The spectral conjugate gradient method in variational adjoint assimilation for model terrain correction I: Theoretical frame

Sulin Tao¹ ² *, Yuhong Li³, Isaac Mugume⁴ and Shuanghe Shen¹

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 e-mail: sulin.tao@outlook.com

Abstract. A spectral conjugate gradient (SCG) method is proposed within the mathematical framework of a variational adjoint assimilation system to correct the bottom terrain of a shallow-water equations model. The formulation of this method is described from the mathematical point of view with determination of the descent direction by using Andrei’s limited-memory form and the step length by solving the tangent linear model. It is proved to benefit from (i) the iterative regularization strategy, (ii) the inverse Hessian approximation involving the second-order information from Broyden-Fletcher-Goldfarb-Shanno class (BFGS), and (iii) the optimal step length. The regularization term introduced to the cost function will allow the incorporation of the known information about the desired bottom terrain and guarantee the uniqueness of the optimal solution.

1. Introduction

A variational assimilation system coupled with an adjoint technique is an efficient means of dealing with inverse problems [1-5]. In this system, minimization of a cost function using descent algorithms constitutes an important numerical problem. The minimization procedure provides solutions that fit the observations by updating parameters iteratively, but it is computationally expensive due to the computation of the Hessian matrix as well as the backward integration of the adjoint model required to evaluate the gradients of the cost function. Fast and accurate algorithms are required to compute the minimizer for the function. Limited-memory optimization techniques are useful for solving large-scale unconstrained minimization problems because they have lower storage requirements and improve the computational stability and efficiency. Limited-memory methods build several rank-one or rank-two matrix updates to eliminate storage of the approximate Hessian matrix; thus, they require fewer function evaluations than the conjugate gradient method but do not require knowledge of the sparsity structure of the Hessian [6].
Nonlinear conjugate gradient (CG) methods provide regularization implicitly by neglecting non-dominant Hessian eigenvectors with low memory requirements and perfect convergence speed. The limited-memory variants of these methods are more powerful and straightforward due to the combination of adjoint equations [7-9] and the implementation of a preconditioned conjugate-gradient vector version of quasi-Newton updates, such as Shanno’s limited-memory form [10-12]. Several conjugate-gradient methods have been applied to large-scale unconstrained minimization problems in meteorology [9], and the Shanno-Phua [13] quasi-Newton memoryless conjugate-gradient technique has proven to be the most consistent. As an approximation of the memoryless quasi-Newton scheme of Perry and Shanno, Hager and Zhang’s CG_DESCENT method greatly outperforms the limited-memory Broyden-Fletcher-Goldfarb-Shanno (LBFGS) and other CG methods for unconstrained optimization problems [14, 15]. The high performance of this method was further investigated by comparing it to the quasi-Newton BFGS, LBFGS, truncated Newton method and hybrid algorithms for an ill-posed parameter estimation test problem [16]. This method is also competitive with other SCG methods because it uses the modified secant condition [17], and theoretically, the CG methods of Hager and Zhang account for the spectral scaling secant equation [18].

The SCG method using Andrei’s formulation [19] appears as a variant of the limited-memory CG method. It is modified from Birgin and Martinez’s scaled CG method [20] based on the inexact Wolfe line search conditions and an interpretation of the secant equation. The new descent direction is updated by using a scaling parameter determined by a spectral gradient or using an anticipative method based on the function values of two successive points and the inverse Hessian approximation, similar to the quasi-Newton BFGS updating expression to overcome the lack of positive definiteness of the matrix defining the search direction [19]. This approach has sufficiently accurate performance, and its perfect convergence rate was also studied by Wong and Xie [21]. This method converges more often for minimizations because of the consideration of the second-order information of the BFGS and the preserved feature of the classical CG method.

In this paper, we introduce a spectral conjugate gradient method within a framework formulated by the correction of the terrain parameter in a numerical weather prediction model from meteorological observations. Firstly, we introduce a one-dimensional shallow-water equations model with rotation to conduct our test problem. Secondly, we propose detailed numerical schemes of the variational adjoint assimilation system including the forward problem, the sensitivity problem, the adjoint problem, the cost functional and its gradient. Finally, we provide the formulation of the SCG method with determination of the descent direction by using Andrei’s limited-memory form [19] and the step length by solving tangent linear model.

2. The test problem

We conduct our test problem by using a one-dimensional shallow-water equations model with rotation and bottom topography describing the motion of a single-layer atmospheric flow [22]. The height of the bottom terrain and the depth of atmospheric flow above the terrain are defined as $H = H(x)$ and $\eta(x,t) > 0$, respectively (Figure 1). Our test problem supposes that the accurate description of bottom topography in the shallow-water equations model was unknown. Thus, we try to seek an optimal terrain parameter for an accuracy improvement of numerical model.

2.1. The forward model

The forward model can be stated as [23]

$$m^k_j = m^{k-2}_j - \frac{\Delta t}{2\Delta x^2} \left[ (u^{k-1}_{j+1} + u^{k-1}_j) (m^{k-1}_{j+1} + m^{k-1}_j) - (u^{k-1}_j + u^{k-1}_{j-1}) (m^{k-1}_j + m^{k-1}_{j-1}) + (\phi^{k-1}_{j+1})^2 \right]$$

$$- \left( \phi^{k-1}_{j-1} \right)^2 - g \frac{\Delta t}{2\Delta x} \left[ (\phi^{k-1}_{j+1} + \phi^{k-1}_j) (H_{j+1} - H_j) + (\phi^{k-1}_j + \phi^{k-1}_{j-1}) (H_j - H_{j-1}) \right]$$

$$+ 2\Delta tf n^{k-1}_j + 2 \frac{\Delta t}{\Delta x^2} K \left( m^{k-2}_{j+1} - 2m^{k-2}_j + m^{k-2}_{j-1} \right) \right] \quad (1a)$$
\[ n_j^k = n_j^{k-2} - \frac{\Delta t}{2\Delta x} \left[ (v_{j+1}^{k-1} + v_j^{k-1}) (m_{j+1}^{k-1} + m_j^{k-1}) - (v_j^{k-1} + v_{j-1}^{k-1}) (m_j^{k-1} + m_{j-1}^{k-1}) \right] \]

\[ -2\Delta tfm_j^{k-2} + 2 \frac{\Delta t}{\Delta x^2} K \left( n_{j+1}^{k-2} - 2n_j^{k-2} + n_{j-1}^{k-2} \right), \]  

\[ \phi_j^k = \phi_j^{k-2} - \frac{\Delta t}{\Delta x} \left( m_{j+1}^{k-1} - m_j^{k-1} \right) + 2 \frac{\Delta t}{\Delta x^2} K \left( \phi_{j+1}^{k-2} - 2\phi_j^{k-2} + \phi_{j-1}^{k-2} \right), \]  

