal model constructed by Chen et al. [18], the practical performances of the L-BFGS method (given by Liu and Nocedal [32]) and two versions of GD method are compared and investigated computationally through a series of ideal experiments in which the OBCs are estimated by assimilating the interior observations with the adjoint method. The two GD methods include the normal one with Wolfe's line search for the step length and the simplified one with a fixed step length which should be predetermined by modeler. The cost function, the gradient norm, and the inversion result are reported and examined for each experiment.

According to the locations of the observations assimilated, all observations can be divided into two groups. Accordingly, all experiments are divided into three groups. In both of the first two groups, the observations that are closer to the unknown boundary are included for assimilation. In this case, as expected, the L-BFGS method has a remarkably better performance than the GD methods, not only in terms of the convergence rate but also in terms of the final result. This is consistent with the classical theory about the convergence properties of the L-BFGS method and the GD method. On the other hand, compared with the simplified GD method, the normal one with Wolfe's line search can really use fewer iteration steps to reach a satisfactory solution, but the values of the minimized cost function and the gradient norm are almost the same. Although great computational efforts have been made to implement Wolfe's line search to optimize the step length, the effect is much smaller than expected, and even the normal GD method finally yields a worse inversion result than the simplified one. In the third group of experiments, only the observations that are further from the unknown boundary are assimilated. In this case, the simplified GD method remains effective and yields satisfactory results, whereas the L-BFGS method fails to converge and the GD method with Wolfe's line search presents an unacceptable slow convergence rate. This demonstrates that in the practical application, the simplified GD method, which is controllable and easy to implement, is competitive to both the classical L-BFGS method and the normal GD method with Wolfe's line search. This suggests that when applied to the practical cases, the advanced L-BFGS algorithm and Wolfe's line search may still need to be improved, while how to take sufficient advantage of some simple methods is still

worth investigating. Nonetheless, the simplified GD method should not be ignored and should be regarded seriously as a choice, especially when the classical advanced optimization techniques fail or perform poorly. Furthermore, how to take sufficient advantage of some other simplified methods is also worth investigating.

Acknowledgment

The authors would like to deeply thank the editor and two anonymous reviewers for their constructive suggestions whereby their paper has been improved greatly. Partial support for this research was provided by the National Natural Science Foundation of China through Grant Nos. 41076006 and 41206001, the State Ministry of Science and Technology of China under Contract nos. 2008AA09A402, 2013AA091201, and 2013AA091202, the Ministry of Education's 111 Project through Grant B07036, and the Fundamental Research Funds for the Central Universities 201261006 and 201262007.

