SHORT COMMUNICATIONS

Application of nonlinear optimization method to sensitivity analysis of numerical model*

XU Hui[1], MU Mu[1]** and LUO Dehai[2]

(1. LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China; 2. Department of Atmosphere Sciences, Ocean University of China, Qingdao 266003, China)

Received September 20, 2003; revised November 26, 2003

Abstract A nonlinear optimization method is applied to sensitivity analysis of a numerical model. Theoretical analysis and numerical experiments indicate that this method can give not only a quantitative assessment whether the numerical model is able to simulate the observations or not, but also the initial field that yields the optimal simulation. In particular, when the simulation results are apparently satisfactory, and sometimes both model error and initial error are considerably large, the nonlinear optimization method, under some conditions, can identify the error that plays a dominant role.

Keywords: **sensitivity analysis, nonlinear optimization, adjoint method, numerical model.**

The accuracy of numerical weather prediction decreases as the forecast time increases due to initial errors and model errors. To reduce these errors, meteorologists have performed numerous sensitivity analyses using the numerical simulation, the adjoint method and the linear singular vectors (LSVS)[1~6], etc. When applied in practice, these methods appear more or less unsatisfactory. For example, the numerical simulation method cannot give all possible significant initial fields or combinations of physical processes and parameters[7]. The adjoint method used by previous authors and the LSVS are based on the linear theory, and only describe the development of small perturbations during the validity period of the tangent linear model[5].

It is suggested by Mu et al.[8] that the nonlinear optimization method can be applied to sensitivity analysis of a numerical model. However, Mu et al. only briefly described the method as one of nonlinear optimization problems in atmospheric and oceanic science, and did not verify the results from the numerical experiments.

This paper, based on Mu et al.[8], introduces the nonlinear optimization method for sensitivity analysis of numerical model in detail. Besides, it further investigates how to identify the error type that plays a dominant role in the prediction results, especially when the simulation results are apparently satisfactory, and both the model error and initial error are notably large. As an example, with the two-dimensional quasi-geostrophic equation, we perform a series of numerical experiments to verify the theoretical results.

1 Method

Given the model $\Phi_t = M_t(\Phi_0)$, $t \in [0, T]$, the observation Φ_0^o at time 0 and Φ_T^o at time T, we want to find the initial field Φ_0 that yields the optimal simulation for Φ_T^o. The problem now becomes an optimization problem: find the optimal Φ_0 such that the cost function

$$J(\Phi_0) = \frac{1}{2}(M_T(\Phi_0) - \Phi_T^o)^{\mathrm{T}} \boldsymbol{W} (M_T(\Phi_0) - \Phi_T^o) \tag{1}$$

is minimum. $\boldsymbol{W}$ is the matrix of weighting coefficients matrix which is the inverse of the covariance matrix of the observational error. In order to minimize $J(\Phi_0)$, we need the information on the gradient of $J(\Phi_0)$ with respect to Φ_0. Because the number of control variables in a typical meteorological model is on the order of 10^6, the adjoint method is introduced

* Supported by the Key Innovation Direction Project of the Chinese Academy of Sciences (Grant No. KZCX2-208) and the National Natural Science Foundation of China (Grant Nos. 40233029 and 40075015)

** To whom correspondence should be addressed. E-mail: mumu@lasg.iap.ac.cn

to compute the gradient[9]. Using variational principle for (1), there is

$$\delta J(\Phi_0) = \langle \boldsymbol{W}(M_T(\Phi_0) - \Phi_T^o),\ \delta\Phi_T\rangle, \quad (2)$$

where $\delta\Phi_T$ is the development at time T of the initial perturbation $\delta\Phi_0$. Introducing the tangent linear operator M_T of M_T and substituting it into (2), we have

$$\delta J(\Phi_0) = \langle \boldsymbol{W}(M_T(\Phi_0) - \Phi_T^o),\ M_T(\Phi_0)\,\delta\Phi_0\rangle. \quad (3)$$

Moreover, introducing the adjoint propagator M_T^* of M_T we obtain

$$\delta J(\Phi_0) = \langle M_T^*(\Phi_0)(\boldsymbol{W}(M_T(\Phi_0) - \Phi_T^o)),\ \delta\Phi_0\rangle. \quad (4)$$

Thus, the gradient of $J(\Phi_0)$ with respect to Φ_0 is equal to

$$J(\Phi_0) = M_T^*(\Phi_0)(\boldsymbol{W}(M_T(\Phi_0) - \Phi_T^o)). \quad (5)$$

Then we can minimize J and find the optimal initial field Φ_0^* with the optimization algorithm.

