Model terrain correction using variational adjoint method with Tikhonov-total variation regularization

Sulin Tao1,2,*, Yuhong Li3 and Isaac Mugume4

1Key Laboratory of Meteorological Disaster, Ministry of Education (KLME), Joint International Research Laboratory of Climate and Environment Change (ILCEC), Collaborative Innovation Centre on Forecast and Evaluation of Meteorological Disasters (CIC-FEMD), School of Applied Meteorology, Nanjing University of Information Science & Technology, Nanjing 210044, P.R.China
2Technology Development Centre, Nanjing Daqiao Machine CO.,Ltd, Nanjing 211101, P.R.China
3Liaoning Institute of Meteorological Sciences, Shenyang 110166, P.R.China
4Department of Geography, Geo-Informatics & Climatic Sciences, Makerere University, P.O. Box 7062, Kampala, Uganda

*Corresponding author’s e-mail: sulin.tao@outlook.com

Abstract. Numerical weather prediction models require optimal topographic data to improve the accuracy of the forecasts. A framework for bottom terrain correction of a shallow-water equations model is proposed with variational adjoint and regularization methods. The Tikhonov-total variation (TV) regularization with dual regularization parameters is introduced as a constraint for unique and stable correction. The limited-memory Broyden-Fletcher-Goldfarb-Shanno (LBFGS) method is used to minimize the cost function. Numerical results indicate that the Tikhonov-TV regularization helps improve the corrections by reducing both overall shifts and surface fluctuations. These positive impacts become much more obvious after considering dual regularization parameters. The corrections at final time of assimilation significantly improve the numerical predictions. Tests confirm the potential of the computational framework for bottom terrain correction of numerical weather prediction models to reduce model errors.

1. Introduction
Mismatches between terrain and meteorological fields are a common cause of model errors [1]. The bottom terrain, one of the critical parameters controlling atmospheric circulation, has a strong dynamic effect on flow movements. The influences of various terrains on the quality of weather forecasting are well documented [2], and optimal topographic data are expected in numerical models. “Optimal data” does not mean terrain with very high spatial resolutions but, rather, topography similar in scale to the grid intervals of prediction models [3]. Usually the terrain data are smoothed before they are used in models, and critical topographic information might be lost. However, a feedback mechanism exists between terrain and meteorological observations, and some of the topographic information is included in the observed meteorological data as a result of the dynamic effect of terrain.

In this paper we will address a test problem to correct the bottom terrain parameter in a numerical model with the help of meteorological observations. This problem forms a typical ill-posed inverse
problem. The variational adjoint method is efficient to solve the inverse problem by minimizing the difference between limited observed data and corresponding model results [4]. This procedure allows for the approaching solutions against the observations by the updated parameters at each iterate. However, it’s difficult to solve large-scale minimization problems, and the critical one is the high computational cost due to the backward integration of adjoint model for the further evaluation of the gradients of the cost function and the computation for the Hessian matrix of the least squares function.

Limited-memory variants of the optimization techniques are widely used for inverse problems due to the consideration of the computation stability and efficiency as the two issues of critical importance during the problem solution. The LBFGS method, a limited-memory quasi-Newton algorithm, applies a modified matrix-updating strategy for the Hessian approximation by the use of curvature information from the most recent iterations without matrix storage and too many vectors of memory [5, 6], and converges almost as rapidly as the quasi-Newton BFGS method [7]. It’s useful for solving large problems. Tests and real-life problems were introduced to examine this method compared with other common limited-memory quasi-Newton algorithms, and it outperforms CONMIN, E04DGF and BBVSCG [8]. Our previous work constructed a discretized assimilation system for a one-dimensional variable coefficient convection-diffusion problem with LBFGS algorithm, and obtained optimal estimations, but require further investigation with more complicated inverse problems [9].

The interest of this study is to present a framework for the identification of the optimal (or matched) bottom terrain to satisfy numerical weather prediction model. We follow the LBFGS method and consider regularization function to reduce the undesired oscillations, to allow the incorporation of a priori information about the desired bottom terrain and to guarantee the uniqueness of the optimal solution.

