Research Article

Inversion Study of Vertical Eddy Viscosity Coefficient Based on an Internal Tidal Model with the Adjoint Method

Guangzhen Jin,¹ Qiang Liu,² and Xianqing Lv¹

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

²College of Engineering, Ocean University of China, Qingdao 266100, China

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

Received 28 March 2014; Revised 17 August 2014; Accepted 18 August 2014

Academic Editor: Fatih Yaman

Copyright © 2015 Guangzhen Jin 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 isopycnic-coordinate internal tidal model with the adjoint method, the inversion of spatially varying vertical eddy viscosity coefficient (VEVC) is studied in two groups of numerical experiments. In Group One, the influences of independent point schemes (IPs) exerting on parameter inversion are discussed. Results demonstrate that the VEVCs can be inverted successfully with IPs and the model has the best performance with the optimal IPs. Using the optimal IPs obtained in Group One, the inversions of VEVCs on two different Gaussian bottom topographies are carried out in Group Two. In addition, performances of two optimization methods of which one is the limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) method and the other is a simplified gradient descent method (GDM-S) are also investigated. Results of the experiments indicate that this adjoint model is capable to invert the VEVC with spatially distribution, no matter which optimization method is taken. The L-BFGS method has a better performance in terms of the convergence rate and the inversion results. In general, the L-BFGS method is a more effective and efficient optimization method than the GDM-S.

1. Introduction

Internal tide, which is the internal wave of tidal frequency, is a ubiquitous phenomenon in the oceans. Rattray [1], Baines [2], Bell [3], Baines [4], Craig [5], Gerkema [6], and Llewellyn Smith and Young [7] have developed theoretical models and obtained some analytical solutions of internal tide on ideal topographies. These theoretical models helped them investigate the generation and propagation of internal tide. Although great progress has been made on the internal tide theory and some analytical works were carried out, only a small amount of solutions can be provided, due to the complexity of the problems. For this reason, quantitative analysis with practical significance still has to rely on the combination of numerical simulation, theoretical analysis, experiment, and observation. Numerical simulation is an effective method in marine research and has been widely used in the internal tide research. Kang et al. [8] investigated the M_2 internal tide near Hawaii with a two-dimensional, two-layered numerical model and confirmed

that the internal tide was generated by barotropic forcing at the Hawaiian Ridge and propagated in north-northeast and south-southwest directions. Based on a high-precision three-dimensional Princeton Ocean Model (POM), Niwa and Hibiya [9] obtained the distribution of the M_2 internal tide in the Pacific Ocean using the TOPEX/Poseidon (T/P) satellite data. Cummins et al. [10] simulated the generation and propagation of internal tides near the Aleutian Ridge using T/P altimeter data. The comparison between the altimeter data and their model results showed good agreement for the phase, which also provided evidence for wave fraction near the Aleutian Ridge. With a three-dimensional POM, Niwa and Hibiya [11] investigated the distribution and the energy of the M_2 internal tide around the continental shelf edge in the East China Sea. Their numerical experiment results indicated that M_2 internal tides are effectively generated over prominent topographies such as sea ridges, island chains, and straits. Jan et al. [12] modified the POM to study the generation of the K_1 internal tide and its influence on surface tide in the South China Sea. The conversion from

the barotropic energy to the baroclinic energy over topographic ridges in the Luzon Strait was also estimated.

Determination of the vertical eddy viscosity coefficient (VEVC), which describes the vertical mixing in the ocean, plays an important role in the study of energy exchange and material transportation. The VEVC is regularly regarded as a constant in numerical models. Schemes to determine the VEVC mainly include the Prandtl mixing-length hypothesis model, the k - ϵ model, the Pacanowski-Philander mixing model [13], and some turbulent closure models that are more complicated. Many studies have been carried out to investigate the variation of the VEVC [14–18]. All these mentioned studies indicate that due to different intensions of the vertical mixing in sea water, the VEVC should not be treated as a constant but a parameter with spatial distribution.

Satellite remote sensing technology and other related technologies provide us with a large number of data. Thus, it is one of the most important missions in physical oceanography to make use of the data efficiently and precisely as well as to combine the observation data with present numerical models. Indeed, data assimilation with the adjoint method provides an effective access to these missions. The use of the adjoint method in marine science can be traced back to 1980s. The adjoint model is capable of optimizing control parameters in numerical simulation. Bennett and McIntosh [19] applied the weak constraint thought to solve the tidal problem and the geostrophic-flow problem. Yu and O'Brien [20] assimilated both meteorological and oceanographic data into an oceanic Ekman layer model and deduced the unknown boundary condition, the unknown vertical eddy viscosity, and the current field. Based on a tidal model with a two-level leapfrog method, Lardner [21] inverted the open boundary conditions in three-test problems. Seiler [22] used the adjoint method to assimilate observations into a quasi-geostrophic ocean model and estimated the lateral boundary values in ideal experiments. Navon [23] wrote a summary of the parameter estimation in meteorology and oceanography in the view of applications with four-dimensional variational data assimilation techniques. Using an automatic differentiation compiler, Ayoub [24] constructed the adjoint model of the Massachusetts Institute of Technology Ocean General Circulation Model and inverted the open boundary conditions in the North Atlantic. Zhang and Lu [25] developed a three-dimensional nonlinear numerical tidal model with the adjoint method and designed several numerical experiments to estimate three kinds of parameters including the open boundary conditions, the bottom friction coefficients, and the vertical eddy viscosity coefficients. Zhang and Lu [26] employed a two-dimensional tidal model to study the inversion of the bottom friction coefficients in the Bohai Sea and the Yellow Sea with the adjoint method. Chen et al. [27] constructed a three-dimensional internal tidal model with the adjoint method and estimated six different kinds of open boundary conditions on fourteen types of topography. Based on a tidal model, Zhang and Chen [28] carried out several semiidealized experiments to estimate the partly and fully spatial varying open boundary conditions. Cao et al. [29] investigated the inversion of open boundary conditions with

a three-dimensional internal tidal model and simulated the M_2 internal tide around Hawaii by assimilating T/P data.

There are two main objectives of this paper. One is to study the inversion of the VEVC with an internal tidal model and the adjoint method. According to the introductions above, a lot of studies have been carried out to investigate the inversion of the control parameters of internal tide such as the open boundary condition [29, 30] and the bottom friction condition [31, 32]. However, few works are found to study the inversion of VEVC. Since VEVC is a decisive factor to describe the vertical mixing in the ocean, it is necessary to pay attention to the inversion of the VEVC. The other objective is to make a computational investigation on the performance of the gradient descent method and the limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) method for the inversion of VEVCs based on the model constructed by Chen et al. [27]. Both of the methods do not require any evaluations of the Hessian matrices but gradient vectors and, thus, are computationally feasible. Chen et al. [30] have made a comparative study on several optimization methods but it is on the inversion of the open boundary conditions which is a one-dimension case. The feasibility of these optimization methods for two-dimensional case such as the inversion of the VEVC needs further studies.

Two groups of numerical experiments are carried out to study the inversion of spatially varying VEVCs based on an isopycnic-coordinate internal tidal model with the adjoint method. In Group One, the influences of independent point schemes (IPSs) exerting on parameter inversion are discussed. Group Two investigates the inversions of VEVCs on two different Gaussian bottom topographies and the performances of two optimization methods which are the GDM-S and the L-BFGS methods.

This paper is organized as follows. Section 2 briefly introduces the adjoint tidal model and the methodology. Two optimization methods, including the GDM-S and the L-BFGS methods, are described in Section 3. Section 4 presents design and process of the experiments in detail. Results of the experiments are discussed in Section 5. Finally, we make a summary and draw some conclusions in Section 6.

