NOTES AND CORRESPONDENCE

A Note on the Uniqueness Problem of Variational Adjustment Approach to Four-Dimensional Data Assimilation

By Yong Li¹

*School of Meteorology, University of Oklahoma, 200 Felgar St., Rm 219, Norman,
Oklahoma 73019, U.S.A.*

(Manuscript received 28 May 1991, in revised form 22 July 1991)

Abstract

A model for nonlinear resonant Rossby waves is utilized to obtain analytical solutions. "Identical twin" approach is adopted, in which one specific solution is chosen as the control which serves as pseudo-observation. Then initial condition is varied and thus specified various solutions are compared with the control. The cost function is computed and plot vs. the initial condition. Multiple minima are revealed in the cost function. In the context of variational adjustment approach to four-dimensional data assimilation, which is usually conducted with some minimization scheme based on the gradient information, this multiple-minimum phenomenon in cost function implies nonuniqueness, *i.e.*, the final convergent solution depends upon the first guess for the initial condition. Such nonuniqueness problem becomes more serious due to the increasing complexity of the cost function as the assimilation period gets longer. The relation to nonlinear energy exchange is discussed.

1. Introduction

Variational data assimilation has experienced dramatic development since its introduction into meteorology by Sasaki (1958, 1970). The adjustment approach, the so-called adjoint data assimilation, has been investigated intensively (Lewis and Derber, 1985; LeDimet and Talagrand, 1986; Courtier and Talagrand, 1990; Thacker and Long, 1988) and is approaching operational use (Navon, *et al.*, 1991). This variational adjustment approach is accomplished through minimizing a cost function, which is defined as the discrepancy between the model forecast and the observations, so that the optimal initial condition can be found. The minimization is usually performed iteratively based on the derivatives of the cost function with respect to the initial condition (Navon and Legler, 1987), which are obtained by integrating the adjoint model backward.

But the existence of multiple minima in the cost function and its impact on uniqueness of solution have not been studied yet, at least not to the ex-

tend desired by the popularity such data assimilation method has already gained, even though they have been speculated. In the following, a model for resonant nonlinear planetary waves was chosen for the investigation of multiple minima and some results are presented.

2. Model and Experimental Design

Considering a solution consisting of three Rossby wave modes, given by

$$\Psi = \sum_{n=1}^3 \Psi_n \quad (1)$$

where $\Psi_n = a_n(t) \cos(k_n x + l_n y + \sigma_n t)$, the equations for the resonant interaction can be written as

$$\frac{da_1(t)}{dt} = \frac{1}{2} \frac{(k_3 l_2 - k_2 l_3) [K_2^2 - K_3^2]}{K_1^2 + \alpha^2} a_2(t) a_3(t) \quad (2a)$$

$$\frac{da_2(t)}{dt} = \frac{1}{2} \frac{(k_1 l_3 - k_3 l_1) [K_3^2 - K_1^2]}{K_2^2 + \alpha^2} a_1(t) a_3(t) \quad (2b)$$

$$\frac{da_3(t)}{dt} = \frac{1}{2} \frac{(k_2 l_1 - k_1 l_2) [K_1^2 - K_2^2]}{K_3^2 + \alpha^2} a_1(t) a_2(t) \quad (2c)$$

where

¹Current affiliation: Supercomputer Computation Research Institute, B-186, Florida State University, Tallahassee, FL 32306, U.S.A.

$$K_1^2 = k_1^2 + l_1^2, \quad K_2^2 = k_2^2 + l_2^2, \quad K_3^2 = k_3^2 + l_3^2$$

A complete derivation of the model and discussion of the physical aspect can be found in Loguet-Higgins and Gill (1967) and also in Pedlosky (1987).