The periodic boundary conditions in the form

\[ u_0^k = u_{J}^k \quad \text{and} \quad v_0^k = v_{J}^k, \quad \phi_0^k = \phi_{J}^k, \quad \phi_0^{j+1} = \phi_{j+1}^{k}, \quad \phi_{J-1}^k = \phi_{J}^k, \]  

are considered for \( k = 0,1,\ldots,N \). We take the Coriolis parameter

\[ f = 7.292 \times 10^{-5} \text{ s}^{-1}, \]

diffusion coefficient \( K = 5.0 \times 10^2 \text{ m}^2 \text{ s}^{-1} \), gravitational acceleration \( g = 9.8 \text{ m s}^{-2} \), domain length \( L = 9.55 \times 10^5 \) m, and time length \( T = 1.2 \times 10^5 \) 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 (1a–c) are defined as follows:

\[ u_j^0 = 0 \text{ m s}^{-1}, \quad v_j^0 = 0 \text{ m s}^{-1} \text{ and } \phi_j^0 = \phi_m, \]  

with \( \phi_m = g\eta_m \), where the mean depth of atmospheric flow \( \eta_m = 1 \) m, and \( j = 1,2,\ldots,J \). The true bottom terrain with the ridge shape \( H_j \) is

\[ 1/2 - \left( x_j - \pi L \right)^2 / (2a^2) \]  

m when \( x_j \in [\pi L - a, \pi L + a] \) and 0 m when \( x_j \in [0, \pi L - a) \cup (\pi L + a, 2\pi L] \), where \( a = 10\Delta x \) and \( j = 1,2,\ldots,J \). Detailed description of the forward model can be found in our previous work [23].

2.2. The cost function and its gradient

We set the wind and geopotential field vectors at the time \( t_k \) as

\[ \mathbf{u}^k = \left( u_1^k, u_2^k, \ldots, u_J^k \right)^T, \quad \mathbf{v}^k = \left( v_1^k, v_2^k, \ldots, v_J^k \right)^T, \quad \Phi^k = \left( \phi_1^k, \phi_2^k, \ldots, \phi_J^k \right)^T \in \Phi^{J+1}. \]  

where \( T \) denotes transposition. Then we define the state vector \( \chi^k \) that combines \( \mathbf{u}^k \), \( \mathbf{v}^k \) and \( \Phi^k \) as

\[ \chi^k = \begin{pmatrix} \mathbf{u}^k \\ \mathbf{v}^k \\ \Phi^k \end{pmatrix} \in H^{J+1}, \]  

and the control vector \( \mathbf{H} \) in the form

\[ \begin{aligned}
\chi^k & = \begin{pmatrix}
\mathbf{u}^k \\
\mathbf{v}^k \\
\Phi^k 
\end{pmatrix} \in H^{J+1}, \\
\mathbf{H} & = \begin{pmatrix}
\mathbf{u}_0^k \\
\mathbf{v}_0^k \\
\Phi_0^k 
\end{pmatrix} \in H^{J+1},
\end{aligned} \]  

Figure 1. The test problem.
The best estimation of the terrain results in seeking a value of \( H^* \) that minimizes a cost function, defined as the weighted sum of the squared discrepancies between the observations, \( \bar{u}^k, \bar{v}^k \) and \( \bar{\phi}^k \in \phi^{inv} \), and the model predictions, \( u^k, v^k \) and \( \phi^k \in \phi^{inv} \), which can be given by

\[
J(\chi, H) = \frac{1}{2} \sum_{t=0}^{N} \left[ \alpha \left( u^k - \bar{u}^k \right)^T \left( u^k - \bar{u}^k \right) + \beta \left( v^k - \bar{v}^k \right)^T \left( v^k - \bar{v}^k \right) + \gamma \left( \phi^k - \bar{\phi}^k \right)^T \left( \phi^k - \bar{\phi}^k \right) \right].
\] (7)

\( \alpha, \beta, \) and \( \gamma \) are weighting factors (or trade-off coefficients) that determine the contributions of the eastward wind, northward wind and geopotential fields, respectively, to the cost function during the minimization procedures. These factors are also used:

(1) to scale the cost function, making it dimensionless, and
(2) to scale the cost function with velocity and geopotential terms of the same order of magnitude.

Thus, we determine the values of the weighting factors according to the maximum square difference between the analysed fields at the initial time \( t_0 \) and that at the final time \( t_N \) using the following formulae:

\[
\alpha = \eta / \max_{j} \left[ \bar{u}^j - \bar{u}^j_N \right]^2, \quad \beta = \eta / \max_{j} \left[ \bar{v}^j - \bar{v}^j_N \right]^2, \quad \gamma = \eta / \max_{j} \left[ \bar{\phi}^j - \bar{\phi}^j_N \right]^2.
\] (8)

where

\[
\eta = \min \left[ \max_{j} \left[ \bar{u}^j - \bar{u}^j_N \right]^2, \max_{j} \left[ \bar{v}^j - \bar{v}^j_N \right]^2, \max_{j} \left[ \bar{\phi}^j - \bar{\phi}^j_N \right]^2 \right].
\] (9)

In this way \( \alpha, \beta, \) and \( \gamma \) are found to be 1 m\(^2\) s\(^2\), \(7.23 \times 10^{-2}\) m\(^2\) s\(^2\), and \(1.22 \times 10^{-1}\) m\(^4\) s\(^4\), respectively.

The test inverse problem of determining the bottom terrain from observations is a typical example of an ill-posed mathematical problem, suggesting the application of a regularization technique \([23-25]\) to ensure a well-posed inverse problem.