2. Numerical Model Introduction

An isopycnic-coordinate internal tidal model with adjoint assimilation method is employed in this paper. There are two parts in the internal tide model. One is forward model with the governing equations and the other is adjoint model with the adjoint equations. The two models are used to simulate the internal tide and to optimize the control variables, respectively. Chen et al. [27] had introduced the two parts in great detail and tested the reasonability and feasibility of the model. The formulation will not be presented in this paper. The derivation of VEVC adjustment, introduction of the two optimization method, test of the adjoint method, and the independent point scheme (IPS) are described in details in this part.

2.1. Test of the Adjoint Method. According to the equations and derivations of Chen et al. [27], the formula to invert the VECV can be derived. The first derivative of Lagrangian function with respect to VECV is obtained as follows:

$$\frac{\partial L}{\partial A_{vi,j,k}} = 0, \quad (1)$$

where $A_{vi,j,k}$ is the value of VECV at grid (i, j) in the k th layer. The gradient of cost functions with respect to the VECV in the grid (i, j, k) can be deduced as follows:

$$\begin{aligned} \frac{\partial J}{\partial A_{vi,j,k}} &= (\rho_k + \rho_{k+1}) \\ &\times \sum_n \left[\frac{(u_{i,j,k}^n - u_{i,j,k+1}^n)(u_{ai,j,k}^n - u_{ai,j,k+1}^n)}{h_{i,j,k} + h_{i+1,j,k} + h_{i,j,k+1} + h_{i+1,j,k+1}} \right. \\ &\quad \left. + \frac{(u_{i-1,j,k}^n - u_{i-1,j,k+1}^n)(u_{ai-1,j,k}^n - u_{ai-1,j,k+1}^n)}{h_{i-1,j,k} + h_{i,j,k} + h_{i-1,j,k+1} + h_{i,j,k+1}} \right] \\ &+ (\rho_k + \rho_{k+1}) \\ &\times \sum_n \left[\frac{(v_{i,j,k}^n - v_{i,j,k+1}^n)(v_{ai,j,k}^n - v_{ai,j,k+1}^n)}{h_{i,j,k} + h_{i+1,j,k} + h_{i,j,k+1} + h_{i+1,j,k+1}} \right. \\ &\quad \left. + \frac{(v_{i-1,j,k}^n - v_{i-1,j,k+1}^n)(v_{ai-1,j,k}^n - v_{ai-1,j,k+1}^n)}{h_{i-1,j,k} + h_{i,j,k} + h_{i-1,j,k+1} + h_{i,j,k+1}} \right], \quad (2) \end{aligned}$$

where ρ_k is the potential density in the k th layer, $u_{i,j,k}^n$ and $v_{i,j,k}^n$ are horizontal velocities at the n th time step, $u_{ai,j,k}^n$ and $v_{ai,j,k}^n$ are the adjoint variables of $u_{i,j,k}^n$ and $v_{i,j,k}^n$, respectively, and $h_{i,j,k}$ is the initial thickness of the k th layer. The detailed derivation of (2) is presented in the appendix.

Accurately programming the adjoint in such problems as the present one is quite tricky and experience has shown that it is essential to check the accuracy of the adjoint computation before proceeding with the minimization runs [33]. The correctness of the adjoint method is verified in this section. Take the first-order term of a Taylor expansion for the cost function and we obtain the following equation:

$$J(\mathbf{p} + \alpha \mathbf{U}) = J(\mathbf{p}) + \alpha \mathbf{U} \cdot \mathbf{G}(\mathbf{p}) + O(\alpha^2). \quad (3)$$

Here, $\mathbf{p}$ is a general point of the control variable, $\mathbf{G}(\mathbf{p}) = \nabla J(\mathbf{p})$ is the computed gradient, and $\mathbf{U}$ is an arbitrary unit vector in the parameter space. Based on (3) a function of α can be written as follows:

$$\Phi(\alpha) = \frac{J(\mathbf{p} + \alpha \mathbf{U}) - J(\mathbf{p})}{\alpha \mathbf{U} \cdot \mathbf{G}(\mathbf{p})}, \quad (4)$$

where α is a small real number that is not equal to zero. If the adjoint methodology is correct, it is supposed that

$\lim_{\alpha \rightarrow 0} \Phi(\alpha) = 1$ according to (4). In this paper, the VECV variable A_v is treated as $\mathbf{p}$ and test of the adjoint method is based on (4).

In order to test the accuracy of the adjoint method, two experiments are carried out in which two different types of $\mathbf{U}$ are used. The different vector directions are $U_1 = G(p)/|G(p)|$ and $U_2 = (\sqrt{1/N}, \sqrt{1/N}, \dots, \sqrt{1/N})$, respectively.

Figure 1 indicates the trends of $\Phi(\alpha)$ as α approaches to 0. It is clear that in both experiments when α is less than 10^{-3} , values of Φ (solid lines) are both very close to 1 (dashed lines). Equation $\lim_{\alpha \rightarrow 0} \Phi(\alpha) = 1$ is proved and the correctness of gradient computed in the adjoint model is verified.

2.2. Independent Point Scheme. The available observation data may not be sufficient and control parameters to be determined may be excessive in practice. That may cause ill-posedness of the inversion problem. Richardson and Panchang [34] first noted that if an adjoint method is applied, when there is a big error in data, the solution will be unstable and not unique. Many researchers have made progress in solving this problem. A lot of work [26, 27, 30, 32] have been done to prove the capability and feasibility of the independent point scheme (IPS) in solving ill-posed problems of inversion.

In this paper, the IPS is used to optimize the control parameter. The basic idea of IPS is as follows: some grids (e.g., (ii, jj)) are selected as the independent points; it is assumed that $P_{ii,jj}$ represents the value of VECV in grid (ii, jj) , so values of VECV in all grids $p_{i,j}$ can be calculated from $P_{ii,jj}$ with linear interpolation method. The computing formula is given as follows:

$$p_{i,j} = \frac{\sum_{ii,jj} W_{i,j,ii,jj} P_{ii,jj}}{\sum_{ii,jj} W_{i,j,ii,jj}}, \quad (5)$$

where $W_{i,j,ii,jj}$ is the weight coefficient of the Cressman form [35]:

$$W_{i,j,ii,jj} = \frac{R^2 - r_{i,j,ii,jj}^2}{R^2 + r_{i,j,ii,jj}^2}, \quad (6)$$

where $r_{i,j,ii,jj}$ is the center distance between (i, j) and (ii, jj) and R is the influence radius. According to (5), the gradient of J with respect to $p_{i,j}$ can be written as

$$\frac{\partial J}{\partial p_{ii,jj}} = \sum_{i,j} W_{i,j,ii,jj} \frac{\partial J}{\partial p_{i,j}} \sum_{i,j} W_{i,j,ii,jj}, \quad (7)$$

where $\partial J / \partial p_{i,j}$ is the gradient of J with respect to the VECV at the grid (i, j) and can be calculated using formula (2).

According to Section 2.1, the correctness of the gradient with respect to the independent points should be tested. Based on (5), the perturbations ($\alpha \mathbf{U}$) applied to the independent points have a linear impact on those applied to the nonindependent points. The convergences for the nonindependent points will remain the same after linear transformation for independent points.

The values of VECV at the independent grids can be calculated inside the model and values at other grids are

FIGURE 1: Variation of $\Phi(\alpha)$ with respect to α .

gained through interpolation using (5). Then the spatial distribution of VEC in the entire area is obtained.

3. Optimization Algorithms

There have been many large-scale optimization methods to solve the minimization problem [36]. Four main methods are line search method (e.g., Wolfe and Goldstein), trust-region method, conjugate gradient method (e.g., Fletcher-Reeves), and quasi-newton method (e.g., BFGS and L-BFGS). However, the number of studies discussing the performances of various optimization methods in the meteorological and oceanographic application is still relatively small [30]. Line search method requires repeated computations for the cost function and the gradient. Therefore it spends too much computation resource during numerical simulations especially for physical oceanography. Besides, line search methods may fail in some cases [37]. The L-BFGS method is commonly used to solve large-scale problems in oceanography and meteorology [30, 38, 39]. Chen et al. [30] have made a computational investigation on the performances of the L-BFGS method and two versions of gradient descent method with an internal tide model. In their work, a simplified gradient descent method is applied which is able to avoid too much computation. The step length is chosen according to the experience of the modeler. Compared with other methods, there are two main advantages of this plan. One is the less computation resource usage and the other is the more controllable optimization process. Many research papers have proved the feasibility of this method [25–27, 31, 32].