2. The test problem

2.1. The forward model

We consider the one-dimensional inviscid shallow-water equations model including the Coriolis term and bottom terrain describing the motion of a single-layer atmospheric flow [10]. The governing equations are in the following form:

\[
\begin{align*}
\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + \frac{\partial \varphi}{\partial x} &= f v - g \frac{\partial H}{\partial x}, \\
\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} &= -fu, \\
\frac{\partial \varphi}{\partial t} + u \frac{\partial \varphi}{\partial x} + \frac{\partial u}{\partial x} &= 0,
\end{align*}
\]

where \(x \in \Omega = (0, 2\pi L)\) and \(t \in (0, T)\). \(u = u(x, t)\) and \(v = v(x, t)\) are two components representing eastward and northward wind fields, respectively. \(\varphi = g \eta(x, t)\) is the geopotential field, where \(g\) is the acceleration due to gravity. \(\eta(x, t) > 0\) is the depth of atmospheric flow above the bottom terrain. \(H = H(x)\) is the height of the bottom terrain. \(f\) represents the Coriolis parameter.

Diffusion is introduced to the model equations (1a–c) to eliminate spurious oscillations that are generated by second-order finite difference schemes near jumps [1]. The diffusion terms of this scheme are discretized via forward time differenting for the stability, while all of the other terms are via centred time and space differencing. We define \(J\) and \(N\) as the number of spatial grids and the number of integration time steps, respectively. Then we discretize the spatial domain \([0, 2\pi L]\) to obtain the spatial interval \(\Delta x = 2\pi L / (J - 1)\) and the time domain \([0, T]\) to obtain the time interval \(\Delta t = T / N\). Thus the corresponding forward model can be stated as
\[
m^k_j = m^{k-1}_j - \frac{\Delta t}{2\Delta x^2} \left( (u^{k-1}_{j+1} + u^{k-1}_{j-1})(m^{k-1}_{j+1} + m^{k-1}_{j-1}) - (u^{k-1}_j + u^{k-1}_{j-1})(m^{k-1}_j + m^{k-1}_{j-1}) + (\phi^{k-1}_{j+1} - \phi^{k-1}_{j-1})^2 \right) - \left( \phi^{k-1}_{j+1} - \phi^{k-1}_{j-1} \right)^2 - g \frac{\Delta t}{2\Delta x} \left[ \left( \phi^{k-1}_{j+1} + \phi^{k-1}_{j-1} \right) (H_{j+1} - H_j) + \left( \phi^{k-1}_j - \phi^{k-1}_{j-1} \right) (H_j - H_{j-1}) \right]
\]
\[
+ 2\Delta t f n_{j+1}^{k-1} + 2 \frac{\Delta t}{\Delta x^2} K \left( m_{j+1}^{k-2} - 2m_{j+1}^{k-2} + m_{j-1}^{k-2} \right),
\]
\[
n^{k}_j = n^{k-1}_j - \frac{\Delta t}{\Delta x} \left( (v^{k-1}_{j+1} + v^{k-1}_{j-1})(m^{k-1}_{j+1} + m^{k-1}_{j-1}) - (v^{k-1}_j + v^{k-1}_{j-1})(m^{k-1}_j + m^{k-1}_{j-1}) \right)
\]
\[
- 2\Delta t f m_{j+1}^{k-1} + 2 \frac{\Delta t}{\Delta x^2} K \left( n_{j+1}^{k-2} - 2n_{j+1}^{k-2} + n_{j-1}^{k-2} \right),
\]
\[
\phi^{k}_j = \phi^{k-1}_j - \frac{\Delta t}{\Delta x} \left( m_{j+1}^{k-2} - m_{j-1}^{k-2} \right) + 2 \frac{\Delta t}{\Delta x^2} K \left( \phi^{k-2}_{j+1} - 2\phi^{k-2}_j + \phi^{k-2}_{j-1} \right),
\]
where \( m^k_j = u^k_j \phi^k_j \) and \( n^{k}_j = v^{k}_j \phi^{k}_j \), for \( k=2,3,\cdots,N \) and \( j=1,2,\cdots,J \). While at the first time step, i.e., \( k=1 \), one-step forward time differences is applied. \( u^1_j \), \( v^1_j \) and \( \phi^1_j \) are obtained with the initial values of the model states \( u^0_j \), \( v^0_j \) and \( \phi^0_j \) by replacing the terms \( u^{k-1}_j \), \( v^{k-1}_j \) and \( \phi^{k-1}_j \) in the equations (2a–c) with \( u^{k-1}_j \), \( v^{k-1}_j \) and \( \phi^{k-1}_j \), respectively. Two virtual grid points are introduced to treat the boundary conditions for this discrete scheme, including (i) the grid point \( x_0 \), which is located on the west side of grid \( x_i \) at a distance of \( \Delta x \), and (ii) the grid point \( x_{J+1} \), which is located on the east side of grid \( x_j \) at the same distance. The periodic boundary conditions in the form
\[
u^0_0 = u^1_1 \quad \text{and} \quad u^0_{J+1} = u^1_0, \quad v^0_0 = v^1_1 \quad \text{and} \quad v^0_{J+1} = v^1_0, \quad \phi^0_0 = \phi^1_1 \quad \text{and} \quad \phi^0_{J+1} = \phi^1_0,
\]
are considered for \( k=0,1,\cdots,N \). The solution at the time \( t_k \) is obtained only in the internal grids \( x_1, x_2, \ldots, x_N \) from the fields at the time \( t_{k-1} \) and \( t_{k-2} \) defined in the grids \( x_{1,1}, x_{1,2}, \ldots, x_{J,1} \), and the values of the fields at the time \( t_k \) in the virtual grids \( x_0 \) and \( x_{J+1} \) are obtained from cyclic boundary conditions (3).