We consider adding a regularization term to the cost function (7) to allow the incorporation of the known information about the desired bottom terrain and guarantee the uniqueness of the optimal solution. In practice the bottom terrain data required by numerical weather prediction models are generated by interpolating and smoothing based on the mismatched terrain (usually the satellite observations). For a specific modeling region, the mean elevation, as a general feature of the bottom terrain, might be obtained from these mismatched terrain data. We thus penalize the cost function (7) by adding a quadratic penalty term and obtain a regularized cost function:

\[
J_{r}(\chi, H) = J(\chi, H) + \frac{\tau}{2} \left( e^T (H - H_b) \right)^2,
\] (10)

where the first term \( J(\chi, H) \) is defined previously as Eq. (7). The second term describes a constraint involving known average elevation of the desired bottom terrain, i.e., the regularization term, which forces the proximity between the corrected \( H \) and known average terrain elevation \( H_b \) (= 6.68 \times 10^{-2} \) m for the true bottom terrain with a ridge shape). \( e \) is a \( J \times 1 \) vector of ones. \( \tau \) is a nonnegative regularization parameter. This regularization term eliminates ill-posedness of the original problem. It penalizes the deviations of bottom terrain from the true one selectively on the null space of the original problem. The method to determine this parameter is proposed in part II of this paper.

We provide the SCG method as an optimization algorithm to minimize the cost function (10). It begins with an initial guess of the control vector \( H \) and provides a sequence of improved bottom terrains until it reaches a solution. The strategy used to move from one iteration to the next makes use of the search step length and the search direction. The step length is determined using the tangent linear model, which can be found in the Appendix A. The search direction is formulated using the gradient of the cost function, which can be computed with adjoint variables provided by an adjoint model. The derivation of this adjoint model can be found in the Appendix B and the evaluation of its accuracy is provided in part II of this paper.
We define the gradient of the cost function $J(\chi, H)$ with respect to the control variable $H$ as $\nabla_H J$. For the cost function (7) the $j$th element of $\nabla_H J$ is then determined by

$$\frac{\partial J}{\partial H_j} = -g \frac{\Delta \sum_{k=1}^N \left[ (\phi_{j,k}^{k-1} + \phi_{j,k}^{k+1})(\lambda_{j,k}^k + \lambda_{j,k}^{k+1}) - (\phi_{j,k+1}^{k} + \phi_{j,k-1}^{k})(\lambda_{j,k}^k + \lambda_{j,k}^{k+1}) \right]}{2\Delta x},$$

with $j = 1, 2, \cdots, J$. Note that $\frac{\partial J}{\partial H_1}$ and $\frac{\partial J}{\partial H_J}$ are computed using the cyclic conditions $\lambda_{j,1}^{k-1} = \lambda_{j,1}^{k+1}$ and $\phi_{j,1}^{k-1} = \phi_{j,1}^{k+1}$ and the cyclic conditions $\lambda_{j,N}^{k-1} + \phi_{j,N}^{k-1} = \lambda_{j,N}^{k+1} + \phi_{j,N}^{k+1}$, respectively. Then for the cost function (10) the $j$th element of $\nabla_H J_t$ is updated by

$$\frac{\partial J_t}{\partial H_j} = \frac{\partial J}{\partial H_j} + \tau \left( \frac{\epsilon^t H - H_b}{J} \right).$$

This gradient is applied to generate the descent direction along which the cost function can be reduced.

3. The spectral conjugate gradient method

In order to get an optimal estimate of bottom terrain $H^t$, we consider the recurrence equation

$$H^{t+1} = H^t + l^t d^t,$$

where $d^t \in d^{t+1}$ is the descent direction at the $i$th (the number of iterations $i = 0, 1, 2, \cdots$) iteration and $l^t$ is the corresponding step length.

Firstly, we compute the optimal step length $l^t$ at the $i$th iteration by solving this problem:

$$l^t = \arg \min \left\{ J_i(\chi^t, H^t + l^t d^t) \right\}.$$

We define the perturbation vectors of eastward wind field, northward wind field, and geopotential field, respectively, as

$$\Delta u^t = (\Delta u_1^t, \Delta u_2^t, \cdots, \Delta u_N^t)^T, \Delta v^t = (\Delta v_1^t, \Delta v_2^t, \cdots, \Delta v_N^t)^T, \Delta \phi^t = (\Delta \phi_1^t, \Delta \phi_2^t, \cdots, \Delta \phi_N^t)^T,$$

and the perturbation vector of bottom terrain as

$$\Delta H = (\Delta H_1, \Delta H_2, \cdots, \Delta H_J)^T.$$

The step length $l^t$ can be achieved with the outputs from the tangent linear model. Using a first-order Taylor series approximation for $u^t(\Delta H_1^t + l^t d^t)$, $v^t(\Delta H_2^t + l^t d^t)$ and $\phi^t(\Delta H_3^t + l^t d^t)$ and specifying the value of $d^t$ with that of $\Delta H^t$ (ignoring the difference in their units), we can obtain

$$u^t(\Delta H_1 + l^t d^t) \approx u^t(\Delta H_1^t) + l^t \Delta u_1^t, v^t(\Delta H_2 + l^t d^t) \approx v^t(\Delta H_1^t) + l^t \Delta v_1^t, \phi^t(\Delta H_3 + l^t d^t) \approx \phi^t(\Delta H_1^t) + l^t \Delta \phi_1^t.$$

Note that in Eq. (17) we should treat the step length $l^t$ as a unitless scalar but in fact $l^t$ has the unit m$^2$.