Generally speaking, numerical methods to solve the minimization problems have the similar iterative formula as follows:

$$A_v^{n+1} = A_v^n + \alpha^n d^n, \quad (8)$$

where A_v^n and A_v^{n+1} are the priori and adjusted values of VEC in the n th iteration, respectively and α^n and d^n represent the iteration step length and the search direction, respectively. There are many methods to determine the search

direction d_n . Two different optimization methods employed in this paper are the GDM-S and the L-BFGS methods.

3.1. Gradient Descent Method (GDM). The GDM is a simple and feasible method to define the search direction as follows:

$$d_n = -g_n = -\frac{(\partial J / A_{vi,j,k})_n}{\|(\partial J / A_{vi,j,k})_n\|}, \quad (9)$$

where i , j , and k are the zonal index, the meridional index, and the layer index of the calculation grid, respectively; n represents the step of iteration. $\|(\partial J / A_{vi,j,k})_n\|$ is the L_2 norm of the gradient of the cost function with respect to the VEC in the n th iteration.

In the GDM-S, the step length α is chosen to be a constant according to the experience of the modeler. We have surveyed the performances of different values of α and take 0.006 as the best choice. The optimized value of VEC is obtained after the VEC in every grid (i, j, k) is adjusted according to (8) and (9).

3.2. L-BFGS Method. L-BFGS is an optimization algorithm in the family of quasi-Newton methods that approximates the BFGS algorithm using a limited amount of computer memory. This method is first described in the work of Nocedal [40], where it is called the SQN method. It is a popular algorithm for parameter estimation in machine learning [41, 42]. Due to its resulting linear memory requirement, the L-BFGS method is particularly well suited for optimization problems with a large number of variables.

It requires the search direction to be

$$d_k = -H_k g_k, \quad (10)$$

where

$$\begin{aligned}
 H_{k+1} = & \left(V_{k-1}^T \cdots V_{k-m}^T \right) H_0 \left(V_{k-m} \cdots V_{k-1} \right) \\
 & + \rho_{k-m} \left(V_k^T \cdots V_{k-m+1}^T \right) s_{k-m} s_{k-m}^T \left(V_{k-m+1} \cdots V_k \right) \\
 & + \rho_{k-m} \left(V_k^T \cdots V_{k-m+1}^T \right) s_{k-m} s_{k-m}^T \left(V_{k-m+1} \cdots V_k \right) \\
 & + \rho_{k-m} \left(V_k^T \cdots V_{k-m+1}^T \right) s_{k-m} s_{k-m}^T \left(V_{k-m+1} \cdots V_k \right) \\
 & \vdots \\
 & + \rho_k s_k s_k^T.
 \end{aligned} \tag{11}$$

Note that g_k is the simplification of $\nabla J(p_k)$ for convenience. Here, $\rho_k = 1/s_k y_k^T$, $V_k = 1 - \rho_k y_k s_k^T$, $s_k = p_{k+1} - p_k = \alpha_k d_k$, and $y_k = g_{k+1} - g_k$. Many studies have shown that typically $3 \leq m \leq 7$, where $m > 7$ does not improve the performance of L-BFGS [43]. So the number of corrections m used in the L-BFGS update of this paper is taken as 5 [44]. The version of L-BFGS used in this paper is described in Liu and Nocedal [44] and the Fortran codes are authorized by Nocedal [45].

4. Design of Experiments

All the experiments in this paper are implemented in an ideal regional area from 116°E to 124.17°E and from 18°N to 23.17°N with in mind the practical sea area located around the Luzon strait. The horizontal resolution is $10' \times 10'$ and there are totally 49×31 grids in the area. The maximum depth is set to be 1000 meters. The horizontal eddy coefficient is chosen to be $A_h = 1000$ and the bottom friction coefficient is taken as $\kappa = 0.003$. The Coriolis coefficient is taken as the local value. Only M_2 tide is considered and its angular frequency is $1.41 \times 10^{-4} \text{ s}^{-1}$. The whole-time step is 496.86 s which is 1/90 of the period of M_2 tide. All the four boundaries are open boundaries and boundary conditions are set to be the local water levels in the Flather form.

Eastern and western boundaries:

$$\zeta - \zeta' = \pm \sqrt{\left(1 - \frac{f^2}{\omega^2}\right)} \frac{H}{g} (U - \bar{U}), \tag{12}$$

positive in eastern boundary and negative in western boundary.

North and south boundaries:

$$\zeta - \zeta' = \pm \sqrt{\left(1 - \frac{f^2}{\omega^2}\right)} \frac{H}{g} (V - \bar{V}), \tag{13}$$

positive in north boundary and negative in south boundary.

ζ is the surface elevation above the undisturbed sea level. U and V are the zonal current velocity and the meridional current velocity, respectively. ζ' , $\bar{U}$, and $\bar{V}$ are the surface elevation and the zonal and meridional current velocity relating to the boundary barotropic tidal force, respectively.

f is the Coriolis coefficient and ω is the tidal frequency of M_2 tide. H is the local water depth and g is the acceleration of gravity.

Similar as Chen et al. [27], the M_2 tidal force at the n th time step is subject to

$$\tilde{\zeta}^{t2} = a_\zeta \cos(\omega n \Delta t) + b_\zeta \sin(\omega n \Delta t), \tag{14}$$

the Fourier coefficients a_ζ at four open boundaries are set to be 0 and b_ζ are set to be 1 (−1) at the north and west (south and east) boundaries.

The T/P altimeter data is widely spread throughout the ocean and it can be used to invert VEV. In this work, we pick 89 calculating points based on the distribution features of T/P altimeter observation as the observation points (Figure 2).

Two kinds of topographies are tested in this paper and they are generated based on the two formulas in (15), respectively (Figure 3). The sea water in the computing area is divided into two layers. The thicknesses of each layers are, respectively, $h_1 = 200 \text{ m}$ and $h_2 = 800 \text{ m}$. The potential densities of corresponding layer are $\rho_1 = 1021 \text{ (kg/m}^{-3}\text{)}$ and $\rho_2 = 1024 \text{ (kg/m}^{-3}\text{)}$, respectively. Consider

$$\begin{aligned}
 H_A = & h_0 \exp\left(-\frac{(x-x_0)^2 + (y-y_0)^2}{2L^2}\right) \\
 & - \text{MaxDepth}, \\
 H_B = & h_0 \left(1 - \exp\left(-\frac{(x-x_0)^2 + (y-y_0)^2}{2L^2}\right)\right) \\
 & - \text{MaxDepth},
 \end{aligned} \tag{15}$$

where h_0 is the height of the topography and MaxDepth is the depth of the water.

For each experiment, the optimization of the VEV can be implemented with the following steps.

Step 1. The prescribed VEV is given and the forward model is run. The whole simulation time is 20 period of the M_2 tide in order to obtain a stable simulation result. The water elevations in the observation positions are treated as the “pseudoobservations.”

Step 2. Initial value of the control parameter (VEVC) is given and forward model is run to get the simulated results of all the state variables such as current velocity and water elevation. The value of cost function J is calculated.

Step 3. Difference between the simulated elevation and the “pseudoobservation” plays as the external force of the adjoint model. Via backward integrating the adjoint equations in a period of the M_2 tide values of the adjoint variables are obtained.

Step 4. Using formula (2), along with the state variables and the adjoint variables obtained in Steps 2 and 3, the gradient of cost function with respect to VEV is calculated.

FIGURE 2: Planform of topography (e.g., topography A) and locations of the observations (white dots).