References

- [1] I. M. Navon, X. Zou, J. Derber, and J. Sela, "Variational data assimilation with an adiabatic version of the NMC spectral model," *Monthly Weather Review*, vol. 120, no. 7, pp. 1433–1446, 1992.
- [2] J. Zou, W. H. Hsieh, and I. M. Navon, "Sequential open-boundary control by data assimilation in a limited-area model," *Monthly Weather Review*, vol. 123, pp. 2899–2909, 1995.
- [3] D. L. T. Anderson, J. Sheinbaum, and K. Haines, "Data assimilation in ocean models," *Reports on Progress in Physics*, vol. 59, no. 10, pp. 1209–1266, 1996.
- [4] R. W. Lardner, A. H. Al-Rabeh, and N. Gunay, "Optimal estimation of parameters for a two-dimensional hydro-dynamical model of the Arabian gulf," *Journal of Geophysical Research*, vol. 98, no. C10, pp. 18229–18242, 1993.
- [5] U. Seiler, "Estimation of open boundary conditions with the adjoint method," *Journal of Geophysical Research*, vol. 98, no. C12, pp. 22855–22870, 1993.
- [6] P. G. J. ten Brummelhuis, A. W. Heemink, and H. F. P. van den Boogaard, "Identification of shallow sea models," *International Journal for Numerical Methods in Fluids*, vol. 17, no. 8, pp. 637–665, 1993.
- [7] A. W. Heemink, E. E. A. Moutaan, M. R. T. Roest, E. A. H. Vollebregt, K. B. Robaczewska, and M. Verlaan, "Inverse 3D shallow water flow modelling of the continental shelf," *Continental Shelf Research*, vol. 22, no. 3, pp. 465–484, 2002.
- [8] I. Shulman, "Local data assimilation in specification of open boundary conditions," *Journal of Atmospheric and Oceanic Technology*, vol. 14, no. 6, pp. 1409–1419, 1997.
- [9] I. Shulman, J. K. Lewis, A. F. Blumberg, and B. N. Kim, "Optimized boundary conditions and data assimilation with application to the M_2 tide in the yellow sea," *Journal of Atmospheric and Oceanic Technology*, vol. 15, no. 4, pp. 1066–1071, 1998.
- [10] V. Taillandier, V. Echevin, L. Mortier, and J. L. Devenon, "Controlling boundary conditions with a four-dimensional variational data-assimilation method in a non-stratified open coastal model," *Ocean Dynamics*, vol. 54, no. 2, pp. 284–298, 2004.
- [11] A. J. Zhang, E. Wei, and B. B. Parker, "Optimal estimation of tidal open boundary conditions using predicted tides and adjoint data assimilation technique," *Continental Shelf Research*, vol. 23, no. 11–13, pp. 1055–1070, 2003.
- [12] I. Y. Gejadze and G. J. M. Copeland, "Open boundary control problem for Navier-Stokes equations including a free surface: adjoint sensitivity analysis," *Computers & Mathematics with Applications*, vol. 52, no. 8–9, pp. 1243–1268, 2006.
- [13] J. C. Zhang and X. Q. Lu, "Parameter estimation for a three-dimensional numerical barotropic tidal model with adjoint method," *International Journal for Numerical Methods in Fluids*, vol. 57, no. 1, pp. 47–92, 2008.
- [14] J. C. Zhang and X. Q. Lu, "Inversion of three-dimensional tidal currents in marginal seas by assimilating satellite altimetry," *Computer Methods in Applied Mechanics and Engineering*, vol. 199, no. 49–52, pp. 3125–3136, 2010.
- [15] A. Z. Cao, Z. Guo, and X. Q. Lu, "Inversion of two-dimensional tidal open boundary conditions of M_2 constituent in the Bohai and Yellow Seas," *Chinese Journal of Oceanology and Limnology*, vol. 30, no. 5, pp. 868–875, 2012.
- [16] Z. Guo, A. Z. Cao, and X. Q. Lv, "Inverse estimation of open boundary conditions in the Bohai Sea," *Mathematical Problems in Engineering*, vol. 2012, Article ID 628061, 9 pages, 2012.
- [17] E. Kazantsev, "Boundary conditions control for a shallow-water model," *International Journal for Numerical Methods in Fluids*, vol. 68, no. 5, pp. 625–641, 2012.
- [18] H. B. Chen, C. B. Miao, and X. Q. Lv, "A three-dimensional numerical internal tidal model involving adjoint method," *International Journal for Numerical Methods in Fluids*, vol. 69, no. 10, pp. 1584–1613, 2012.
- [19] Z. Wang, I. M. Navon, X. Zou, and F. X. Le Dimet, "A truncated Newton optimization algorithm in meteorology applications with analytic Hessian/vector products," *Computational Optimization and Applications*, vol. 4, no. 3, pp. 241–262, 1995.
- [20] Z. Wang, K. Droegemeier, and L. White, "The adjoint Newton algorithm for large-scale unconstrained optimization in meteorology applications," *Computational Optimization and Applications*, vol. 10, no. 3, pp. 283–320, 1998.
- [21] X. Zou, I. M. Navon, M. Berger, K. H. Phua, T. Schlick, and F. X. Le Dimet, "Numerical experience with limited-memory quasi-Newton and truncated Newton methods," *SIAM Journal on Optimization*, vol. 3, no. 3, pp. 582–608, 1993.
- [22] X. Q. Lu and J. C. Zhang, "Numerical study on spatially varying bottom friction coefficient of a 2D tidal model with adjoint method," *Continental Shelf Research*, vol. 26, no. 16, pp. 1905–1923, 2006.
- [23] J. C. Zhang, X. Q. Lu, P. Wang, and Y. P. Wang, "Study on linear and nonlinear bottom friction parameterizations for regional tidal models using data assimilation," *Continental Shelf Research*, vol. 31, no. 6, pp. 555–573, 2011.
- [24] A. F. Bennett and P. C. McIntosh, "Open ocean modeling as an inverse problem: tidal theory," *Journal of Physical Oceanography*, vol. 12, no. 10, pp. 1004–1018, 1982.
- [25] F. X. Le Dimet and O. Talagrand, "Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects," *Tellus A*, vol. 38, no. 2, pp. 97–110, 1986.
- [26] E. Courtier and O. Talagrand, "Variational assimilation of meteorological observations with the adjoint equations—part 2: numerical results," *Quarterly Journal of the Royal Meteorological Society*, vol. 113, no. 478, pp. 1329–1347, 1987.