We then minimize the cost function $J_i(\chi^t, H^t + l^t d^t)$ with respect to $l^t$ and thus obtain the step length

$$l^t = -\left[ \sum_{k=1}^N (\Delta u_k^t)^T W_u^t (u_k^t - \bar{u}_k^t) + (\Delta v_k^t)^T W_v^t (v_k^t - \bar{v}_k^t) + (\Delta \phi_k^t)^T W_{\phi}^t (\phi_k^t - \bar{\phi}_k^t) \right]$$

$$+ \tau \left[ \frac{\epsilon^t d^t}{J} \left( \frac{\epsilon^t H - H_b}{J} \right) \right] + \tau \left[ \frac{\epsilon^t d^t}{J} \right] + \tau \left( \frac{\epsilon^t d^t}{J} \right)^2,$$

where the variables $\Delta u^t$, $\Delta v^t$ and $\Delta \phi^t \in H^{t+1}$ are calculated via the tangent linear model which can be found in the Appendix A.
Secondly, we introduce the spectral conjugate gradient method by using a BFGS-like formula to compute the descent direction. Unlike the LBFGS algorithm, the SCG method only requires information in the last two iterations and uses a restart strategy to ensure a descent direction when the step length goes beyond the strong Wolf condition. This method formulates the descent direction \( d_i \) by the application of the spectral gradient as a scalar parameter.

Let \( s_i = H_i - H_{i-1} \) and \( y_i = \nabla u_i J - \nabla u_{i-1} J \); the spectral gradient can be of the form

\[
\theta^i = \left[ (s_{i-1})^T s_{i-1} \right] \left[ (y_{i-1})^T y_{i-1} \right],
\]

where \((y_{i-1})^T s_{i-1} > 0\). For \( i = 0 \), the descent direction \( d_i \) is updated by

\[
d_0 = -\nabla u_0 J_0,
\]

and for \( i \geq 1 \), this direction is defined by

\[
d_i = -\theta^i \nabla u_i J_i + \theta^i \left( \nabla u_i J_i \right)^T s_{i-1} y_{i-1} - \left( \nabla u_i J_i \right)^T s_{i-1} s_{i-1} - \theta^i \left( y_{i-1} \right)^T y_{i-1} \left( \nabla u_i J_i \right)^T s_{i-1} s_{i-1} + \theta^i \left( y_{i-1} \right)^T s_{i-1} y_{i-1} s_{i-1} \]

(21).

However, if \((y_{i-1})^T s_{i-1} \leq 0\), the direction is generated by

\[
d_i = -\nabla u_i J_i.
\]

Furthermore, when the angle between the current direction \( d_i \) and \(-\nabla u_i J\) is not acute enough, i.e., if the inequality \((d_i)^T \nabla u_i J > -10^{-3} \|d_i\| \|\nabla u_i J\|\) is satisfied, we “restart” the algorithm using the direction given by Eq. (22). The computational algorithm of this SCG method can be expressed as in Figure 2.

**Figure 2.** Flow chart of the SCG method.
Remark 1. The SCG method always provides a descent direction.
We consider the line search subjected to the strong Wolfe conditions
\[
J'(H^i - l'd') - J'(H^i) \leq c_1 l' \left( \nabla_{H^i} J \right)^T d' \quad \text{and} \quad \left[ \nabla_{H^i} J (H^i - l'd') \right]^T d' \leq c_2 \left( \nabla_{H^i} J \right)^T d',
\]
where \( 0 < c_1 < c_2 < 1 \), then have \( (y^i)^T s^i > 0 \). Thus for \( i = 0 \), multiplying the direction (20) by \( \left( \nabla_{H^i} J, i \right)^T \), we have
\[
\left( \nabla_{H^i} J, i \right)^T d^i = -\left\| \nabla_{H^i} J, i \right\|^2 \leq 0; \quad (24)
\]
and for \( i \geq 1 \), multiplying the direction (21) by \( \left( \nabla_{H^i} J, i \right)^T \), we have
\[
\left( \nabla_{H^i} J \right)^T d^i = \left\{ -\theta \left[ \nabla_{H^i} J \right]^T \left[ (y^{i-1})^T s^{i-1} \right]^2 - \left[ (\nabla_{H^i} J)^T s^{i-1} \right]^2 \left[ (y^{i-1})^T s^{i-1} \right] - \theta \left( (y^{i-1})^T y^{i-1} \right) \right\}
\]
\[
+ \theta \left[ (\nabla_{H^i} J)^T s^{i-1} y^{i-1} \right] \left[ (\nabla_{H^i} J)^T s^{i-1} y^{i-1} \right] - \left[ (\nabla_{H^i} J)^T s^{i-1} \right]^2 \left[ (y^{i-1})^T s^{i-1} \right] \left[ (y^{i-1})^T s^{i-1} \right] \right\} \right)^2 \left[ (y^{i-1})^T s^{i-1} \right],
\]
Defining
\[
a = \left[ (s^{i-1})^T y^{i-1} \right] \nabla_{H^i} J \quad \text{and} \quad b = \left[ (\nabla_{H^i} J)^T s^{i-1} \right] y^{i-1},
\]
and according to the inequality
\[
a^T b \leq \frac{1}{2} \left( \|a\|^2 + \|b\|^2 \right), \quad (27)
\]
we have
\[
\left( \nabla_{H^i} J \right)^T d^i \leq \left\{ \begin{array}{c}
-\theta \left[ \nabla_{H^i} J \right]^T \left[ (y^{i-1})^T \right]^2 + \theta \left[ (s^{i-1})^T y^{i-1} \nabla_{H^i} J \right]^2 \\
+ \theta \left[ (\nabla_{H^i} J)^T s^{i-1} y^{i-1} \right] \left[ (\nabla_{H^i} J)^T s^{i-1} y^{i-1} \right] - \left[ (\nabla_{H^i} J)^T s^{i-1} \right]^2 \left[ (y^{i-1})^T s^{i-1} \right]^2 \\
- \theta \left[ (y^{i-1})^T s^{i-1} \right] \left[ (\nabla_{H^i} J)^T s^{i-1} \right] \left[ (y^{i-1})^T s^{i-1} \right] \right\} \right)^2 \left[ (y^{i-1})^T s^{i-1} \right],
\]
\[
< 0.
\]
Moreover, when \( (y^{i-1})^T s^{i-1} \leq 0 \), i.e. the line search doesn’t satisfy the strong Wolfe conditions, we also have
\[
\left( \nabla_{H^i} J, i \right)^T d^i = -\left\| \nabla_{H^i} J, i \right\|^2 \leq 0, \quad (29)
\]
for \( i \geq 0 \). Therefore, the direction formulated by the SCG method we introduced is a descent direction for every \( i = 0, 1, 2, \ldots \).