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

Step 6. If the stopping criterion of iteration is reached, bring the iteration to an end and return the optimized parameter. Otherwise, update all the parameters and go back to Step 2.

In the experiments of this paper, all initial values of VEC are set to 0.005 and the total number of iterations is allowed to be 100 at most. The chosen convergence criterion is that the last two values of the cost function are sufficiently close, which is defined by

$$|J_{\text{end}} - J_{\text{end}-1}| < 10^{-9}, \quad (16)$$

where J_{end} and $J_{\text{end}-1}$ are the last and the second last values of the cost function, respectively.

Two groups of numerical experiments are carried out: the influence of IPSs on the inversion of VEC is studied in Group One; in Group Two the ability of this internal tide model to invert different kinds of VEC with spatial distribution is examined. Two kinds of spatial distribution of VEC are prescribed and given in Figure 4. For both distributions, the VEC value ranges from $3 \times 10^{-3} \text{ m}^2/\text{s}$ to $7 \times 10^{-3} \text{ m}^2/\text{s}$.

In Group One, nine experiments are carried out to discuss the influence of IPS on the inversion of VEC. The distance between independent points (IP) ranges from $20'$ (length of 2 grids) to $100'$ (length of 10 grids) and details are listed in Table 1. Topography A is applied and distribution (a) is chosen to be the prescribed spatial distribution of VEC.

In Group Two, four numerical experiments (NEs) are carried out which are numbered as NE1~NE4, respectively. Each experiment is implemented with GDM-S and the L-BFGS methods. Information of topographies and prescribed VECs for all NEs is listed in Table 2. The IPSs are the optimal schemes obtained in Group One. Other parameters are exactly the same as Group One.

All the experiments in Group One and Group Two are carried out following Steps 1 to 6 described in Section 4. Version of L-BFGS method used in this paper is from the work of Liu and Nocedal [44].

TABLE 1: Settings of independent point schemes in Group One.

IPS	1	2	3	4	5	6	7	8	9
Number of IPs	360	160	96	60	40	35	24	24	15
Distance between IPs ($'$)	20	30	40	50	60	70	80	90	100

TABLE 2: Topographies and distributions of VEC in Group Two.

Experiment	Topography	Distribution
NE1	A	a
NE2	A	b
NE3	B	a
NE4	B	b

5. Experiment Results and Discussions

5.1. Results and Discussions of Group One. Figure 5 illustrates the relationships between mean absolute errors (MAE, which reflects the error between the inversion result and the given VEC) and the distance between independent points in Group One.

As is shown in Figure 5, MAEs with different optimization methods vary as the IPSs change. The minimum MAEs with the two methods are not the same (dashed lines). The minimum MAE with the GDM-S is $2.66 \times 10^{-4} \text{ m}^2/\text{s}$ while that with the L-BFGS method reaches $9.14 \times 10^{-5} \text{ m}^2/\text{s}$. The corresponding IP distances are $90'$ (length of 9 grids, GDM-S) and $30'$ (length of 3 grids, L-BFGS), respectively. According to the MAEs illustrated in Figure 5, the optimal IPSs of the two methods are selected as IPS8 (GDM-S) and IPS 2 (L-BFGS).

5.2. Results and Discussions of Group Two. With the respective optimal IPSs and the iteration process in Section 4, the VECs are inverted with GDM-S method (I) and L-BFGS method (II), respectively, and the inversion results are shown in Figure 6.

Comparison of the inversion results with the prescribed VECs indicates that all the given spatial distributions of VEC are successfully inverted after 100 iteration steps. The main features of all distributions can be recovered very well. Surfaces of the inverted VEC with the L-BFGS are much smoother than those with the GDM-S. Compared against the inversion results with the GDM-S (left panels), patterns with the L-BFGS method (right panels) are closer to the prescribed VEC. More statistic data will be presented in the next paragraphs.

MAEs of the four numerical experiments after assimilation are calculated and listed in Table 3. The initial values of MAE in all NEs are $1 \times 10^{-3} \text{ m}^2/\text{s}$. Based on the results shown in Table 3, all the MAEs after assimilation are more than one order of magnitude lower than the initial values, which means the success for both of the two optimization methods. Note that the MAEs with the L-BFGS method are quite close (less than $1 \times 10^{-8} \text{ m}^2/\text{s}$ and cannot be distinguished in the table) and are less than half of those with the GDM-S. This result demonstrates the ability of the L-BFGS method to deduce the overall errors.

FIGURE 3: Topographies A and B. Note that the abscissa and ordinate axes are labeled with zonal index and meridional index, respectively.

FIGURE 4: Planform of two prescribed spatial distributions of VECV.

TABLE 3: Inversion errors of VECV in Group Two (unit: m^2/s).

Method	Experiment			
	NE1	NE2	NE3	NE4
GDM-S(I)	$2.42\text{E}-04$	$2.43\text{E}-04$	$2.23\text{E}-04$	$2.24\text{E}-04$
L-BFGS(II)	$9.14\text{E}-05$	$9.14\text{E}-05$	$9.14\text{E}-05$	$9.14\text{E}-05$

To compare the effectiveness of the two methods to invert the VECV, we make statistics on the percentages of the grids

FIGURE 5: MAEs versus IP distance in Group One. The abscissas indicates distance between adjacent IPs (unit: $^\circ$) while the ordinate indicates MAE of inversion results. The solid lines are values of different experiments and the dashed lines indicate the minimum values of two solid lines, respectively.

at which the MAEs are less than $1 \times 10^{-4} \text{ m}^2/\text{s}$, which is listed in Table 4. With the GDM-S, the inversion errors are deduced by one order of magnitude at about 40% of the total grids. By contrast, ratios of all NEs with the L-BFGS method are 79.76%, without differences between NEs. This phenomenon indicates that the L-BFGS method is effective at more computation grids than the GDM-S. Furthermore, the L-BFGS method maintains its effectiveness no matter which topography is applied.

Combining the inversion patterns, the inversion errors of the VECV, and the effectiveness analyses, conclusions can be drawn that the L-BFGS has a better performance in reducing the inversion errors.

Finally we come to the optimization history for all the experiments carried out in Group Two. The variations 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

FIGURE 6: Planform of inversion results in Group Two.

TABLE 4: Effectiveness analyses of inversions in Group Two.

Method	Experiment			
	NE1	NE2	NE3	NE4
GDM-S(I)	42.53%	42.24%	40.99%	40.91%
L-BFGS(II)	79.76%	79.76%	79.76%	79.76%

VEVCs and that of the inversion error, are plotted in Figures 7(a), 7(b), and 7(c), respectively, as a function of the iteration step.

Note that all experiments with the L-BFGS method reach the convergence criterion and stop after 4 iterations, which indicates that the computation time for the L-BFGS method is one twenty-fifth of that for the GDM-S. Figure 7(a) indicates that all the cost functions are in downward trends throughout the iteration process and decrease by more than 2 (7) orders of magnitude for the GDM-S (L-BFGS method), which means that the final differences between simulation value and the observation of these two methods are less than one-tenth and one-thousandth of their initial values, respectively. As is shown in Figure 7(b), the L_2 norms of gradient of the cost function with respect to the VEV decrease by more than 1 order of magnitude (GDM-S) and 4 orders of magnitude (L-BFGS), compared with their respective initial values. This indicates that the inversion result is becoming increasingly closer to the given VEV during the iteration. As shown in Figure 7(c), the inversion errors with the two methods keep declining throughout the iterations until the stopping criterions are satisfied. In general, the cost functions, norms of gradient, and the inversion errors of the VEV have steady descent, which demonstrates that this adjoint model is capable to invert the VEV. What is more, both the GDM-S and the L-BFGS methods are effective in terms of the inversion of the control parameters with spatially distributions of internal tide.