Model-generated observations of the shallow-water system are taken to be the solutions to the forward model equations (2a–c) with the Coriolis parameter \( f = 7.292 \times 10^{-5} \text{ s}^{-1} \), the diffusion coefficient \( K = 2.5 \times 10^4 \text{ m}^2 \text{ s}^{-1} \), the gravitational acceleration \( g = 9.8 \text{ m s}^{-2} \), the domain length \( L = 9.55 \times 10^5 \text{ m} \), and the time length \( T = 1.2 \times 10^5 \text{ s} \). Let \( J = 100 \), and \( N = 100 \); over the domain \([0,2\pi L] \times [0, T] \), the initial values of the states for the model equations (2a–c) are defined as follows:

\[
u^0_0 = U \cos \left( \frac{x_0}{L} \right), \quad v^0_0 = 0 \quad \text{and} \quad \phi^0_0 = \phi - \frac{U^2}{8} \cos \left( \frac{2x_0}{L} \right) + \left( \phi - \frac{U^2}{8} \right)^{1/2} U \cos \left( \frac{x_0}{L} \right),
\]

with \( \phi = g \eta_m \), where the constant \( U = 10 \text{ m s}^{-1} \), the mean depth of atmospheric flow \( \eta_m = 100 \text{ m} \), and \( j=1,2,\cdots,J \). The true bottom terrain with the mountain peak shape \( H_j \) is 50 \left[ 1 - \left( x_j - \pi L \right)^2 / a^2 \right] \text{ m when} x_j \in [\pi L - a, \pi L + a] \text{ and} 0 \text{ m when} x_j \in [0, \pi L - a) \cup (\pi L + a, 2\pi L] \text{, where} a = 10\Delta x \text{ and} j=1,2,\cdots,J \). The elevations at the two virtual grid points are treated as \( H_0 = H_j \) and \( H_{J+1} = H_1 \).

2.2. The cost function

Note that the bottom terrain elevation under consideration is defined up to an arbitrary constant: the solutions of the model (1a–c) with the bottom terrain \( H(x) \) and \( H(x) + H_c \) (where \( H_c \) denotes a constant) will be identical. Thus, the model terrain correction problem is ill-posed, the Hessian matrix has null space, and the optimal solution is neither unique nor stable. If \( H(x) \) is the optimal solution obtained by data assimilation procedure, then \( H(x) + H_c \) is another optimal solution with the same
value of the cost function but different root mean square (RMS) errors with respect to the true bottom terrain. This exactly suggests the regularization function. A common regularization function will impose a smooth character on the solution. It might not be effective to identify the matched bottom terrain without losing important terrain information, such as the mountain topography. The total variation (TV) regularization function turns out to be the right choice to preserve the jump discontinuities and successfully applied in image restoration [11] and geophysical-property reconstruction [12-14].