Remark 2. The SCG method inherits the second-order information from BFGS and computes a new descent direction \( d' \) using only the current vectors \( H^i, \nabla_{H^i} J \) and the previous vectors \( H^{i-1}, \nabla_{H^{i-1}} J \).
It preserves the features of the classical nonlinear CG methods and provides regularization implicitly by neglecting non-dominant Hessian eigenvectors, ensuring low memory requirements and perfect convergence speed. In this algorithm, at each iteration, we follow Andrei’s update formula to define the inverse Hessian approximation matrix as \( Q' \):
\[ Q^i = \theta I - \theta \frac{y^{i-1}(s^{i-1})^T}{(y^{i-1})^T s^{i-1}} - \theta \frac{(y^{i-1})^T s^{i-1}}{(y^{i-1})^T s^{i-1}} + \theta \frac{y^{i-1} y^{i-1}^T s^{i-1} (s^{i-1})^T}{(y^{i-1})^T s^{i-1}} + \frac{s^{i-1} (s^{i-1})^T}{(y^{i-1})^T s^{i-1}}, \] (30)

which satisfies \( Q^i y^{i-1} = s^{i-1} \), where \( I \) is the identity matrix. This updating expression indicates the involvement of the second-order information from BFGS and further implies a quasi-Newton BFGS method in comparison with the best update of the Broyden class \( Q^{*,i} \):

\[ Q^{*,i} = Q^{-1,*} - \frac{Q^{-1,*} y^{i-1} (s^{i-1})^T}{(y^{i-1})^T s^{i-1}} - \frac{(y^{i-1})^T s^{i-1} Q^{-1,*}}{(y^{i-1})^T s^{i-1}} + \frac{(y^{i-1})^T (s^{i-1})^T}{(y^{i-1})^T s^{i-1}} + \frac{(s^{i-1})^T}{(y^{i-1})^T s^{i-1}}, \] (31)

where \( \rho^{*-1} = 1/\left[(y^{i-1})^T s^{i-1}\right] \) and \( V^{i-1} = I - \rho^{*-1} y^{i-1} (s^{i-1})^T \). The LBFGS method requires the storage of the vectors \( y^i \) and \( s^i \) obtained in the most recent \( m \) (e.g., \( = 3, 5, \) or 7) iterations to construct an approximation matrix of the inverse Hessian:

\[ Q^{*,i} = \left[ (V^i)^T \cdots (V^{i-n})^T \right] Q^{0} \left( (V^{i-n})^T \cdots (V^i) \right) + \rho^{i-n} \left[ (V^i)^T \cdots (V^{i-n+2})^T \right] s^{i-n} \left( s^{i-n+1}^T (V^{i-n+1}) \cdots (V^i) \right) \]
\[ + \rho^{i-n-1} \left[ (V^i)^T \cdots (V^{i-n+2})^T \right] s^{i-n} \left( s^{i-n+1}^T (V^{i-n+2}) \cdots (V^i) \right) + \cdots + \rho s^i \left( s^i \right)^T, \] (32)

for \( i \geq m \), where \( m = m - 1 \), and \( Q^{0} = \left[ \left( y^i \right)^T \cdot s^i \right]/\left( y^i \right)^T \cdot s^i . \) When \( i < m \), the approximation matrix is defined by Eq. (31). However, the SCG method updates the inverse Hessian approximation \( Q^i \) with the identity matrix multiplied by the scalar \( \theta \) and requires only the information of the current vectors \( H^i, V^i, J^i \) and the previous vectors \( H^{i-1}, V^{i-1}, \) and \( J^{i-1} \) at every iteration. This trait is believed to be one of the reasons for the superior performance of the SCG method. In addition, the BFGS updating formula is known to have very effective self-correcting properties. If the matrix \( Q^{*,i} \) incorrectly estimates the curvature in the cost function and if this bad estimate slows the iteration, then the inverse Hessian approximation will tend to correct itself within a few steps [26].

Then, we can start the iteration from an initial guess of the bottom terrain to seek the optimal result by considering the stopping criterion for the recurrent process, which should meet one of the following three rules: (i) the relative change of the cost function value from one iteration to the next satisfies \( \left| J_x \left( H^i \right) - J_x \left( H^{i+1} \right) \right| < 10^{-8} \); (ii) the gradient of the cost function satisfies \( \left\| V^i H^x J_x \right\| \leq 10^{-4} \left\| V^i H^x J_x \right\| \); or (iii) the maximum number of iterations is reached. The symbols \( \| \) and \( || \) denote the absolute value and the standard Euclidean norm, respectively.

The whole computational procedure of the SCG method with the optimal step length is as follows.

**Step 1.** Choose an initial guess \( H^0 \) and set the number of iterations \( i = 0 \).

**Step 2.** Solve the forward model (1a–c) with initial values of the states \( u^0, v^0, \) and \( \Phi^0 \) and initial guess \( H^0 \) to compute the model states from \( t_0 \) to \( t_N \).

**Step 3.** Check whether \( J_x \left( H^0 \right) \) is smaller than the tolerance \( \varepsilon = 10^{-8} \); continue if not.

**Step 4.** Solve the adjoint model (39a–c) to generate the adjoint vectors \( \lambda = (\lambda_1, \lambda_2, \cdots, \lambda_J)^T \in \lambda^{J \times 1}, \mu = (\mu_1, \mu_2, \cdots, \mu_J)^T \in \mu^{J \times 1}, \) and \( \sigma = (\sigma_1, \sigma_2, \cdots, \sigma_J)^T \in \sigma^{J \times 1} \) from \( t_N \) to \( t_0 \).

**Step 5.** Calculate the gradient of the cost function \( V^i H^x J_x \) as in Eq. (12).
Step 6. Compute the descent direction \( d^i \) using Eq. (20) when \( i = 0 \) and, otherwise, using Eq. (21) when \( (y^{i-1})^T s^{i-1} > 0 \) or Eq. (22) when \( (y^{i-1})^T s^{i-1} \leq 0 \). “Restart” the algorithm using the direction given by Eq. (22) if the inequality \( (d^i)^T V_H J_i > -10^{-3} \|d^i\|_2 \|V_H J_i\|_2 \) is satisfied.