It is also clear in Figure 7 that the convergence rate for the cost functions, norms of gradient, and the inversion errors of the VEV is much faster with the L-BFGS method than those with the GDM-S, which is consistent with the classic theories about the convergence rate of the quasi-Newton method and the GDM [36]. This trend is also consistent with the results of numerical experiments to invert the open boundary conditions in Chen et al. [30]. With no doubt, the L-BFGS method is a more effective and efficient optimization method to invert the spatially varying VEV. However, the GDM-S is easier to understand and to implement in the model. Moreover, the step length and the search direction in the process of GDM-S can be freely controlled by the modelers, which is very convenient in practice. Therefore, for the inversion of the VEV, the GDM-S should also be regarded as a choice.

6. Summary and Conclusions

Based on an isopycnic-coordinate internal tidal model, the inversion of VEV is studied in this paper. A series of numerical experiments are carried out to examine the influence factors on the inversion of VEVs in four aspects:

independent point schemes (IPS), topography, the spatial distribution of VEV, and the optimization methods. For each experiment, the cost function, the L_2 norm of gradient of cost function with respect to the VEV, and the inversion error are calculated and analyzed in details.

The IPS is introduced and discussed in Group One. All the VEVs can be inverted successfully with IPS. MAE is regarded as the comparison criterion of the result. After comparing the 9 experiments, the correctness of the IPS is confirmed and the optimal IPSs are selected for the GDM-S and the L-BFGS methods, respectively.

Based on the optimal IPSs in Group One, two kinds of VEV distributions are successfully inverted with this adjoint model on two kinds of topography in Group Two. MAEs after optimization are at the level of 10^{-4} (10^{-5}) for the GDM-S (L-BFGS), which is one (two) order(s) of magnitude lower than the initial value. All the cost functions and their gradient norms with respect to the VEV lead satisfactory declines no matter which optimization method is taken. Compared with the GDM-S, the L-BFGS method has a remarkably better performance, not only in terms of the convergence rate but also in terms of the final inversion results. The computation time for the L-BFGS method is much shorter than that for the GDM-S. To sum up, the L-BFGS method is a more effective and efficient method than the GDM-S in terms of the inversion of the VEV. Nevertheless, the GDM-S is more convenient and controllable so it should not be ignored and should be taken seriously as a choice for the inversion of the VEV with spatially distribution.

The success of numerical experiments lays a solid foundation for the practical experiments and encourages us to carry out experiments in practical sea area with measured data and the real T/P altimeter data.

Appendix

Derivation of (2)

Let us start with the governing equations in Chen et al. [27].

Layer 1 (surface layer)

$$\frac{\partial q_1}{\partial t} + \frac{1}{R \cos \varphi} \frac{\partial (q_1 u_1)}{\partial \lambda} + \frac{1}{R \cos \varphi} \frac{\partial (q_1 v_1 \cos \varphi)}{\partial \varphi} = 0, \quad (\text{A.1a})$$

$$\begin{aligned} \frac{\partial u_1}{\partial t} + \frac{u_1}{R \cos \varphi} \frac{\partial u_1}{\partial \lambda} + \frac{v_1}{R} \frac{\partial u_1}{\partial \varphi} - \frac{u_1 v_1 \tan \varphi}{R} - f v_1 - A_{h1} \Delta u_1 \\ + \frac{\tau_{1\lambda}}{q_1} + \frac{g}{R \cos \varphi} \left[\frac{\partial}{\partial \lambda} \sum_{m=1}^l \left(\frac{q_m}{\rho_m} - h_m \right) + \frac{\rho_1}{\rho_k} \frac{\partial \bar{\zeta}}{\partial \lambda} \right] = 0, \end{aligned} \quad (\text{A.1b})$$

$$\begin{aligned} \frac{\partial v_1}{\partial t} + \frac{u_1}{R \cos \varphi} \frac{\partial v_1}{\partial \lambda} + \frac{v_1}{R} \frac{\partial v_1}{\partial \varphi} + \frac{u_1^2 \tan \varphi}{R} - f u_1 - A_{h1} \Delta v_1 \\ + \frac{\tau_{1\varphi}}{q_1} + \frac{g}{R} \left[\frac{\partial}{\partial \varphi} \sum_{m=1}^l \left(\frac{q_m}{\rho_m} - h_m \right) + \frac{\rho_1}{\rho_k} \frac{\partial \bar{\zeta}}{\partial \varphi} \right] = 0. \end{aligned} \quad (\text{A.1c})$$

FIGURE 7: Optimization history for experiments of Group Two, about (a) the cost function normalized by its initial value J_0 , (b) the L_2 norm of gradient of the cost function with respect to the VEVC, and (c) the MAEs between the inverted and prescribed VEVCs.

Layer k ($k = 2, \dots, l-1$)

$$\frac{\partial q_k}{\partial t} + \frac{1}{R \cos \varphi} \frac{\partial (q_k u_k)}{\partial \lambda} + \frac{1}{R \cos \varphi} \frac{\partial (q_k v_k \cos \varphi)}{\partial \varphi} = 0, \quad (\text{A.2a})$$

$$\begin{aligned} & \frac{\partial u_k}{\partial t} + \frac{u_k}{R \cos \varphi} \frac{\partial u_k}{\partial \lambda} + \frac{v_k}{R} \frac{\partial u_k}{\partial \varphi} - \frac{u_k v_k \tan \varphi}{R} - f v_k - A_{hk} \Delta u_k \\ & - \frac{\tau_{(k-1)\lambda} - \tau_{k\lambda}}{q_k} + \frac{g}{R \cos \varphi} \\ & \times \left[\sum_{m=1}^{k-1} \left(\frac{1}{\rho_k} - \frac{1}{\rho_m} \right) \frac{\partial q_m}{\partial \lambda} \right. \\ & \left. + \frac{\partial}{\partial \lambda} \sum_{m=1}^l \left(\frac{q_m}{\rho_m} - h_m \right) + \frac{\rho_1}{\rho_k} \frac{\partial \bar{\zeta}}{\partial \lambda} \right] = 0, \end{aligned} \quad (\text{A.2b})$$

$$\begin{aligned} & \frac{\partial v_k}{\partial t} + \frac{u_k}{R \cos \varphi} \frac{\partial v_k}{\partial \lambda} + \frac{v_k}{R} \frac{\partial v_k}{\partial \varphi} + \frac{u_k^2 \tan \varphi}{R} + f u_k \\ & - A_{hk} \Delta v_k - \frac{\tau_{(k-1)\varphi} - \tau_{k\varphi}}{q_k} + \frac{g}{R} \\ & \times \left[\sum_{m=1}^{k-1} \left(\frac{1}{\rho_k} - \frac{1}{\rho_m} \right) \frac{\partial q_m}{\partial \varphi} \right. \\ & \left. + \frac{\partial}{\partial \varphi} \sum_{m=1}^l \left(\frac{q_m}{\rho_m} - h_m \right) + \frac{\rho_1}{\rho_k} \frac{\partial \bar{\zeta}}{\partial \varphi} \right] = 0. \end{aligned} \quad (\text{A.2c})$$