In this study, we set

$$\begin{aligned} k_1 &= k, & l_1 &= l, & \sigma_1 &= \sigma \\ k_2 &= -k, & l_2 &= l, & \sigma_2 &= -\sigma \\ k_3 &= 0, & l_3 &= -2l, & \sigma_3 &= 0 \end{aligned}$$

$$\sigma = \frac{\beta k}{k^2 + l^2 + \alpha^2}$$

which obviously satisfies the resonant condition: $k_1 + k_2 + k_3 = 0$, $l_1 + l_2 + l_3 = 0$ and $\sigma_1 + \sigma_2 + \sigma_3 = 0$. So the streamfunction can be written as

$$\begin{aligned} \Psi &= a_1(t) \cos(kx + ly + \sigma t) \\ &+ a_2(t) \cos(kx - ly + \sigma t) + a_3(t) \cos(2ly) \end{aligned}$$

The third mode in this system represents meridionally varying zonal currents and the first two modes are westward propagating waves. The equations are nondimensionalized before they are solved.

$$\begin{aligned} (x, y) &= L(x', y'), \quad t = Tt', \\ (a_1, a_2, a_3) &= A(a'_1, a'_2, a'_3), \quad (k, l) = K(k', l') \end{aligned}$$

Select the following scales

$$\begin{aligned} L &\sim 10^7 \text{ m}, \quad K = 2\pi/L \sim 10^{-6} \text{ m}^{-1}, \quad A \sim 10^7 \text{ m}^2 \cdot \text{s}^{-1}, \\ T &= L/U \sim 10^6 \text{ s} \quad \text{and} \quad \alpha^2 = f^2/(gD) \sim 10^{-12} \text{ m}^{-2} \end{aligned}$$

The characteristic scale A is so chosen that the velocity scale is in the order of $10 \text{ m} \cdot \text{s}^{-1}$. Substituting the scales into (2) and setting the nondimensional wave number $k=1=1.0$ yield the nondimensionalized equations with the primes all neglected for simplicity.

$$\frac{da_1(t)}{dt} = \frac{20}{3} a_2(t) a_3(t) \quad (3a)$$

$$\frac{da_2(t)}{dt} = -\frac{20}{3} a_1(t) a_3(t) \quad (3b)$$

$$\frac{da_3(t)}{dt} = 0 \quad (3c)$$

The third mode becomes a steady mode which neither gains nor loses energy in the resonant interaction. But its presence is essential. Without this mode, mode 1 and mode 2 are also of constant amplitude and then the whole system becomes two westward propagating Rossby waves without interacting each other and behaving just as linear waves. As it was pointed out, mode 3 represents meridionally varying zonal currents which does not change with time. The development of long waves is related with zonal currents such as described by instability

theories (Kuo, 1949; Charney, 1947). A full representation of these dynamical mechanisms of the instability theory in this simple system is out of hand and actually won't be necessary. Nevertheless, it is attempted to reflect the spirit of such a relation between zonal current and waves by crudely specifying the correlation of strengths:

$$a_3(0) = R_0 \sqrt{a_1^2(0) + a_2^2(0)} \quad (4)$$

where R_0 is an empirical parameter and set to be unity in this paper. Then the solutions of the model become

$$\begin{aligned} a_1(t) &= a_1(0) \cos\left(\frac{20}{3} R_0 \sqrt{a_1^2(0) + a_2^2(0)} t\right) \\ &+ a_2(0) \sin\left(\frac{20}{3} R_0 \sqrt{a_1^2(0) + a_2^2(0)} t\right) \end{aligned} \quad (5a)$$

$$\begin{aligned} a_2(t) &= a_2(0) \cos\left(\frac{20}{3} R_0 \sqrt{a_1^2(0) + a_2^2(0)} t\right) \\ &- a_1(0) \sin\left(\frac{20}{3} R_0 \sqrt{a_1^2(0) + a_2^2(0)} t\right) \end{aligned} \quad (5b)$$

$$a_3(t) = R_0 \sqrt{a_1^2(0) + a_2^2(0)} \quad (5c)$$

The “identical twin” approach is applied. Initial conditions $a_1(0)=2.0$ and $a_2(0)=1.0$ are chosen for the control. The solution from (5) with this set of initial conditions is taken to represent a hypothetical true state $a_{1ob}(t)$, $a_{2ob}(t)$, and $a_{3ob}(t)$, the “observation”.