Step 7. Set \( \Delta H^i = d^i \), and solve the tangent linear model (33a–c) to obtain \( \Delta u^i \), \( \Delta v^i \), and \( \Delta \phi^i \) from \( t_0 \) to \( t_N \) and compute the step length \( l^i \).

Step 8. Update the new estimate \( H^{i+1} = H^i + l^i d^i \) with the descent direction \( d^i \) and step length \( l^i \).

Step 9. Check the stopping criterion. If it is satisfied, stop. Otherwise, set \( i = i + 1 \) and go to step 4.

4. Conclusion

We provide a SCG method within the mathematical framework of a variational adjoint assimilation system to correct the bottom terrain of a shallow-water equations model. We describe the formulation of this method from the mathematical point of view with determination of the descent direction by using Andrei’s limited-memory form and the step length by solving the tangent linear model. It is proved to benefit from (i) the iterative regularization strategy, (ii) the inverse Hessian approximation involving the second-order information from BFGS, and (iii) the optimal step length. The regularization term introduced to the cost function will allow the incorporation of the known information about the desired bottom terrain and guarantee the uniqueness of the optimal solution. We will evaluate the performance of the SCG method by comparison with that of the LBFGS method [6] and the CG_DESCENT method [14] in the related numerical tests in part II of this paper.

5. Acknowledgments

We thank the contribution of Prof. Yuepeng Wang of the School of Mathematics and Statistics, Nanjing University of Information Science & Technology. This work is supported by the National Natural Science Foundation of China (41375115), National Key Technology Research and Development Program of the Ministry of Science and Technology of China (2012BAC23B01), and a project funded by Norway, Improving Weather Information Management in East Africa for Effective Service Provision through the Application of Suitable ICTs (UGA-13/0018).

Appendix

A. The tangent linear model

We consider a numerical solution of Eqs. (1a–c) with initial conditions Eq. (3) and a prescribed terrain elevation \( H \). Infinitely small perturbation \( \Delta H \) added to the terrain \( H \) will cause perturbations in the solutions \( \Delta u \), \( \Delta v \) and \( \Delta \phi \). To derive the tangent linear model equations describing evolutions of \( \Delta u \), \( \Delta v \) and \( \Delta \phi \), we write down the forward equations for perturbed variables, subtract equations for unperturbed variables and keep the terms that are linear with respect to the perturbations.

\( \Delta u^i \), \( \Delta v^i \) and \( \Delta \phi^i \) are then obtained using the discrete schemes as follows:
$$\Delta u^j_j \phi^k_j + u^j_j \Delta \phi^k_j = \Delta u^{k-2}_j \phi^{k-2}_j + u^{k-2}_j \Delta \phi^{k-2}_j - \Delta t / (2\Delta x) \left[ (\Delta u^{k-1}_j \phi^{k-1}_j + u^{k-1}_j \Delta \phi^{k-1}_j) + (u^{k-1}_j + u^{k}_j) \left( u^{k-1}_j \phi^{k-1}_j + u^{k}_j \phi^{k-1}_j \right) \right] - \Delta t / (2\Delta x) \left[ (\Delta u^{k-1}_j \phi^{k-1}_j + u^{k-1}_j \Delta \phi^{k-1}_j) + (u^{k-1}_j + u^{k}_j) \left( u^{k-1}_j \phi^{k-1}_j + u^{k}_j \phi^{k-1}_j \right) \right] - 2 u^{k}_j \phi^{k-1}_j \Delta \phi^{k-1}_j$$

$$\Delta v^j_j \phi^k_j + v^j_j \Delta \phi^k_j = \Delta v^{k-2}_j \phi^{k-2}_j + v^{k-2}_j \Delta \phi^{k-2}_j - \Delta t / (2\Delta x) \left[ (\Delta v^{k-1}_j \phi^{k-1}_j + v^{k-1}_j \Delta \phi^{k-1}_j) + (v^{k-1}_j + v^{k}_j) \left( v^{k-1}_j \phi^{k-1}_j + v^{k}_j \phi^{k-1}_j \right) \right] - \Delta t / (2\Delta x) \left[ (\Delta v^{k-1}_j \phi^{k-1}_j + v^{k-1}_j \Delta \phi^{k-1}_j) + (v^{k-1}_j + v^{k}_j) \left( v^{k-1}_j \phi^{k-1}_j + v^{k}_j \phi^{k-1}_j \right) \right] - 2 v^{k}_j \phi^{k-1}_j \Delta \phi^{k-1}_j + \Delta \phi^{k}_j$$

$$\Delta \phi^k_j = \Delta \phi^{k-2}_j - \Delta t / (2\Delta x) \left[ (\Delta u^{k-1}_j \phi^{k-1}_j + u^{k-1}_j \Delta \phi^{k-1}_j) + (u^{k-1}_j + u^{k}_j) \left( u^{k-1}_j \phi^{k-1}_j + u^{k}_j \phi^{k-1}_j \right) \right] + \Delta \phi^{k}_j$$

for $k = 2,3,\cdots,N$, and $j = 1,2,\cdots,J$. $u^j_j$, $\Delta v^j_j$, and $\Delta \phi^j_j$ are obtained with the initial values $\Delta u^0_j = 0$, $\Delta v^0_j = 0$, and $\Delta \phi^0_j = 0$, respectively, by replacing the terms $\Delta u^{k-2}_j$, $\Delta v^{k-2}_j$, and $\Delta \phi^{k-2}_j$ in Eqs. (33a–c) with $\Delta u^{k-1}_j$, $\Delta v^{k-1}_j$, and $\Delta \phi^{k-1}_j$ when $k = 1$.

We introduce a treatment similar to that considered in the forward model by defining the periodic boundary conditions for these Eqs. (33a–c) as follows:

$$\begin{cases}
\Delta u^0_j = \Delta u^J_j, & \Delta v^0_j = \Delta v^J_j, \quad \Delta \phi^0_j = \Delta \phi^J_j, \\
\Delta u^{k-1}_j = \Delta u^{k+1}_j, & \Delta v^{k-1}_j = \Delta v^{k+1}_j, \quad \Delta \phi^{k-1}_j = \Delta \phi^{k+1}_j,
\end{cases}$$

for $k = 0,1,\cdots,N$. Similar to the cyclic conditions for $H_j$, $\Delta H_j$ should obey cyclic conditions $\Delta H_0 = \Delta H_J$ and $\Delta H_{J+1} = \Delta H_1$. The eastward wind perturbation $\Delta u^k_j$, northward wind perturbation $\Delta v^k_j$, and geopotential perturbation $\Delta \phi^k_j$ are calculated via Eqs. (33a–c) when $\Delta H_j$ is updated by the descent direction introduced in Section 3.