2 Theoretical analysis

Let $E = \min J(\Phi_0)$. For a given error bound ε, there are two cases for E:

$$\begin{cases} E > \varepsilon, \\ 0 < E \leqslant \varepsilon. \end{cases}$$

When testing a model, $E > \varepsilon$ means that even if we get the optimal initial field Φ_0^*, the model is not able to simulate the observation Φ_T^o properly in the given error bound ε. Namely, no matter how we adjust Φ_0, a satisfactory simulation for Φ_T^o cannot be obtained. Then we can conclude that the model error is considerably large so that the model needs to be improved.

Then we consider the case $0 < E \leqslant \varepsilon$. Now the numerical solution $\Phi_T = M_T(\Phi_0)$ and the observation Φ_T^o have no apparent difference, which indicates that a satisfactory simulation can be obtained by adjusting the initial field Φ_0. It should be pointed out that, in this case, the model errors could be large too, which will be discussed later. With a given norm, defining a maximum allowable initial error ε_0, we have three cases now:

$$\begin{cases} \| \Phi_0^* - \Phi_0^o \| \ll \varepsilon_0, & \text{(a)} \\ \| \Phi_0^* - \Phi_0^o \| \sim \varepsilon_0, & \text{(b)} \\ \| \Phi_0^* - \Phi_0^o \| \gg \varepsilon_0. & \text{(c)} \end{cases}$$

When the model error is small, the model can simulate the movement of atmosphere very well. We can estimate the observation based on (a), (b) and (c). In (a), a satisfactory simulation for the observation Φ_T^o can be obtained from the existing observation Φ_0^o directly. We do not need to treat the initial field Φ_0 of the model particularly, and some ordinary interpolation is enough. In (b), a satisfactory simulation for Φ_T^o cannot be obtained from Φ_0^o directly. But if we improve the initial field Φ_0 of the model (for example, by assimilation method), a satisfactory simulation can be obtained too. In (c), the existing observation lacks enough information, and cannot represent the real weather and climate processes. If we want to obtain a satisfactory simulation for Φ_T^o, we should intensify the observational network to get more detailed observation than the existing one.

In case that the observational error is small, the observation is close to the real development of atmosphere. We can evaluate the model error based on (a), (b) and (c). In (a), the model error is small, and a satisfactory simulation for the observation Φ_T^o is easily obtained by adjusting the initial field Φ_0. In (b), there are certain model errors, but a satisfactory simulation can also be obtained by adjusting Φ_0 in the allowable error bound. It is in fact that an inaccurate model plus an inaccurate initial field produces a satisfactory simulation. In (c), the difference between Φ_0^* and Φ_0^o is too large, Φ_0^o is close to the real state, so Φ_0^* has no physical significance. We can conclude that the model error is large, a satisfactory simulation for Φ_T^o is illusive, and more work should be done to improve the numerical model.

In practice, it is common to evaluate the model error by comparing the numerical simulation with a relatively accurate observation, or assess the observation by comparing it with a relatively accurate numerical simulation. In these two cases, we can obtain some significant conclusions by the above analytical method. However, due to some objective reasons, when the model and observation are both considerably inaccurate, the applicability of the nonlinear optimization method to sensitivity analysis is limited. In this case, when $\min J > \varepsilon$, as mentioned above, the model error is considerably large, and the model needs to be improved. When $0 < \min J \leqslant \varepsilon$, a satisfactory simulation can be obtained by adjusting the initial field Φ_0, but both model error and observational error may be large. Let ε_0 be the observational precision that is usually known. If $\| \Phi_0^* - \Phi_0^o \| \gg \varepsilon_0$ appears, the optimal initial field Φ_0^* is far from the real state of atmosphere, and has no physical significance, which implies that the model error is large. In other cases, we cannot obtain some significant conclusions by the nonlinear optimization method.

In summary, we can get some instructive conclusions by the nonlinear optimization method for sensitivity analysis of a numerical model. In the next section, we will continue to discuss it with numerical experiments.

