A Simple Method of Calculating the Optimal Step Size in 4DVAR Technique①

Wang Yunfeng (王云峰), Wu Rongsheng (伍荣生), Wang Yuan (王 元), and Pan Yinong (潘益农)

Laboratory of Meso-scale Severe Weather, Nanjing University, Nanjing, 210093

(Received July 6, 1999; revised April 14, 2000)

ABSTRACT

In four-dimensional variational data assimilation (4DVAR) technology, how to calculate the optimal step size is always a very important and indeed difficult task. It is directly related to the computational efficiency. In this research, a new method is proposed to calculate the optimal step size more effectively. Both nonlinear one-dimensional advection equation and two-dimensional inertial wave equation are used to test and compare the influence of different methods of the optimal step size calculations on the iteration steps, as well as the simulation results of 4DVAR processes. It is in evidence that the different methods have different influences. The calculating method is very important to determining whether the iteration is convergent or not and whether the convergence rate is large or small. If the calculating method of optimal step size is properly determined as demonstrated in this paper, then it can greatly enlarge the convergence rate and further greatly decrease the iteration steps. This research is meaningful since it not only makes an important improvement on 4DVAR theory, but also has useful practical application in improving the computational efficiency and saving the computational time.

Key words: 4DVAR, Optimal step size, Iterative convergence rate

1. Introduction

With highly developing of modern scientific technology and incessant updating of the observational system, many kinds of new sources of observational data are readily obtained, such as satellite data and radar data. It has been a critical problem to puzzle researchers to effectively employ these data in numerical calculation. 4DVAR technique has gradually been a very important tool in the main stream of numerical weather prediction, because it is of strong capability of combining and extracting the useful information from all kinds of observational data. The data are usually in a very complex form since it is observed with different approaches in different areas, and more importantly it is spaced in different time with different precision.

The application of four-dimensional variational data assimilation in numerical forecast and simulation was best demonstrated and developed by Lewis and Derber(1985), Le Dimet and Talagrand(1986), and Talagrand and Courtier(1987). Recently, it has become a popular interesting topic in atmospheric research and oceanic research (cf. Derber, 1987; Navon et al.,

①This work was supported by the National Natural Science Foundation under grants: 49735180 and 49675259, the "973 Project" CHERES(G1998040907), the Project of Natural Science Foundation of Jiangsu Province(BK99020), and the Project Sponsored by the Scientific Research Foundation for the Returned Overseas Chinese Scholars.

1987; Chao et al., 1992; Zou et al., 1995; Fillon and Errico, 1997; and the reference therein). In China, many interesting research works also have been implemented (cf. Gao et al., 1995; Zhu, 1995; Chen et al., 1998; Wang et al., 1999; and so on).

It is well known that an iterative process is needed in 4DVAR in order to obtain the initial conditions in the optimized fashion. No matter the iterations are convergent or not, and whether the convergence rate is large or not, they all depend on the proper determination of the optimal step size. Furthermore, how to calculate the optimal step size, in one sense, has been the crucial issue in the minimizing processes of 4DVAR algorithm. Though many papers had studied the 4DVAR technique, few of them had even considered to determine the optimal step size properly. And in fact, few practical methods had been given because of the great difficulty in such a research field. For instance, Gao et al. (1994) used the one-dimensional searching method, but sometimes the method is inefficient. Derber (1987) provided in other way a guess expression to roughly calculate the optimal step size in a quasi-geostrophic model, and the method had been used in some succeeding researches (Navon et al., 1992). Note that Derber's method was only a guess expression and there exist limits in applying. Wang et al. (1999) suggested a simple formula to design the optimal step size, which has relatively low calculation efficiency in some iterative processes. In this paper, we present a new method based on the character of 4DVAR, to obtain the proper optimal step size, which greatly releases the disadvantages mentioned above to a considerable extent. And numerical sensitivity tests are implemented to check the efficiency and accuracy of this new method.