B. The adjoint model and accuracy evaluation

We deduce the adjoint model from the cost function $J(\chi, \mathbf{H})$ constrained by the discrete forward model (1a–c). We define
\[ U_j^i = \frac{m_j^k - m_{j-1}^{k-2} + \Delta t / (2\Delta x) \left[ (u_{j+1}^i + u_j^i) \left( m_{j+1}^{k-1} + m_j^{k-1} \right) - (u_{j-1}^i + u_j^i) \left( m_{j-1}^{k} + m_j^{k-1} \right) \right] + (\varphi_{j+1}^i - \varphi_{j-1}^i)^2}{(H_j - H_{j-1})} \] 

\[ + \frac{g\Delta t}{(2\Delta x) \left[ (\varphi_{j+1}^i + \varphi_{j-1}^i) \left( H_{j+1} - H_j \right) \right] + (\varphi_{j+1}^i + \varphi_{j-1}^i), \] (35a)
We further let $\partial\mathcal{L} / \partial u_j^k = 0$ and obtain the adjoint equation (39a). We can obtain the other two adjoint equations (39b, c) in a similar manner.

$$\begin{align*}
\lambda_j^k \phi_j^k &= -\lambda_j^{k+2} \phi_j^{k+2} - \Delta t / (2\Delta x) \left[ \lambda_j^{k+1} (u_j^{k+1} \phi_j^{k+1} + u_{j+1}^{k+1} \phi_{j+1}^{k+1}) + \lambda_j^{k+1} (u_j^{k+1} \phi_j^{k+1} + u_{j-1}^{k+1} \phi_{j-1}^{k+1}) + \lambda_j^{k+1} \phi_j^{k+1} \\
&\quad - \lambda_j^{k+1} \phi_j^{k+1} (u_j^{k+1} + u_{j-1}^{k+1}) - \lambda_j^{k+1} \phi_j^{k+1} (u_{j+1}^{k+1} + u_j^{k+1}) + 2\Delta t / (\Delta x^2) (\lambda_j^{k+2} - 2\lambda_j^{k+1}) + \lambda_j^{k+1} \phi_j^{k+1} (v_j^{k+1} + v_{j+1}^{k+1}) + \lambda_j^{k+1} \phi_j^{k+1} (v_j^{k+1} + v_{j-1}^{k+1}) - \lambda_j^{k+1} \phi_j^{k+1} (v_{j+1}^{k+1} + v_j^{k+1}) \\
&\quad + \alpha (u_j^{k+1} - \bar{u}_j^{k+1}) + \mu_j^{k+2} \phi_j^{k+2} - \Delta t / (2\Delta t) \left[ \mu_j^{k+1} (u_j^{k+1} \phi_j^{k+1} + u_{j+1}^{k+1} \phi_{j+1}^{k+1}) + \mu_j^{k+1} (u_j^{k+1} \phi_j^{k+1} + u_{j-1}^{k+1} \phi_{j-1}^{k+1}) - \mu_j^{k+1} (u_j^{k+1} \phi_j^{k+1}) + u_{j+1}^{k+1} (\phi_j^{k+1} + \phi_{j+1}^{k+1}) \right] + \Delta t / (2\Delta x^2) (\mu_j^{k+2} \phi_j^{k+2} - 2\mu_j^{k+1} \phi_j^{k+1} + \mu_j^{k+1} \phi_j^{k+1}) + \beta (v_j^{k+1} - \bar{v}_j^{k+1}) \\
\sigma_j^k &= \sigma_j^{k+2} / \Delta t / (\Delta x) \left[ (\sigma_j^{k+1} u_j^{k+1} - \sigma_j^{k+1} u_{j+1}^{k+1}) + 2\Delta t / (\Delta x^2) (\sigma_j^{k+2} - 2\sigma_j^{k+1} + \sigma_j^{k+1}) \right] - \lambda_j^{k+1} u_j^{k+1} \\
&\quad + \lambda_j^{k+1} u_j^{k+1} - \Delta t / (2\Delta t) \left[ \lambda_j^{k+1} (u_j^{k+1} + u_{j-1}^{k+1}) + \lambda_j^{k+1} u_j^{k+1} (u_{j+1}^{k+1} + u_j^{k+1}) - \lambda_j^{k+1} u_j^{k+1} (u_j^{k+1} + u_{j-1}^{k+1}) \right] - 2\Delta t / (2\Delta x) (\lambda_j^{k+1} (H_j - H_{j+1})) \\
&\quad + \lambda_j^{k+1} (H_{j+1} - H_j) + \lambda_j^{k+1} (H_{j-1} - H_{j+1}) + \lambda_j^{k+1} (H_{j+1} - H_j) + 2\Delta t v_j^{k+1} \lambda_j^{k+1} \\
&\quad + 2\Delta t / (\Delta x^2) (\lambda_j^{k+1} u_j^{k+1} - 2u_j^{k+1} \lambda_j^{k+1} + u_j^{k+1} \lambda_j^{k+1}) - \mu_j^{k+1} \phi_j^{k+1} + \mu_j^{k+1} \phi_j^{k+1} - \Delta t / (2\Delta x) (\lambda_j^{k+1} u_j^{k+1} + \lambda_j^{k+1} u_j^{k+1} (u_j^{k+1} + u_{j-1}^{k+1}) - \lambda_j^{k+1} u_j^{k+1} (u_j^{k+1} + u_{j-1}^{k+1})) \\
&\quad + 2\Delta t u_j^{k+1} \lambda_j^{k+1} + 2\Delta t / (\Delta x^2) (\lambda_j^{k+1} u_j^{k+1} - 2u_j^{k+1} \lambda_j^{k+1} + u_j^{k+1} \lambda_j^{k+1}) + \gamma (\phi_j^{k+1} - \bar{\phi}_j^{k+1}) \\
\end{align*}$$