Here defining the state vector \( \chi \in \mathbb{R}^{31 \times 1} \) by combining \( j_u \), \( j_v \) and \( j_\phi \) in this form
\[
(1,1,\ldots,1,1,j_u^k,\ldots,j_v^k,j_\phi^k,\ldots,1,1)^T
\]
and the control vector \( H \in \mathbb{R}^{1 \times 1} \) in the form
\[
(1,1,\ldots,1,1,H_1^k,\ldots,H_{Nt}^k)^T
\]
then, the best correction of the terrain results in seeking a value of \( H^* \) that minimizes a cost function
\[
J(\chi,H) = J_1 + \mu(J_2 + J_3),
\]
(5)
which combines a data misfit \( J_1 \), a priori constraint \( J_2 \) and a TV function \( J_3 \). The data misfit \( J_1 \) is defined as the weighted sum of the squared discrepancies between the observations, \( \hat{u}_j^k, \hat{v}_j^k \) and \( \bar{\phi}_j^k \), and the model predictions, \( u_j^k, v_j^k \) and \( \phi_j^k \). The combination of \( J_2 \) and \( J_3 \) are introduced as penalized cost function, also as the Tikhonov-TV regularization term, which ensures a well-posed inverse problem. The prior constraint \( J_2 \) formulates the absolute constraint by forcing the proximity between the identified and known terrain elevations, and the TV function \( J_3 \) provides the relative constraint from the relative variation for spatially adjacent elevations. \( \mu \) is a nonnegative regularization parameter. In this study, it’s determined after a great many trials and errors.

The data misfit \( J_1 \) is given by
\[
J_1 = \frac{1}{2} \left[ \sum_{k=1}^{N} \sum_{j=1}^{J} \alpha \left( u_j^k - \hat{u}_j^k \right)^2 \delta_j^q \delta_j^s + \sum_{k=1}^{N} \sum_{j=1}^{J} \beta \left( v_j^k - \hat{v}_j^k \right)^2 \delta_j^q \delta_j^s + \sum_{k=1}^{N} \sum_{j=1}^{J} \gamma \left( \phi_j^k - \bar{\phi}_j^k \right)^2 \delta_j^q \delta_j^s \right],
\]
(6)
where \( \alpha, \beta \) and \( \gamma \) are weighting factors that determine the contributions of eastward wind, northward wind and geopotential fields, respectively, to the cost function during the minimization procedures, and they are also used to scale the cost function, making it dimensionless. These weighting factors can be determined as the inverse of the maximum difference between the analysed fields at the initial time \( t_0 \) and that at the final time \( t_{Nt} \) by the formulae: \( \alpha = 1/\max_j |\hat{u}_j^k - \hat{u}_j^{k-1}| \), \( \beta = 1/\max_j |\hat{v}_j^k - \hat{v}_j^{k-1}| \) and \( \gamma = 1/\max_j |\bar{\phi}_j^k - \bar{\phi}_j^{k-1}| \). In this way, \( \alpha, \beta \) and \( \gamma \) are found to be \( 1.2 \times 10^{-3} \) m\(^2\) s\(^{-2}\), \( 1.8 \times 10^{-2} \) m\(^2\) s\(^{-2}\) and \( 2.2 \times 10^{-4} \) m\(^4\) s\(^{-4}\), respectively. \( q \) and \( s \) are variables of the Kronecker delta,
\[
\delta_j^q = \begin{cases} 0 & q \neq j \\ 1 & q = j \end{cases}, \quad \text{and} \quad \delta_j^s = \begin{cases} 0 & s \neq k \\ 1 & s = k \end{cases},
\]
(7)
respectively defined to control the density of the observations used in our assimilation problem.

The prior constraint \( J_2 \) is denoted by
\[
J_2 = \frac{1}{2} \sum_{j=1}^{J} \left( H_j^k - \hat{H}_j^k \right)^2 \delta_j^p,
\]
(8)
where \( p \in \{1,2,4,6,8,10\} \) is an index variable for six known grid points at the horizontal coordinate, and \( \hat{H}_j^k \) denotes the prior knowledge of true bottom terrain from these grid points.

The TV function \( J_3 \) is defined as
\[
J_3 = \sum_{j=1}^{J} \left| H_j^k - H_{j+1}^k \right|.
\]
(9)
If \( H(x) \) is discontinuous with jumps, then the TV function \( J_3 \) denotes the sum of magnitudes of the jumps; and if \( H(x) \) has many large amplitude oscillations, then the \( J_3 \) will become large. This is the property accounting for the effectiveness of the TV function as a regularization function. However, when \( H_j = H_{j+1} \), the \( J_3 \) will be not differentiable. An approximation in the form

\[
J_3 = \sum_{j=1}^{J-1} \sqrt{(H_j - H_{j+1})^2 + \beta^2}
\]  

(10)
is considered to avoid the difficulty, where \( \beta \) is a small positive parameter.

We consider the LBFGS method as the optimization algorithm to minimize the cost function (5). It begins with an initial guess of the control vector \( H \), and provides a sequence of improved bottom terrain until it reaches a solution. The strategy used to move from one iterate to the next makes use of the gradient of the cost function.