Layer l (bottom layer)

$$\frac{\partial q_l}{\partial t} + \frac{1}{R \cos \varphi} \frac{\partial (q_l u_l)}{\partial \lambda} + \frac{1}{R \cos \varphi} \frac{\partial (q_l v_l \cos \varphi)}{\partial \varphi} = 0, \quad (\text{A.3a})$$

$$\begin{aligned} & \frac{\partial u_l}{\partial t} + \frac{u_l}{R \cos \varphi} \frac{\partial u_l}{\partial \lambda} + \frac{v_l}{R} \frac{\partial u_l}{\partial \varphi} - \frac{u_l v_l \tan \varphi}{R} - f v_l \\ & - A_{hl} \Delta u_l - \frac{\tau_{(l-1)\lambda} - \tau_{bl\lambda}}{q_l} + \frac{g}{R \cos \varphi} \end{aligned} \quad (\text{A.3b})$$

$$\times \left[\frac{\partial}{\partial \lambda} \sum_{m=1}^l \left(\frac{q_m}{\rho_l} - h_m \right) + \frac{\rho_1}{\rho_l} \frac{\partial \bar{\zeta}}{\partial \lambda} \right] = 0,$$

$$\begin{aligned} & \frac{\partial v_l}{\partial t} + \frac{u_l}{R \cos \varphi} \frac{\partial v_l}{\partial \lambda} + \frac{v_l}{R} \frac{\partial v_l}{\partial \varphi} + \frac{u_l^2 \tan \varphi}{R} + f u_l \\ & - A_{hl} \Delta v_l - \frac{\tau_{(l-1)\varphi} - \tau_{bl\varphi}}{q_l} + \frac{g}{R} \\ & \times \left[\frac{\partial}{\partial \varphi} \sum_{m=1}^l \left(\frac{q_m}{\rho_l} - h_m \right) + \frac{\rho_1}{\rho_l} \frac{\partial \bar{\zeta}}{\partial \varphi} \right] = 0. \end{aligned} \quad (\text{A.3c})$$

The variables and background of the governing equations have been introduced in Chen's [27] work in details. We will not repeat them in this part. The interface and friction terms are expressed by

$$\begin{aligned} (\tau_{k\lambda}, \tau_{k\varphi}) &= A_{vk} \frac{\rho_{k+1/2}}{h_{k+1/2}} (u_k - u_{k+1}, v_k - v_{k+1}), \\ k &= 1, \dots, l-1, \end{aligned} \quad (\text{A.4})$$

where A_v is the vertical eddy viscosity coefficient, $\rho_{k+1/2} = (\rho_k + \rho_{k+1})/2$, and $h_{k+1/2} = (h_k + h_{k+1})/2$.

The cost function is defined as

$$\begin{aligned}
 J(q, u, v; \mathbf{p}) &= \frac{1}{2} \int \left\{ K_\zeta \sum_{k=1}^l \left[\sum_{m=k}^l \left(\frac{q_m}{\rho_m} - h_m \right) \hat{\zeta}_k \right]^2 \right. \\
 &\quad \left. + K_u \sum_{k=1}^l (u_k - \hat{u}_k)^2 + K_v \sum_{k=1}^l (v_k - \hat{v}_k)^2 \right\} d\sigma, \quad (\text{A.5})
 \end{aligned}$$

which is exactly the same as that in [27]. Then the Lagrangian function is defined as

$$\begin{aligned}
 L(q, u, v; q_a, u_a, v_a; \mathbf{p}) &= J(q, u, v; \mathbf{p}) \\
 &+ \int [q_{a1} \cdot (\text{A.1a}) + u_{a1} q_1 \cdot (\text{A.1b}) + v_{a1} q_1 \cdot (\text{A.1c})] d\sigma \\
 &+ \dots \\
 &+ \int [q_{ak} \cdot (\text{A.2a}) + u_{ak} q_k \cdot (\text{A.2b}) + v_{ak} q_k \cdot (\text{A.2c})] d\sigma \\
 &+ \dots \\
 &+ \int [q_{al} \cdot (\text{A.3a}) + u_{al} q_l \cdot (\text{A.3b}) + v_{al} q_l \cdot (\text{A.3c})] d\sigma, \quad (\text{A.6})
 \end{aligned}$$

where

$$\begin{aligned}
 (\text{A.1a}) &= - \left[\frac{\partial q_1}{\partial t} + \frac{1}{R \cos \varphi} \frac{\partial (q_1 u_1)}{\partial \lambda} + \frac{1}{R \cos \varphi} \frac{\partial (q_1 v_1 \cos \varphi)}{\partial \varphi} \right], \quad (\text{A.7})
 \end{aligned}$$

$$\begin{aligned}
 (\text{A.1b}) &= - \left\{ \frac{\partial u_1}{\partial t} + \frac{u_1}{R \cos \varphi} \frac{\partial u_1}{\partial \lambda} + \frac{v_1}{R} \frac{\partial u_1}{\partial \varphi} - \frac{u_1 v_1 \tan \varphi}{R} \right. \\
 &\quad \left. - f v_1 - A_{h1} \Delta u_1 + \frac{\tau_{1\lambda}}{q_1} + \frac{g}{R \cos \varphi} \right. \\
 &\quad \left. \times \left[\frac{\partial}{\partial \lambda} \sum_{m=1}^l \left(\frac{q_m}{\rho_m} - h_m \right) + \frac{\rho_1}{\rho_k} \frac{\partial \bar{\zeta}}{\partial \lambda} \right] \right\}, \quad (\text{A.8})
 \end{aligned}$$

$$\begin{aligned}
 (\text{A.1c}) &= - \left\{ \frac{\partial v_1}{\partial t} + \frac{u_1}{R \cos \varphi} \frac{\partial v_1}{\partial \lambda} + \frac{v_1}{R} \frac{\partial v_1}{\partial \varphi} + \frac{u_1^2 \tan \varphi}{R} \right. \\
 &\quad \left. - f u_1 - A_{h1} \Delta v_1 + \frac{\tau_{1\varphi}}{q_1} + \frac{g}{R} \right. \\
 &\quad \left. \times \left[\frac{\partial}{\partial \varphi} \sum_{m=1}^l \left(\frac{q_m}{\rho_m} - h_m \right) + \frac{\rho_1}{\rho_k} \frac{\partial \bar{\zeta}}{\partial \varphi} \right] \right\}, \quad (\text{A.9})
 \end{aligned}$$

same definitions are applied in (A.2a)~(A.2c) and (A.3a)~(A.3c). Then the Lagrangian function $L(q, u, v; q_a, u_a, v_a; \mathbf{p})$ can be written as

$$\begin{aligned}
 L(q, u, v; q_a, u_a, v_a; \mathbf{p}) &= J(q, u, v; \mathbf{p}) \\
 &- \int \left[\frac{u_{a1} A_{vi+1/2, j, 1} \rho_{1+1/2}}{h_{i+1/2, j, 1+1/2}} (u_1 - u_2) \right. \\
 &\quad \left. + \frac{u_{a1} A_{vi+1/2, j, 1} \rho_{1+1/2}}{h_{i+1/2, j, 1+1/2}} (v_1 - v_2) - F_1 \right] d\sigma \\
 &+ \dots \\
 &- \int \left[\frac{u_{ak} A_{vi+1/2, j, k} \rho_{k+1/2}}{h_{i+1/2, j, k+1/2}} (u_k - u_{k+1}) \right. \\
 &\quad \left. - \frac{u_{ak} A_{vi+1/2, j, k-1} \rho_{k-1/2}}{h_{i+1/2, j, k-1/2}} (u_{k-1} - u_k) \right] d\sigma \\
 &- \int \left[\frac{v_{ak} A_{vi+1/2, j, k} \rho_{k+1/2}}{h_{i+1/2, j, k+1/2}} (v_k - u v_{k+1}) \right. \\
 &\quad \left. - \frac{v_{ak} A_{vi+1/2, j, k-1} \rho_{k-1/2}}{h_{i+1/2, j, k-1/2}} (v_{k-1} - v_k) - F_k \right] d\sigma \\
 &+ \dots \\
 &+ \int \left[\frac{u_{al} A_{vi+1/2, j, l-1} \rho_{l-1/2}}{h_{i+1/2, j, l-1/2}} (u_{l-1} - u_l) \right. \\
 &\quad \left. + \frac{u_{al} A_{vi+1/2, j, l-1} \rho_{l-1/2}}{h_{i+1/2, j, l-1/2}} (v_{l-1} - v_l) - F_l \right] d\sigma, \quad (\text{A.10})
 \end{aligned}$$

where $A_{vi+1/2, j, k} = (A_{vi, j, k} A_{vi+1, j, k})/2$, $h_{i+1/2, j, k+1/2} = (h_{i, j, k} + h_{i+1, j, k} + h_{i, j, k+1} + h_{i+1, j, k+1})/4$, and functions F_1 , F_k , and F_l are, respectively, defined as