The cost function to be examined is defined as

$$\begin{aligned} J &= \int_0^T \left\{ (a_1(t) - a_{1ob}(t))^2 + (a_2(t) - a_{2ob}(t))^2 \right. \\ &\quad \left. + (a_3(t) - a_{3ob}(t))^2 \right\} dt \end{aligned} \quad (6)$$

Initial conditions are varied to illustrate the shape of the cost function in order to investigate the uniqueness of minimization process in the variational data assimilation. Various values of T , are chosen to demonstrate the relation between the complexity of the cost function and the assimilation period.

3. Results and Analysis

Figure 1 shows the cost function with an assimilation period $T=0.2$. We can see that there is only one minimum. This minimum corresponds to the control. More detailed results (not shown) indicate that such single-minimum property persists when T is increased up to 0.7 (every additional 0.1), at least within the ranges for initial condition used in this study. But more than one minimum are revealed when T is equal to 0.8 (Fig. 2). The global minimum, of course, still corresponds to the control

Fig. 1. Cost function vs. initial values for a_1 and a_2 . Minima are marked by +. Assimilation period $T=0.2$.

which is $(a_1(0), a_2(0)) = (2.0, 1.0)$. Beside this, a local minimum exists, which is located approximately at $(a_1(0), a_2(0)) = (0.9, -0.2)$. The cost function becomes of more complexity as the period of assimilation increases. This can be seen in Fig. 3 with $T=1.2$.

The existence of multiple minima has tremendous impact on variational data assimilation based on the iterative minimization of a cost function with the gradient information. The final convergent solution depends on the first guess. For example, a first guess located within the basin of the local minimum will lead the minimization process to this local one at $(0.9, -0.2)$ instead of the perfect one at $(2.0, 1.0)$ as shown in Fig. 2. This, of course, will degrade the quality of assimilation or even fail it. When the first guess is lying within the basin of the global minimum, right convergence is guaranteed. But when the observation coverage is not so satisfactory such as in the situation of oceanic modeling or the atmosphere over vast ocean area, or when the model is a rough approximation to the realistic motion, the closeness of first guess to the optimal one is in question.

It is interesting to see how this multiple-minimum phenomenon is related with the nonlinear energy exchange. From the resonant conditions, we can easily obtain

$$k_3 l_2 - k_2 l_3 = k_1 l_3 - k_3 l_1 = k_2 l_1 - k_1 l_2 = b$$

and from the model Eq. (2), we have

$$(K_1^2 + \alpha^2) \frac{da_1^2}{dt} = b [K_2^2 - K_3^2] a_1 a_2 a_3$$

Fig. 2. Cost function vs. initial values for a_1 and a_2 . Minima are marked by +. Assimilation period $T=0.8$.

$$(K_2^2 + \alpha^2) \frac{da_2^2}{dt} = b [K_3^2 - K_1^2] a_1 a_2 a_3$$

$$(K_3^2 + \alpha^2) \frac{da_3^2}{dt} = b [K_1^2 - K_2^2] a_1 a_2 a_3$$

It is obvious that

$$\frac{d}{dt} [(K_1^2 + \alpha^2) a_1^2 + (K_2^2 + \alpha^2) a_2^2 + (K_3^2 + \alpha^2) a_3^2] = 0$$

This is equivalent to the conservation of total energy. $(K_i^2 + \alpha^2) a_i^2$ is the combination of kinetic energy and potential energy contributed by i -th mode. One mode transfers its energy to the other two modes and the cascade is reversible. The direction and efficiency of energy exchange depends on both the amplitudes of these three modes and the wave numbers. When any one of them is absent, the energy exchange can not be carried on. Even when three modes are all present but their total wave numbers are equal, energy exchange is prohibited. Then all three modes have constant amplitudes and they propagate linearly. In this case the cost function defined becomes a quadratic function of the initial conditions no matter how long the model is integrated and thus has only one minimum.