2.3. The gradient of the cost function

The gradient of the cost function (5) with respect to the control variable \( H \)

\[
\nabla_H J = \nabla_H J_1 + \mu(H_2 + \nabla_H J_3),
\]

(11)
is required to search for \( H \) that minimizes \( J \) under the constraints of model equations, the prior knowledge and the relative variation for spatially adjacent elevations.

The \( j \)th element of \( \nabla_H J_1 \in \mathbb{R}^{j+1} \) can be denoted as

\[
\frac{\partial J_1}{\partial H_j} = -g \frac{\Delta t}{2 \Delta x} \sum_{i \in I} \left[ (\varphi_{j-1}^k + \varphi_{j-1}^k)(\lambda_j^k + \lambda_{j+1}^k) - (\varphi_{j-1}^{k-1} + \varphi_{j-1}^{k+1})(\lambda_{j+1}^k + \lambda_j^k) \right],
\]

(12)

for \( j = 1, 2, \ldots, J \). The adjoint variables \( \lambda_j^k \) is determined by the adjoint model deduced from the cost function constrained by the discrete forward model (2a–c). The \( j \)th element of \( \nabla_H J_2 \) is

\[
\frac{\partial J_2}{\partial H_j} = (H_j - \tilde{H}_p) \delta_j^p,
\]

(13)

for \( j = 1, 2, \ldots, J \).

The TV function (10) can be rewritten as

\[
J_3 = \sum_{j=1}^{J-1} F_j(H_j, H_{j+1}),
\]

(14)

by defining \( \sqrt{(H_j - H_{j+1})^2 + \beta^2} \) as \( F_j(H_j, H_{j+1}) \). Then, the gradient of \( J_3 \) with respect to \( H \),

\[
\nabla_H J_3 = \left( \frac{\partial J_3}{\partial H_1}, \frac{\partial J_3}{\partial H_2}, \ldots, \frac{\partial J_3}{\partial H_J} \right)^T
\]

(15)
can be deduced from the first-order derivative of \( F_j(H_j, H_{j+1}) \) in the form

\[
\frac{\partial J_3}{\partial H_j} = \frac{\partial F_j(H_{j-1}, H_j)}{\partial H_j} + \frac{\partial F_j(H_j, H_{j+1})}{\partial H_j} = \frac{H_{j-1} - H_j}{\sqrt{(H_{j-1} - H_j)^2 + \beta^2}} + \frac{H_j - H_{j+1}}{\sqrt{(H_j - H_{j+1})^2 + \beta^2}},
\]

(16)

when \( j \in [2, J - 1] \). However, when \( j = 1 \), the \( \partial J_3 / \partial H_1 \) is denoted by the first-order derivative of \( F_1(H_1, H_2) \) with respect to \( H_1 \), i.e.,

\[
\frac{\partial J_1}{\partial H_1} = \frac{\partial F_1(H_1, H_2)}{\partial H_1} = \frac{H_1 - H_2}{\sqrt{(H_1 - H_2)^2 + \beta^2}},
\]

(17)
and analogously, when \( j = J \), the \( \partial J_j / \partial H_j \) can be represented by the first-order derivative of \( F_j(H_{j-1}, H_j) \) with respect to \( H_j \) in the form

\[
\frac{\partial J_j}{\partial H_j} = \frac{\partial F_{j-1}(H_{j-1}, H_j)}{\partial H_j} = -\frac{H_{j-1} - H_j}{\sqrt{(H_{j-1} - H_j)^2 + \beta^2}}.
\] (18)

Once the gradient has been determined, the descent direction, along which the cost function \( J \) is reduced, can be updated at each iterate by the LBFGS method [7].

3. Results

Matched bottom terrain should conform to numerical models. Comparing with the normal treatments of bottom terrain in numerical models, the results might be useful to reduce model errors if the optimal bottom terrain data could be obtained on the basis of the conventional meteorological observations and numerical models. The conventional observations for weather forecasts mainly consist of surface observations, but these data are usually inhomogeneous both in quality and quantity. This property might increase the level of difficulty in our test problem.