2. Derivation of the optimal step size

Supposing that $Y(t)$ is the control variable of the model, the process to obtain the optimal initial field can be realized by the steepest descent algorithms or the conjugate gradient algorithms, such as

$$Y^{\nu+1}(t_0) = Y^{\nu}(t_0) - \rho^{\nu}[\nabla J]^{\nu}, \tag{1}$$

where ν represents the iteration step, and J is the cost function in general. ∇J is the gradient vector of the cost function J with respect to the initial field $Y(t_0)$. ρ is the optimal step size. From Equation (1), the problem of calculating the initial field $Y(t_0)$ has been shifted into the problem of calculating ∇J and ρ. Since ∇J can be obtained by integrating inversely the adjoint model from the final time to the initial time, the remnant is to calculate the optimal step size. The iteration convergence rate will greatly be determined by the design of optimal step size. In this section, we give a detail discussion of the optimal step size from the point of mathematical view, and two calculating methods are introduced.

(I) The simplest calculation of the optimal step size is based on the Newton–Raphson iteration through the expression $\rho = \frac{1}{J}$. This method avoids the complex calculating process and reduces much of computer resource so that the optimizing process can be performed effectively. The big problem in the Newton–Raphson iteration is that though the iteration is convergent, the convergence processing is quite slow. Especially at the beginning when J is very large and thus ρ is very small, the convergence rate is heavily decreased. In order to accelerate the convergence, we might take the first modification in the method of Newton–Raphson iteration through multiplying a proper constant coefficient with the guessing value of the optimal step size from the Newton–Raphson method. However, the constant coefficient is high artificially determined such that somewhat uncertain is inevitable. If we

want to reach the prescribed threshold of convergence in just fewer iteration steps as possible as we can, we should find another effective and less-subjective method to obtain the proper optimal step size.

(II) Supposing that $\nabla J[Y(t_0)]$ at ν iteration step has been obtained by integrating inversely the adjoint model, then the initial field at $\nu+1$ iteration step could be obtained from Equation (1) if the optimal step size ρ^{ν} at ν iteration step is calculable.

Introducing a guess value of optimal step size, denoting ρ^{ν}_{*} at ν iteration step, the guess initial field could be calculated as follows:

$$Y^{\nu+1}_{*}(t_0)= Y^{\nu}(t_0)- \rho^{\nu}_{*}[\nabla J]^{\nu}. \tag{2}$$

Note that for the first iteration step, we take $\rho^{1}_{*}=\frac{1}{J}$; for the succeeding iteration steps, $\rho^{1}_{*}=\rho^{\nu-1}$.

In time-forward integrating processes, forecasting model equations in the continuous form should be written in the discrete form, and the discrete model could be regarded as a linear or nonlinear algorithm of K transfer from one space to another space. When K acts on the initial field $Y(t_0)$, the succeeding model solutions $Y(t_i)$ $(i=1,\cdots,n)$ at later times can be obtained. It leads to

$$Y^{i+1}(t_i)= K[Y^{i+1}(t_0)], \tag{3}$$

and

$$Y^{i+1}_{*}(t_i)= K[Y^{i+1}_{*}(t_0)]. \tag{4}$$

Equations (3) and (4) are also applicable at the other iteration steps for $i=1,\cdots,\nu$.

If the algorithm of K acts on all terms of Equation (1) and Equation (2) respectively, we have

$$K[Y^{i+1}(t_0)]= K\{Y^{\nu}(t_0)- \rho^{\nu}[\nabla Y(t_0)]^{\nu}\}, \tag{5}$$

and

$$K[Y^{i+1}_{*}(t_0)]= K\{Y^{\nu}_{*}(t_0)- \rho^{\nu}_{*}[\nabla Y(t_0)]^{\nu}\}. \tag{6}$$