for $k = N, N - 1, \ldots, 2$ and $j = 1, 2, \ldots, J$. To obtain $\lambda_j^{k+2}$, $\mu_j^{k+2}$, and $\sigma_j^{k+2}$ using the Eqs. (39a–c), we introduce two virtual time levels, $t_{k+1}$ and $t_{k-1}$, with initial values of the adjoint variables $\lambda_j^{N+1} = 0$, $\mu_j^{N+1} = 0$, and $\sigma_j^{N+1} = 0$. For $k = 1, 0$, the first time step equations are deduced for the computation of $\lambda_j^0$, $\mu_j^0$, and $\sigma_j^0$ from $\lambda_j^1$, $\mu_j^1$, and $\sigma_j^1$. Note that the introduction of the regularization term into the cost function (10) does not change the form of the adjoint model (39a–c) because the derivatives of the term $\xi / 2 \left[ (C_b H - H_b)^T (C_b H - H_b) \right]$ with respect to $u_j^k$, $v_j^k$, and $\phi_j^k$ equal zero.

In addition, we assume that the boundary conditions (2) are substituted directly into the model equations and that the periodic boundary conditions for the corresponding adjoint Eqs. (39a–c) are in the form

$$\begin{align*}
\lambda_j^k = \lambda_j^k, \quad \mu_j^k = \mu_j^k, \quad \sigma_j^k = \sigma_j^k \quad \text{and} \quad \lambda_j^{k+1} = \lambda_j^{k+1}, \quad \mu_j^{k+1} = \mu_j^{k+1}, \quad \sigma_j^{k+1} = \sigma_j^{k+1},
\end{align*}$$

for $k = N, N - 1, \ldots, 0$. The cyclic conditions for $H_j$, i.e., $H_0 = H_J$ and $H_{J+1} = H_1$, accounted for in this adjoint model facilitate the computation of the adjoint variables in the grid nodes $j = 1$ and $j = J$. 

12
References

[1] Y.-K. Sasaki, An objective analysis based on the variational method, J. Meteorol. Soc. JPN 2 (1958) 77-88.
[2] J.M. Lewis and J.C. Derber, The use of adjoint equations to solve a variational adjustment problem with advective constraints, Tellus A 37 (1985) 309-322.
[3] F.X.L. Dimet and O. Talagrand, Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects, Tellus A 38 (1986) 97-110.
[4] O. Talagrand and P. Courtier, Variational assimilation of meteorological observations with the adjoint vorticity equation. I: theory, Q. J. Roy. Meteor. Soc. 113 (1987) 1311-1328.
[5] P. Courtier and O. Talagrand, Variational assimilation of meteorological observations with the adjoint vorticity equation. II: Numerical results, Q. J. Roy. Meteor. Soc. 113 (1987) 1329-1347.
[6] D. Liu and J. Nocedal, On the limited memory BFGS method for large scale optimization, Math. Program. 45 (1989) 503-528.
[7] Y. Jarny, M.N. Ozisik, and J.P. Bardon, A general optimization method using adjoint equation for solving multidimensional inverse heat conduction, Int. J. Heat Mass Tran. 34 (1991) 2911-2919.
[8] O.M. Alifanov, Inverse Heat Transfer Problem, Berlin, Heidelberg: Springer, 1994.
[9] I.M. Navon and D.M. Legler, Conjugate-gradient methods for large-scale minimization in meteorology, Mon. Weather Rev. 115 (1987) 1479-1502.
[10] D.F. Shanno, Conjugate gradient methods with inexact searches, Math. Oper. Res. 3 (1978) 244-256.
[11] D.F. Shanno, On the convergence of a new conjugate gradient algorithm, SIAM J. Numer. Anal. 15 (1978) 1247-1257.
[12] A. Buckley and A. LeNir, Algorithm 630: BBVSCG-a variable-storage algorithm for function minimization, ACM T. Math. Software (TOMS) 11 (1985) 103-119.
[13] D.F. Shanno and K.H. Phua, Remark on Algorithm 500: minimization of unconstrained multivariate functions, ACM T. Math. Software (TOMS) 6 (1980) 618-622.
[14] W.W. Hager and H. Zhang, Algorithm 851: CG_DESCENT, a conjugate gradient method with guaranteed descent, ACM T. Math. Software (TOMS) 32 (2006) 113-137.
[15] W.W. Hager and H. Zhang, A new conjugate gradient method with guaranteed descent and an efficient line search, SIAM J. Optimiz. 16 (2005) 170-192.
[16] 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, Optim. Method Softw. 24 (2009) 63-87.
[17] I.E. Livieris and P. Pintelas, A new class of spectral conjugate gradient methods based on a modified secant equation for unconstrained optimization, J. Comput. Appl. Math. 239 (2013) 396-405.
[18] H. Liu, H. Wang, and Q. Ni, On Hager and Zhang's conjugate gradient method with guaranteed descent, Appl. Math. Comput. 236 (2014) 400-407.
[19] N. Andrei, Scaled conjugate gradient algorithms for unconstrained optimization, Comput. Optim. Appl. 38 (2007) 401-416.
[20] E.G. Birgin and J.M. Martinez, A spectral conjugate gradient method for unconstrained optimization, Appl. Math. Opt. 43 (2001) 117-128.
[21] J.C.F. Wong and J.L. Xie, Inverse determination of a heat source from natural convection in a porous cavity, Comput. Fluids 52 (2011) 1-14.
[22] 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.
[23] S.L. Tao, Y.H. Li, and I. Mugume, Model terrain correction using variational adjoint method with Tikhonov-total variation regularization, J. Phys. Conf. Ser. 1176 (2019) 022034.
[24] A.N. Tikhonov and V.Y. Arsenin, Solutions of ill-posed problems, New York: Halsted Press, 1977.
[25] M.S. Zhdanov, Geophysical inverse theory and regularization problems. Methods in geochemistry and geophysics, The Netherlands: Elsevier, 2002.
[26] J. Nocedal and S.J. Wright, Numerical Optimization. Springer Series in Operations Research, ed. P. Glynn and S.M. Robinson, New York: Springer-Verlag, 1999.