$$\begin{aligned}
 F_1 &= q_{a1} \cdot (\text{A.1a}) + u_{a1} q_1 \cdot \left[(\text{A.1b}) + \frac{\tau_{1\lambda}}{q_1} \right] \\
 &+ v_{a1} q_1 \cdot \left[(\text{A.1c}) + \frac{\tau_{1\varphi}}{q_1} \right], \quad (\text{A.11})
 \end{aligned}$$

$$F_k = q_{ak} \cdot (A.2a) + u_{ak} q_k \cdot \left[(A.2b) - \frac{\tau_{(k-1)\lambda} - \tau_{k\lambda}}{q_1} \right] + v_{ak} q_k \cdot \left[(A.2c) - \frac{\tau_{(k-1)\varphi} - \tau_{k\varphi}}{q_k} \right], \quad (A.12)$$

$$F_l = q_{al} \cdot (A.3a) + u_{al} q_l \cdot \left[(A.3b) - \frac{\tau_{(l-1)\lambda}}{q_l} \right] + v_{al} q_l \cdot \left[(A.3c) - \frac{\tau_{(l-1)\varphi}}{q_l} \right]. \quad (A.13)$$

Note that (A.11)~(A.13) do not contain the variable A_v , which means

$$\frac{\partial F_1}{\partial A_{vi,j,k}} = \frac{\partial F_k}{\partial A_{vi,j,k}} = \frac{\partial F_l}{\partial A_{vi,j,k}} = 0. \quad (A.14)$$

The Lagrangian function can be written as

$$\begin{aligned} L(q, u, v; q_a, u_a, v_a; \mathbf{p}) \\ &= J(q, u, v; \mathbf{p}) \\ &- \int \left[\frac{u_{a1,j,1} A_{vi+1/2,j,1} \rho_{1+1/2}}{h_{i+1/2,j,1+1/2}} (u_{i,j,1} - u_{i,j,2}) \right. \\ &\quad \left. + \frac{u_{a1,j,1} A_{vi+1/2,j,1} \rho_{1+1/2}}{h_{i+1/2,j,1+1/2}} (v_{i,j,1} - v_{i,j,2}) - F_1 \right] d\sigma \\ &+ \dots \\ &- \int \left[\frac{u_{ai,j,k} A_{vi+1/2,j,k} \rho_{k+1/2}}{h_{i+1/2,j,k+1/2}} (u_{i,j,k} - u_{i,j,k+1}) \right. \\ &\quad \left. - \frac{u_{ai,j,k} A_{vi+1/2,j,k-1} \rho_{k-1/2}}{h_{i+1/2,j,k-1/2}} (u_{i,j,k-1} - u_{i,j,k}) \right] d\sigma \\ &- \int \left[\frac{v_{ai,j,k} A_{vi+1/2,j,k} \rho_{k+1/2}}{h_{i+1/2,j,k+1/2}} (v_{i,j,k} - v_{i,j,k+1}) \right. \\ &\quad \left. - \frac{v_{ak} A_{vi+1/2,j,k-1} \rho_{k-1/2}}{h_{i+1/2,j,k-1/2}} (v_{i,j,k-1} - v_{i,j,k}) - F_k \right] d\sigma \\ &+ \dots \\ &+ \int \left[\frac{u_{ai,j,l} A_{vi+1/2,j,l-1} \rho_{l-1/2}}{h_{i+1/2,j,l-1/2}} (u_{i,j,l-1} - u_{i,j,l}) \right. \\ &\quad \left. + \frac{u_{ai,j,l} A_{vi+1/2,j,l-1} \rho_{l-1/2}}{h_{i+1/2,j,l-1/2}} (v_{i,j,l-1} - v_{i,j,l}) - F_l \right] d\sigma. \end{aligned} \quad (A.15)$$

Finally, according to the typical theory of Lagrangian multiplier method, we have the following first-order derivate of Lagrangian function with respect to the control parameter A_v :

$$\frac{\partial L}{\partial A_{vi,j,k}} = 0, \quad (A.16)$$

then the gradient of the cost function with respect to the variable A_v can be deduced from (A.16):

$$\begin{aligned} \frac{\partial J}{\partial A_{vi,j,k}} \\ &= \sum_n \left[u_{ai,j,k}^n \frac{\rho_{k+1/2}}{h_{i+1/2,j,k+1/2}} (u_{i,j,k}^n - u_{i,j,k+1}^n) \right. \\ &\quad \left. + u_{ai-1,j,k}^n \frac{\rho_{k+1/2}}{h_{i-1/2,j,k+1/2}} (u_{i-1,j,k}^n - u_{i-1,j,k+1}^n) \right] \\ &- \sum_n \left[u_{ai,j,k+1}^n \frac{\rho_{k+1/2}}{h_{i+1/2,j,k+1/2}} (u_{i,j,k}^n - u_{i,j,k+1}^n) \right. \\ &\quad \left. + u_{ai-1,j,k+1}^n \frac{\rho_{k+1/2}}{h_{i-1/2,j,k+1/2}} (u_{i-1,j,k}^n - u_{i-1,j,k+1}^n) \right] \\ &+ \sum_n \left[v_{ai,j,k}^n \frac{\rho_{k+1/2}}{h_{i+1/2,j,k+1/2}} (v_{i,j,k}^n - v_{i,j,k+1}^n) \right. \\ &\quad \left. + v_{ai-1,j,k}^n \frac{\rho_{k+1/2}}{h_{i-1/2,j,k+1/2}} (v_{i-1,j,k}^n - v_{i-1,j,k+1}^n) \right] \\ &- \sum_n \left[v_{ai,j,k+1}^n \frac{\rho_{k+1/2}}{h_{i+1/2,j,k+1/2}} (v_{i,j,k}^n - v_{i,j,k+1}^n) \right. \\ &\quad \left. + v_{ai-1,j,k+1}^n \frac{\rho_{k+1/2}}{h_{i-1/2,j,k+1/2}} (v_{i-1,j,k}^n - v_{i-1,j,k+1}^n) \right] \\ &= (\rho_k + \rho_{k+1}) \\ &\times \sum_n \left[\frac{(u_{i,j,k}^n - u_{i,j,k+1}^n)(u_{ai,j,k}^n - u_{ai,j,k+1}^n)}{h_{i,j,k} + h_{i+1,j,k} + h_{i,j,k+1} + h_{i+1,j,k+1}} \right. \\ &\quad \left. + \frac{(u_{i-1,j,k}^n - u_{i-1,j,k+1}^n)(u_{ai-1,j,k}^n - u_{ai-1,j,k+1}^n)}{h_{i-1,j,k} + h_{i,j,k} + h_{i-1,j,k+1} + h_{i,j,k+1}} \right] \\ &+ (\rho_k + \rho_{k+1}) \\ &\times \sum_n \left[\frac{(v_{i,j,k}^n - v_{i,j,k+1}^n)(v_{ai,j,k}^n - v_{ai,j,k+1}^n)}{h_{i,j,k} + h_{i+1,j,k} + h_{i,j,k+1} + h_{i+1,j,k+1}} \right. \\ &\quad \left. + \frac{(v_{i-1,j,k}^n - v_{i-1,j,k+1}^n)(v_{ai-1,j,k}^n - v_{ai-1,j,k+1}^n)}{h_{i-1,j,k} + h_{i,j,k} + h_{i-1,j,k+1} + h_{i,j,k+1}} \right]. \end{aligned} \quad (A.17)$$

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

Partial support for this research was provided by the National Natural Science Foundation of China through Grant 41371496, the State Ministry of Science and Technology of China through Grant 2013AA122803, and the Fundamental Research Funds for the Central Universities 201262007 and 201362033.