We perform the numerical experiments on a Core i3-2100 CPU, 3.10 GHz clock and 2 GB RAM memory. We introduce the LBFGS method with \( m = 3, 5 \) and 7 updates, where the parameter \( m \) represents the number of the last iterations. We use the northward wind field, contaminate the observations available at every two grid points and two time steps by adding random Gaussian noise with a mean of 0 m s\(^{-1}\) and standard deviation of 0.1 m s\(^{-1}\), and set the small positive parameter \( \beta \) of the TV function \( J_3 \) to \( 10^{-3} \) m. The maximum iteration number is 300. The initial guess is given with elevation at each grid equal to 0 m. The RMS error between the initial guess and the true terrain before correction is 16.33 m.

3.1. Using single regularization parameter

We introduce the Tikhonov-TV regularization term, \( \mu(J_2 + J_3) \), to induce stability and guarantee uniqueness, and the other two forms of the regularization cost function,

\[
J(\chi, H) = J_1 + \mu J_2,
\] (19)

and

\[
J(\chi, H) = J_1 + \mu J_3,
\] (20)

for the further investigation of the effect of the absolute and relative constraints, respectively, from \( J_2 \) and \( J_3 \). The regularization parameters can be found in Table 1.

| LBFGS method | The cost function J |
|--------------|---------------------|
|              | \( J_1 + \mu J_2 \) | \( J_1 + \mu J_3 \) | \( J_1 + \mu (J_2 + J_3) \) | \( J_1 + \mu_1 J_2 + \mu_2 J_3 \) |
| \( m = 3 \)  | \( 9.0 \times 10^{-3} \) | \( 9.0 \times 10^{-3} \) | \( 7.0 \times 10^{-3} \) | \( 9.0 \times 10^{-2}, 7.0 \times 10^{-3} \) |
| \( m = 5 \)  | \( 1.0 \times 10^{-2} \) | \( 5.0 \times 10^{-3} \) | \( 9.0 \times 10^{-3} \) | \( 5.0 \times 10^{-3}, 9.0 \times 10^{-4} \) |
| \( m = 7 \)  | \( 9.0 \times 10^{-3} \) | \( 3.0 \times 10^{-4} \) | \( 5.0 \times 10^{-2} \) | \( 5.0 \times 10^{-3}, 3.0 \times 10^{-4} \) |

The regularization terms contribute two improvements to the corrected bottom terrain, including (i) the reduced overall shift between the corrections and the true terrain, and (ii) the smooth bottom terrain surface. The first improvement mainly benefits from the prior constraint \( J_2 \). It is proved by the comparison between the corrections for the bottom terrain before adding regularization term (Figure 1a) and the corrections provided by the LBFGS codes with \( m = 5 \) and 7 updates, which are much closer to the true bottom (Figure 1b). However, only with this constraint, the minimization algorithms fail to reduce the fluctuations along the bottom terrain surface. These fluctuations are smoothed by only using the TV function \( J_3 \) (Figure 1c). The three minimization codes provide the bottom terrain corrections with smooth surface. Meanwhile, the defects of the application of only the relative...
constraint are obvious for the remaining large overall shift between the corrections and the true terrain (Figure 1c).

It seems that we can consider both the prior constraint $J_2$ and the TV function $J_3$ to maintain the two improvements. The application of the cost function (5) with Tikhonov-TV regularization terms do improve the corrections with reduced overall shift and smooth surface, but the improvements are too small and still unacceptable (Figure 1d). It should account for the small improvements that we consider the two regularization terms of equal importance by using the single regularization parameter $\mu$. However, the experiments using $J_1 + \mu J_2$ and $J_1 + \mu J_3$ help us to understand that the effects of the prior constraint and the TV function are different, and the separate regularization parameter should be considered.

Figure 1. The corrected bottom terrain at final time of assimilation using (a) $J_1$, (b) $J_1 + \mu J_2$, (c) $J_1 + \mu J_3$, and (d) $J_1 + \mu (J_2 + J_3)$.

3.2. Using dual regularization parameters
We rewrite the cost function (5) in the form
\[
J(\boldsymbol{x}, \boldsymbol{H}) = J_1 + \mu_1 J_2 + \mu_2 J_3, \tag{21}
\]
where $\mu_1 J_2$ and $\mu_2 J_3$ appear as dual constraints with two nonnegative regularization parameters $\mu_1$ and $\mu_2$. We investigate whether the proper choice for the values of the two regularization parameters could improve the quality of identification. We also determine the regularization parameters $\mu_1$ and $\mu_2$ using trials and errors, and the values can be found in Table 1.