3 Numerical experiment

The nondimensional governing equation is

$$\begin{cases} \frac{\partial P}{\partial t} + \partial(\Phi, P) = 0, \\ P = \nabla^2\Phi - F\Phi + f + fh, \\ \Phi\mid_{t=0} = \Phi_0, \end{cases} \tag{6}$$

where P is the potential vorticity, Φ the stream function, F the Planetary Froude number, f the Coriolis parameter, H the characteristic vertical depth of the barotropic atmosphere and h the topography. The Jacobian operator is $\partial(\Phi, P) = \Phi_x P_y - \Phi_y P_x$.

Eq. (6) is solved under a double periodical boundary condition. Arakawa finite difference scheme is used to discretize the Jacobian operator. The temporal discretization is Adamas-Bashforth scheme. The familiar five-point difference scheme is employed to discretize the Laplacian operator. We take the space domain $[0, 6.4] \times [0, 3.2]$, the space step $d = 0.2$ and the time step $t = 0.018$, corresponding to the dimensional case $[0, 6400\text{ km}] \times [0, 3200\text{ km}]$, 200 km and 30 minutes, respectively. The parameters are $F = 0.102$, $f = 10.0$, and $H = 1.0$. The topography is $h(y) = h_0 \times 0.112 \times (\sin(4\pi y/3.2) + 1.0)$ where the model error is introduced by changing h_0, and $h_0 = 1.0$ denotes the model is accurate.

The "observation" is obtained by adding some normal random perturbations to the true value Φ_t^{tru} obtained by integrating the accurate model with $\Phi_0^{tru}(x, y) = \sin(2\pi y/3.2) + 0.1 \times \sin(2\pi x/6.4) + 0.25$. That is, $\Phi_t^{o(a)} = \Phi_t^{tru} + a \times \Phi_t'$ where Φ_t' is a normal random perturbation, a a coefficient denoting the error and $\Phi_t^{o(0.0)}$ equal to Φ_t^{tru}.

The cost function is as (1) where $\boldsymbol{W}$ is diagonal and normalized. Euclidian norm is employed. The optimization algorithm adopts the limited memory Broyden-Fletcher-Goldfarb-Shanna method (the limited memory BFGS method)[10].

Numerical experiments include two parts: to assess the observation when the model error is small and to evaluate the model error when the observational error is small.

3.1 When the model error is small

When $h_0 = 1.0$ and there is no model error, let $a = 0.05$, 1.0 or 2.0. Then we have obtained three different observations $\Phi_t^{o(0.05)}$, $\Phi_t^{o(1.0)}$ and $\Phi_t^{o(2.0)}$ for 3-day, 5-day and 7-day numerical experiments. The 7-day results are shown in Table 1, where Φ_7^f is the 7-day model forecast of the initial observation $\Phi_0^{o(a)}$.

Table 1. The 7-day results without model error ($\varepsilon = 3.1$, $\varepsilon_0 = 1.6$)

	$E = \min J$	$\| \Phi_0^* - \Phi_0^{o(a)} \|$	$\frac{\| \Phi_7^f - \Phi_7^{o(a)} \|}{\| \Phi_7^{o(a)} \|}$ (%)
$a = 0.05$	7.926×10^{-5}	0.1409	1.6
$a = 1.0$	4.362×10^{-3}	1.648	41
$a = 2.0$	4.706×10^{-2}	2.532	63

Here, as to $a = 0.05$, there are $\min J < \varepsilon$ and $\| \Phi_0^* - \Phi_0^{o(a)} \| \ll \varepsilon_0$ for 7-day run. The initial observation $\Phi_0^{o(a)}$ can be used as the initial value of the model directly to yield the satisfactory simulation for the observation $\Phi_7^{o(a)}$. As to $a = 1.0$, there are $\min J < \varepsilon$ and $\| \Phi_0^* - \Phi_0^{o(a)} \| \sim \varepsilon_0$. The satisfactory simulation for $\Phi_7^{o(a)}$ cannot be obtained directly from $\Phi_0^{o(a)}$. But if we improve the initial field Φ_0 of the model by an assimilation method, a satisfactory 7-day simulation can be obtained too. As to $a = 2.0$, there are $\min J < \varepsilon$ and $\| \Phi_0^* - \Phi_0^{o(a)} \| \gg \varepsilon_0$. Because the existing observation lacks enough information, we should improve the existing observational data to obtain a satisfactory simulation for $\Phi_7^{o(a)}$.

When $h_0 = 0.99$ and the model has some small errors, the numerical results are the same as when $h_0 = 1.0$ and the model has no model error. Tables are neglected.