4. Concluding Remarks

A relatively simple model for nonlinear resonant Rossby waves has been used to study the uniqueness problem of the variational adjustment approach to data assimilation. It has been found that multiple minima appear in the cost function when the assimilation period is long enough. But this can not

Fig. 3. Cost function vs. initial values for a_1 and a_2 . Minima are marked by +. Assimilation period $T=1.2$.

lead us to a solution by simply shortening the assimilation time since the original purpose of data assimilation is to introduce the time-marching information from the observation into the model. However, to avoid using an unnecessarily long assimilation period helps keeping the simplicity of the cost function. How to find an optimal assimilation period is an important issue, which does not only influence the computational efficiency but also affects the uniqueness of convergence. When the nonlinear energy exchange is prohibited for some set of wave numbers, the Rossby waves behave linearly and the cost function has only one minimum.

It should be pointed out that the simple model used in this study can not fully represent the sophisticated model cases and the relation between zonal currents and the waves is physically too crude to be realistic. Therefore, we can only view the results conceptually instead of make direct quantitative interpretation to data assimilation in operational systems.

Acknowledgement

The author wants to express special thanks to Dr. Y. Sasaki for so many stimulating discussions, carefully reading the manuscript and generously providing advice and encouragement. Discussions with Drs. K. Droegemeier, J. Lewis, D. Lilly and L. White are very beneficial. This research was supported by the National Science Foundation in United States under Grant ATM86-57013, and by Grant ATM88-09862 to the Center for Analysis and Prediction of Storms, University of Oklahoma.

References

- Charney, J.G., 1947: The dynamics of long waves in a baroclinic westerly current, *J. Meteor.*, **4**, 135–162.
- Courtier, P. and O. Talagrand, 1990: Variational assimilation of meteorological observations with direct and adjoint shallow-water equations. *Tellus*, **42A**, 531–549.
- Kuo, H.L., 1949: Dynamic instability of two-dimensional nondivergent flow in a barotropic atmosphere. *J. Meteor.*, **6**, 105–122.
- LeDimet, F.X. and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus*, **38A**, 97–110.
- Lewis, J.M. and J.C. Derber, 1985: The use of adjoint equations to solve a variational adjustment problem with advective constraints. *Tellus*, **37A**, 309–322.
- Longuet-Higgins, M.S. and A.E. Gill, 1967: Resonant interaction between planetary waves. *Proc. Roy. Soc. London*, **A299**, 120–140.
- Navon, I.M. and D.M. Legler, 1987: Conjugate-gradient method for large-scale minimization in meteorology. *Mon. Wea. Rev.*, **115**, 1479–1502.
- Navon, I.M., X. Zou, K. Johnson, J. Derber and J. Sela, 1991: Variational Data Assimilation with an Adiabatic Version of the NMC Spectral Model. (Submitted to) *Mon. Wea. Rev.*
- Pedlosky, J., 1987: *Geophysical Fluid Dynamics (second edition)*, Springer-Verlag, New York.
- Sasaki, Y., 1958: An objective analysis based on variational method. *J. Meteor. Soc. Japan*, **36**, 1–12.
- Sasaki, Y., 1970: Some basic formalisms in numerical variational analysis. *Mon. Wea. Rev.*, **98**, 875–883.
- Thacker, W.C. and R.B. Long, 1988: Fitting dynamics to data. *J. Geophys. Res.*, **93**, 1227–1239.

4次元データ同化の変分法による解の一意性問題

Yong Li

(フロリダ州立大学スーパーコンピュータ研究所)

非線型共鳴ロスビー波の一つのモデルを解析解を得るために用いた。一つの解を観測値の代りのコントロールとして用い、次に初期値を変化させた解をコントロールと比較した。評価関数を初期条件について計算し図示したところ、多重極小値の存在が明らかにされた。4次元データ同化を、評価関数の傾度から極小化する変分法の手法では、この多重極小値は解の非一意性、すなわち極限の収束解が初期値の第1推定値に依存することを意味する。この非一意性の問題は、同化の期間が長くなるにつれて評価関数の凹凸がますます複雑になることから、より重大となる。非線型エネルギー交換との関係も議論される。