The corrected bottom terrain benefits from the application of the Tikhonov-TV regularization function with dual regularization parameters. The corrections by the LBFGS codes with $m = 5$ and 7 updates have perfect shapes overlapping that of the true bottom terrain, while that by the LBFGS code
with $m = 3$ updates still exhibits an unacceptable shape (Figure 2a). The superior corrections by the LBFGS codes with $m = 5$ and 7 updates also appear in the variations in the values of the RMS error of the bottom terrain correction with the number of iterations, which are reduced by approximately 2 orders of magnitude (Figure 2b). After 300 iterations, the RMS errors of the bottom terrain respectively related with the LBFGS codes with $m = 5$ and 7 updates are $6.492 \times 10^{-1}$ m and $2.746 \times 10^{-1}$ m, about an order of magnitude smaller than that determined by using the cost function with single regularization parameter. In addition, the value of the cost function and the norm of the gradient are reduced by at least 4 and 2 orders of magnitude, respectively, after 300 iterations when we use the LBFGS codes with $m = 5$ and 7 updates (Figure 2c and d).

Figure 2. Terrain correction using the optimal dual regularization parameters and LBFGS codes with $m=3$, 5 and 7 updates. (a) The identified bottom terrain at final time of assimilation, (b) variations in the RMS errors, (c) variations in the value of the cost function and (d) variations in the norm of the gradient with the number of iterations.

Suppose that the eastward and northward wind fields and geopotential fields at the time $t_k$ are in the form, $u^k = (u^k_1, u^k_2, \ldots, u^k_J)^T$, $v^k = (v^k_1, v^k_2, \ldots, v^k_J)^T$ and $\phi^k = (\phi^k_1, \phi^k_2, \ldots, \phi^k_J)^T$, and we define wind field $w^k$ by combining eastward and northward wind fields $u^k$ and $v^k$ and then, the RMS error for wind field as $\sqrt{\|w^k - \tilde{w}^k\|^2 / J}$ and that for geopotential field $\sqrt{\|\phi^k - \tilde{\phi}^k\|^2 / J}$. $\tilde{w}^k$ and $\tilde{\phi}^k$ represent the solutions of the forward model with the true bottom terrain, and $w^k$ and $\phi^k$ represent the solutions of the forward model with the initial guess with terrain elevation at each grid equal to 0 m. The symbol $\|\|$ denotes the standard Euclidean norm. Mismatched bottom terrains do lead to large prediction errors,
especially for forward integration after $2.4 \times 10^4$ s. The maximum RMS error for wind field approaches 10 m s$^{-1}$, and that for geopotential field is over 100 m$^2$ s$^{-2}$.

The corrected bottom terrain at final time of assimilation can greatly improve the numerical predictions of the forward model (Figure 3). The LBFGS code with $m = 7$ updates provides the best corrections for bottom terrain, and the integration of the shallow-water equations model with these corrections generates the smallest prediction RMS errors of the wind and geopotential fields. The LBFGS codes with $m = 5$ updates also provide good corrections for bottom terrain and satisfactory predictions. The RMS errors for these predictions are slightly larger than those of the best predictions. The LBFGS code with $m = 3$ updates provides the worst corrections for bottom terrain and further the corresponding worst predictions.

![Figure 3](image.png)

**Figure 3.** Temporal variation in the RMS errors between the predictions using the true terrain and that using mismatched terrain and the optimal terrain at final time of assimilation. (a) the wind field and (b) the geopotential field.

4. Conclusions and discussions

We construct a variational adjoint assimilation system with Tikhonov-TV regularization to seek an optimal bottom terrain for the shallow-water equations model. The LBFGS codes provide unacceptable corrections with only data misfit function. After considering the absolute constraint from the prior knowledge of the true bottom terrain, expected improvements exist in the bottom corrections. Further using Tikhonov-TV regularization function with dual regularization parameters, the corrected bottom terrain improves significantly, especially by the LBFGS codes with $m = 5$ and 7 updates. These corrections also significantly improve the numerical predictions.

The ill-posed property of the problem under consideration strongly suggests the regularization technique. Regularization of this problem refers to solving the related regularized problem, which is a well-posed problem providing the physically meaningful solution to the given ill-posed problem [15]. The Tikhonov-TV regularization function, considering the property of the absolute and relative constraints, provides positive impact on improving bottom terrain corrections. The application of the absolute constraint from the prior knowledge of the true bottom terrain helps improving the corrections much more significantly comparing with the using of the relative constraints from the relative variation for spatially adjacent elevations, but the relative constraint provides positive effect on reducing the fluctuations and meanwhile, preserve the jump information of the mountain terrain. Martins [12] and Lima *et al.* [13] develop an inversion method with TV regularization and confirm that their method can delineate discontinuous basements presenting large slips or sequences of small slip step faults. The known depths and the TV function are treated as a combination regularization constraint with same weight. The same treatment in our study, i.e., the consideration of single regularization parameter, however, fails to correction the bottom terrain with acceptable RMS errors. This exactly suggests the dual regularization parameters because the effect of the absolute and relative constraints differs with each other. Tests by using Tikhonov-TV regularization function with dual
regularization parameters confirm the remarkable contribution of the application of dual regularization parameters.