References

- [1] M. Rattray, "On the coastal generation of internal tides," *Tellus*, vol. 12, no. 1, pp. 54–62, 1960.
- [2] P. G. Baines, "The generation of internal tides by flat-bump topography," *Deep-Sea Research and Oceanographic Abstracts*, vol. 20, no. 2, pp. 179–205, 1973.
- [3] T. H. Bell, "Topographically generated internal waves in the open ocean," *Journal of Geophysical Research*, vol. 80, pp. 320–327, 1975.
- [4] P. G. Baines, "On internal tide generation models," *Deep Sea Research A: Oceanographic Research Papers*, vol. 29, no. 3, pp. 307–338, 1982.
- [5] P. D. Craig, "Solutions for internal tidal generation over coastal topography," *Journal of Marine Research*, vol. 45, pp. 83–105, 1987.
- [6] T. Gerkema, "A unified model for the generation and fission of internal tides in a rotating ocean," *Journal of Marine Research*, vol. 54, no. 3, pp. 421–450, 1996.
- [7] S. G. Llewellyn Smith and W. R. Young, "Conversion of the barotropic tide," *Journal of Physical Oceanography*, vol. 32, no. 5, pp. 1554–1566, 2002.
- [8] S. K. Kang, M. G. G. Foreman, W. R. Crawford, and J. Y. Cherniawsky, "Numerical modeling of internal tide generation along the Hawaiian Ridge," *Journal of Physical Oceanography*, vol. 30, no. 5, pp. 1083–1098, 2000.
- [9] Y. Niwa and T. Hibiya, "Numerical study of the spatial distribution of the M_2 internal tide in the Pacific Ocean," *Journal of Geophysical Research C: Oceans*, vol. 106, no. 10, pp. 22441–22449, 2001.
- [10] P. F. Cummins, J. Y. Cherniawsky, and M. G. G. Foreman, "North Pacific internal tides from the Aleutian ridge: altimeter observations and modeling," *Journal of Marine Research*, vol. 59, no. 2, pp. 167–191, 2001.
- [11] Y. Niwa and T. Hibiya, "Three-dimensional numerical simulation of M_2 internal tides in the East China Sea," *Journal of Geophysical Research C: Oceans*, vol. 109, no. 4, Article ID C04027, 2004.
- [12] S. Jan, C.-S. Chern, J. Wang, and S.-Y. Chao, "Generation of diurnal K_1 internal tide in the Luzon Strait and its influence on surface tide in the South China Sea," *Journal of Geophysical Research C: Oceans*, vol. 112, no. 6, Article ID C06019, 2007.
- [13] R. C. Pacanowski and S. G. H. Philander, "Parameterization of vertical mixing in numerical models of tropical oceans," *Journal of Physical Oceanography*, vol. 11, no. 11, pp. 1443–1451, 1981.
- [14] B. Henderson-Sellers, "A simple formula for vertical eddy diffusion coefficients under conditions of nonneutral stability," *Journal of Geophysical Research*, vol. 87, pp. 5860–5864, 1982.
- [15] N. S. Heaps, "Three-dimensional model for tides and surges with vertical eddy viscosity prescribed in two layers—I. Mathematical formulation," *Geophysical Journal of the Royal Astronomical Society*, vol. 64, pp. 291–302, 1981.
- [16] N. S. Heaps and J. E. Jones, "Three-dimensional model for tides and surges with vertical eddy viscosity prescribed in two layers—II. Irish Sea with bed friction layer," *Geophysical Journal of the Royal Astronomical Society*, vol. 64, pp. 303–320, 1981.
- [17] V. P. Kochergin, "Three-dimensional prognostic models. In: three-dimensional coastal ocean models," *Coastal and Estuarine Science*, vol. 4, pp. 201–208, 1987.
- [18] T. Pohlmann, "Calculating the annual cycle of the vertical eddy viscosity in the North Sea with a three-dimensional baroclinic shelf sea circulation model," *Continental Shelf Research*, vol. 16, no. 2, pp. 147–161, 1996.
- [19] 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.
- [20] L. Yu and J. J. O'Brien, "Variational estimation of the wind stress drag coefficient and the oceanic eddy viscosity profile," *Journal of Physical Oceanography*, vol. 21, pp. 709–719, 1991.
- [21] R. W. Lardner, "Optimal control of open boundary conditions for a numerical tidal model," *Computer Methods in Applied Mechanics and Engineering*, vol. 102, no. 3, pp. 367–387, 1993.
- [22] U. Seiler, "Estimation of open boundary conditions with the adjoint method," *Journal of Geophysical Research*, vol. 98, no. 12, pp. 22855–22870, 1993.
- [23] 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.
- [24] N. Ayoub, "Estimation of boundary values in a North Atlantic circulation model using an adjoint method," *Ocean Modelling*, vol. 12, no. 3–4, pp. 319–347, 2006.
- [25] J. Zhang and X. 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.
- [26] 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.
- [27] H. Chen, C. Miao, and X. 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.
- [28] J. Zhang and H. Chen, "Semi-idealized study on estimation of partly and fully space varying open boundary conditions for tidal models," *Abstract and Applied Analysis*, vol. 2013, Article ID 282593, 14 pages, 2013.
- [29] A. Cao, H. Chen, J. Zhang, and X. Lv, "Optimization of open boundary conditions in a 3D internal tidal model with the adjoint method around Hawaii," *Abstract and Applied Analysis*, vol. 2013, Article ID 950926, 11 pages, 2013.
- [30] H. Chen, C. Miao, and X. Lv, "Estimation of open boundary conditions for an internal tidal model with adjoint method: a comparative study on optimization methods," *Mathematical Problems in Engineering*, vol. 2013, Article ID 802136, 12 pages, 2013.
- [31] X. Lu and J. 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.
- [32] J. Zhang, X. 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.
- [33] R. W. Lardner and S. K. Das, “Optimal estimation of eddy viscosity for a quasi-three-dimensional numerical tidal and storm surge model,” *International Journal for Numerical Methods in Fluids*, vol. 18, no. 3, pp. 295–312, 1994.
  - [34] J. E. Richardson and V. G. Panchang, “A modified adjoint method for inverse eddy viscosity estimation for use in coastal circulation models,” in *Proceedings of the 2nd International Conference on Estuarine and Coastal Modeling*, pp. 733–745, November 1992.
  - [35] G. P. Cressman, “An operational objective analysis system,” *Monthly Weather Review*, vol. 87, no. 10, pp. 367–374, 1959.
  - [36] S. J. Wright and J. Nocedal, *Numerical optimization*, Springer, New York, NY, USA, 1999.
  - [37] W. F. Mascarenhas, “The BFGS method with exact line searches fails for non-convex objective functions,” *Mathematical Programming*, vol. 99, no. 1, pp. 49–61, 2004.
  - [38] 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 and Software*, vol. 24, no. 1, pp. 63–87, 2009.
  - [39] X. Zou, I. M. Navon, M. Berger, K. H. Phua, T. Schlick, and F. 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.
  - [40] J. Nocedal, “Updating quasi-Newton matrices with limited storage,” *Mathematics of Computation*, vol. 35, no. 151, pp. 773–782, 1980.
  - [41] R. Malouf, “A comparison of algorithms for maximum entropy parameter estimation,” in *Proceedings of the 6th Conference on Natural Language Learning (CoNLL ’02)*, pp. 49–55, 2002.
  - [42] G. Andrew and J. Gao, “Scalable training of L1-regularized log-linear models,” in *Proceedings of the 24th International Conference on Machine Learning (ICML ’07)*, pp. 33–40, June 2007.
  - [43] 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.
  - [44] D. C. Liu and J. Nocedal, “On the limited memory BFGS method for large scale optimization,” *Mathematical Programming*, vol. 45, no. 1–3, pp. 503–528, 1989.
  - [45] L. J. Nocedal, “BFGS subroutine, software for large-scale unconstrained optimization,” <http://www.ece.northwestern.edu/~nocedal/lbfgs.html>.