3.2 When the observational error is small

In the case without the observational error, namely, when $a=0.0$ and the observation is $\Phi_t^{o(0.0)}$, we introduce three model errors $h_0=0.99$, $h_0=1.3$ and $h_0=1.8$ respectively for 3-day, 5-day and 7-day numerical experiments. The 7-day results are shown in Table 2.

Table 2 The 7-day results without observational error ($\varepsilon=3.1$, $\varepsilon_0=0.90$)

	$E=\min J$	$\|\Phi_0^*-\Phi_0^{o(0.0)}\|$	$\frac{\|\Phi_7^f-\Phi_7^{o(0.0)}\|}{\|\Phi_7^{o(0.0)}\|}(\%)$
$h_0=0.99$	1.398×10^{-7}	3.480×10^{-2}	1.5
$h_0=1.3$	5.998×10^{-4}	0.8891	61
$h_0=1.8$	4.508×10^{-1}	1.441	127

Here, when $h_0=0.99$, there are $\min J<\varepsilon$ and $\|\Phi_0^*-\Phi_0^{o(0.0)}\|\ll\varepsilon_0$. The model error is very small, and the satisfactory simulation for the observation $\Phi_7^{o(a)}$ can be easily obtained by adjusting the initial field Φ_0. When $h_0=1.3$, there are $\min J<\varepsilon$ and $\|\Phi_0^*-\Phi_0^{o(0.0)}\|\sim\varepsilon_0$. There exist certain model errors, but the satisfactory simulation can also be obtained by adjusting the initial field in the allowable error bound. When $h_0=1.8$, there are $\min J<\varepsilon$ and $\|\Phi_0^*-\Phi_0^{o(a)}\|\gg\varepsilon_0$. The optimal initial field Φ_0^* has no physical significance, the model error is large and the satisfactory simulation is illusive.

When the observation has small error (for the observation $\Phi_t^{o(0.05)}$), the numerical results are the same as when the observation has no error (for the observation $\Phi_t^{o(0.0)}$). Tables are neglected.

4 Conclusion and discussion

With the two-dimensional quasi-geostrophic model, we explore the application of the nonlinear optimization method to sensitivity analysis of a numerical model. The results suggest that by this method the ability of a model simulating the observation can be assessed quantitatively, and even the initial field that yields the optimal simulation can also be found out, which therefore become the main advantages of the nonlinear optimization method compared to other sensitivity methods. Besides, for the situation that the modeling is apparently satisfactory, we also explore the model error and initial error of the numerical model. It is shown that, in this case, the model error and initial error may be large. However, there is evidence that nonlinear optimization method, under some conditions, can identify the error that contributes dominantly to the uncertainties of the simulation results.

Although the two-dimensional quasi-geostrophic model is simple, the theoretical analysis of the numerical results is obtained logically. Therefore, it is reasonable to consider that for the complex models, nonlinear optimization method is of importance in the investigation of sensitivity analysis of a numerical model.

References

1 Hall, M. C. G. et al. Sensitivity analysis of a radiative-convective model by the adjoint method. J. Atmos. Sci., 1982, 39: 2038.

2 Hall, M. C. G. Application of adjoint sensitivity theory to an atmospheric general circula- tion model. J. Atmos. Sci., 1986, 43: 2644.

3 Errico, R. M. et al. Sensitivity analysis using an adjont of the PSU- NCAR mesoscale model. Mon. Wea. Rev., 1992, 120: 1644.

4 Rabier, F. et al. An application of adjoint models to sensitivity analysis. Beitr. Phys. Atmosph., 1992, 65: 177.

5 Rabier, F. et al. Sensitivity of forecast errors to initial conditions. Quart. J. Roy. Meteor. Soc., 1996, 122: 121.

6 Gelaro, R. et al. Sensitivity analysis of forecast errors and the construction of optimal perturbations using singular vectors. J. Atmos. Sci., 1998, 15: 1012.

7 Zou, X. et al. Incomplete observations and control of gravity waves in variational data assimilation. Tellus, 1992, 44A: 273.

8 Mu, M. et al. Nonlinear optimization problems in atmospheric and oceanic sciences, East-West Journal of Mathematics, 2002, Special Volume: 155.

9 Talagrand, O. et al. Variational assimilation of meteorological observation with the adjoint vorticity equation I: Theory. Quart. J. Roy. Meteor. Soc., 1987, 113: 1311.

10 Dong, C. L. et al. On the limited memory BFGS method for large scale optimization. Mathematical Programming, 1989, 45: 503.