Our tests confirm the potential of the framework for bottom terrain correction of numerical weather prediction models using variational adjoint method with Tikhonov-TV regularization. This framework, of course, is not confined to treat the problem of bottom terrain correction for numerical models. The most immediate use, for instance, might be to address the questions related to the shallow lakes, such as, the doubt whether lake-air parameterizations established for deep lakes are applicable to shallow lakes [16]. Whatever problems we face, most of all, we should take the different effects of the absolute and relative constraints into account by using different regularization parameters, if the relative variation for spatially adjacent grid points must be considered. Only then can we obtain acceptable corrections or identifications for the desired parameters, but the determination of the regularization parameter, especially the multiple regularization parameters, requires high computational cost. The applications of the techniques to achieve order reduction in the complicated assimilation frameworks form the main issue in our future studies.

Acknowledgements
This work was supported by the National Natural Science Foundation of China (41375115) and the Project funded by Norway for Improving Weather Information Management in East Africa for effective service provision through the application of suitable ICTs (UGA-13/0018).

References
[1] A.K. Griffith and N.K. Nichols, Adjoint methods in data assimilation for estimating model error, Flow Turbul. Combust. 65 (2000) 469-488.
[2] C.-J. Da et al., Dynamic modification of the shallow water equation on the slowly changing underlying surface condition. Acta Phys. Sin.-Ch. Ed. 62 (2013) (039202)1-6.
[3] T.T. Warner, Numerical Weather and Climate Prediction, Cambridge University Press, The Edinburgh Building, Cambridge CB2 8RU, UK, 2011.
[4] Y.-K. Sasaki, An objective analysis based on the variational method, J. Meteorol. Soc. JPN 2 (1958) 77-88.
[5] D. Liu and J. Nocedal, On the limited memory BFGS method for large scale optimization, Math. Program. 45 (1989) 503-528.
[6] J. Vlček and L. Lukšan, A conjugate directions approach to improve the limited-memory BFGS method, Appl. Math. Comput. 219 (2012) 800-809.
[7] Y.P. Wang, S.L. Tao and Q. Chen, Retrieving the variable coefficient for a nonlinear convection–diffusion problem with spectral conjugate gradient method, Inverse Probl. Sci. En. 23 (2015) 1342-1365.
[8] X. Zou et al., Numerical experience with limited-memory quasi-Newton and truncated Newton methods, SIAM J. Optimiz. 3 (1993) 582-608.
[9] Y.P. Wang and S.L. Tao, Application of regularization technique to variational adjoint method: a case for nonlinear convection–diffusion problem, Appl. Math. Comput. 218 (2011) 4475-4482.
[10] C.A. Parrett and M.J.P. Cullen, Simulation of hydraulic jumps in the presence of rotation and mountains, Q. J. Roy. Meteor. Soc. 110 (1984) 147-165.
[11] T.F. Chan and S. Esedoglu, Aspects of total variation regularized L1 function approximation, SIAM J. Appl. Math. 65 (2005) 1817-1837.
[12] C.M. Martins et al., Total variation regularization for depth-to-basement estimate: Part 1 Mathematical details and applications, Geophysics 76 (2011) I1-I12.
[13] W.A. Lima et al., Total variation regularization for depth-to-basement estimate: Part 2 Physicogeologic meaning and comparisons with previous inversion methods, Geophysics 76 (2011) I13-I20.
[14] E.L. Evans, J.P. Loveless and B.J. Meade, Total variation regularization of geodetically and geologically constrained block models for the Western United States, Geophys. J. Int. 202 (2015) 713-727.
[15] I.M. Navon, Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography, Dyn. Atmos. Oceans 27 (1997) 55-79.

[16] X. Lee et al., The Taihu eddy flux network: An observational program on energy, water, and greenhouse gas fluxes of a large freshwater lake, B. Am. Meteorol. Soc. 95 (2014) 1583-1594.