- [27] X. Zou, I. M. Navon, and F. X. Le Dimet, "Incomplete observations and control of gravity waves in variational data assimilation," *Tellus A*, vol. 44, no. 4, pp. 273–296, 1992.
- [28] M. Ghil and P. Malanotte-Rizzoli, "Data assimilation in meteorology and oceanography," in *Advances in Geophysics*, B. Saltzman, Ed., vol. 33, pp. 141–266, Academic Press, New York, NY, USA, 1991.
- [29] I. M. Navon, "Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography," *Dynamics of Atmospheres and Oceans*, vol. 27, no. 1–4, pp. 55–79, 1998.
- [30] L. S. Yu and J. J. O'Brien, "On the initial condition parameter estimation," *Journal of Physical Oceanography*, vol. 22, pp. 1361–1364, 1992.
- [31] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer Series in Operations Research and Financial Engineering, Springer, New York, NY, USA, 2nd edition, 2006.
- [32] D. C. Liu and J. Nocedal, "On the limited memory BFGS method for large scale optimization," *Mathematical Programming*, vol. 45, no. 3, pp. 503–528, 1989.
- [33] A. K. Alekseev, I. M. Navon, and J. L. Steward, "Comparison of advanced large-scale minimization algorithms for the solution of inverse ill-posed problems," *Optimization Methods & Software*, vol. 24, no. 1, pp. 63–87, 2009.
- [34] J. Nocedal, A. Sartenaer, and C. Zhu, "On the behavior of the gradient norm in the steepest descent method," *Computational Optimization and Applications*, vol. 22, no. 1, pp. 5–35, 2002.
- [35] M. N. Vrahatis, G. S. Androulakis, J. N. Lambrinos, and G. D. Magoulas, "A class of gradient unconstrained minimization algorithms with adaptive stepsize," *Journal of Computational and Applied Mathematics*, vol. 114, no. 2, pp. 367–386, 2000.
- [36] P. Wolfe, "Convergence conditions for ascent methods," *SIAM Review*, vol. 11, pp. 226–235, 1969.
- [37] P. Wolfe, "Convergence conditions for ascent methods. II. Some corrections," *SIAM Review*, vol. 13, pp. 185–188, 1971.
- [38] G. Zhou, "A descent algorithm without line search for unconstrained optimization," *Applied Mathematics and Computation*, vol. 215, no. 7, pp. 2528–2533, 2009.
- [39] Z.-J. Shi and J. Shen, "Convergence of descent method without line search," *Applied Mathematics and Computation*, vol. 167, no. 1, pp. 94–107, 2005.
- [40] R. A. Flather, "A tidal model of the north-west European continental shelf," *Memoires de la Societe Royale des Sciences de Liege*, vol. 6, no. 10, pp. 141–164, 1976.
- [41] C. B. Miao, H. B. Chen, and X. Q. Lu, "An isopycnic-coordinate internal tide model and its application to the South China Sea," *Chinese Journal of Oceanology and Limnology*, vol. 29, no. 6, pp. 1339–1356, 2011.
- [42] E. D. Palma and R. P. Matano, "On the implementation of passive open boundary conditions for a general circulation model: the barotropic mode," *Journal of Geophysical Research C*, vol. 103, no. 1, pp. 1319–1341, 1998.
- [43] J. Nycander and K. Ds, "Open boundary conditions for barotropic waves," *Journal of Geophysical Research*, vol. 108, no. C5, 2002.
- [44] J. J. Leendertse, R. C. Alexander, and S. K. Liu, "A three dimensional model for estuaries and coastal seas: volume I, principles of computations," Report R-1417-OWRR, Rand Corporation, Santa Monica, Calif, USA, 1973.
- [45] J. O. Backhaus, "A three-dimensional model for the simulation of shelf sea dynamics," *Deutsche Hydrographische Zeitschrift*, vol. 38, no. 4, pp. 165–187, 1985.
- [46] V. Casulli, "A semi-implicit finite difference method for the non-hydrostatic free-surface flow," *International Journal for Numerical Methods in Fluids*, vol. 30, no. 4, pp. 425–440, 1999.
- [47] H. L. Yuan and C. H. Wu, "An implicit three-dimensional fully non-hydrostatic model for free-surface flows," *International Journal for Numerical Methods in Fluids*, vol. 46, no. 7, pp. 709–733, 2004.
- [48] P. J. Roache, *Fundamentals of Computational Fluid Dynamics*, Hermosa Publishers Press, Albuquerque, NM, USA, 1982.
- [49] G. Fairweather and I. M. Navon, "A linear ADI method for the shallow-water equations," *Journal of Computational Physics*, vol. 37, no. 1, pp. 1–18, 1980.
- [50] T. J. Simons, "Verification of numerical models of Lake Ontario—part I: circulation in spring and early summer," *Journal of Physical Oceanography*, vol. 4, no. 4, pp. 507–523, 1974.
- [51] L. J. Nocedal, "BFGS subroutine, software for large-scale unconstrained optimization," <http://www.ece.northwestern.edu/~nocedal/lbfgs.html>.
- [52] H. B. Chen, X. Q. Lv, and Y. S. Qiao, "Application of gradient descent method to the sedimentary grain-size distribution fitting," *Journal of Computational and Applied Mathematics*, vol. 233, no. 4, pp. 1128–1138, 2009.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