In order to estimate the optimal step size, it is assumed that for a very small perturbation, the corresponding model solutions will vary linearly in the search direction (cf. Derber, 1987). While the calculating values of ρ^{ν} and ρ^{ν}_{*} at each iteration step are used to derive the initial fields, they can be taken account of constant values. Substituting Equations (3) and (4) into Equations (5) and (6), we obtain

$$Y^{\nu}(t_i)- Y^{\nu+1}(t_i)= \rho^{\nu} K\{[\nabla Y(t_0)]^{\nu}\}, \tag{7}$$

and

$$Y^{\nu}(t_i)- Y^{\nu+1}_{*}(t_i)= \rho^{\nu}_{*} K\{[\nabla Y(t_0)]^{\nu}\}. \tag{8}$$

Multiplying Equations (7) and (8) by the factor of $[Y^{\nu}(t_i)- Y^{\nu+1}_{*}(t_i)]^{T}$, respectively, results in

$$[Y^{\nu}(t_i)- Y^{\nu+1}(t_i)][Y^{\nu}(t_i)- Y^{\nu+1}_{*}(t_i)]^{T}= \rho^{\nu} K\{[\nabla Y(t_0)]^{\nu}\}[Y^{\nu}(t_i)- Y^{\nu+1}_{*}(t_i)]^{T}, \tag{9}$$

and

$$[Y^{\nu}(t_i)-Y_*^{\nu+1}(t_i)][Y^{\nu}(t_i)-Y_*^{\nu+1}(t_i)]^T=\rho_*^{\nu}K\{[\nabla Y(t_0)]^{\nu}\}[Y^{\nu}(t_i)-Y_*^{\nu+1}(t_i)]^T, \quad (10)$$

where superscript "T" denotes transposition. To be sure that at ν+1 iterative times, the relationship holds

$$Y^{\nu+1}(t_i)=Y_{\rm obs}(t_i). \quad (11)$$

We can get the following equation by substituting Eq. (11) into Eq. (9) and then comparing with Equation (10):

$$\frac{\rho^{\nu}}{\rho_*^{\nu}}=\frac{[Y^{\nu}(t_i)-Y_{\rm obs}(t_i)][Y^{\nu}(t_i)-Y_*^{\nu+1}(t_i)]^T}{[Y^{\nu}(t_i)-Y_*^{\nu+1}(t_i)][Y^{\nu}(t_i)-Y_*^{\nu+1}(t_i)]^T}. \quad (12)$$

Equation (12) also can be rewritten as

$$\rho^{\nu}=\rho_*^{\nu}\frac{<\hat{\Phi}(t_i),\Phi'(t_i)>}{<\Phi'(t_i),\Phi'(t_i)>}, \quad (13)$$

where $\hat{\Phi}(t_i)=Y^{\nu}(t_i)-Y_{\rm obs}(t_i)$ is the difference between the model solution $Y^{\nu}(t_i)$ and the observation $Y_{\rm obs}(t_i)$; whereas $\Phi'(t_i)=Y^{\nu}(t_i)-Y_*^{\nu+1}(t_i)$ is the difference between the model solution $Y^{\nu}(t_i)$ and the guess model solution $Y_*^{\nu+1}(t_i)$. $Y^{\nu}(t_i)$ can be obtained by integrating the forecasting model with the initial field $Y^{\nu}(t_0)$; on the other hand, $Y_*^{\nu+1}(t_i)$ can be obtained by integrating the forecast model with the guess initial field $Y_*^{\nu+1}(t_0)$. Notation "$<.>$" denotes an inner product.

From Equation (13), the generalized expression of the proper optimal step size should be a sum of ρ^{ν} at all times with different weighting coefficient, namely

$$\rho^{\nu}=\rho_*^{\nu}\sum_{i=1}^{N}W_i\frac{<\hat{\Phi}(t_i),\Phi'(t_i)>}{<\Phi'(t_i),\Phi'(t_i)>}, \quad (14)$$

where W_i is the weighting coefficient at each integrating step, which can be determined by some optimal control methods (Hamming, 1989). Nevertheless, the simplest form can be given by $\frac{1}{N}$. That is to say, the weighting coefficients at all times are the same. Our research has shown that even if the simplest form $W_i=\frac{1}{N}$ is taken, the satisfying results can be obtained (not shown). Two numerical tests will be used to verify the accuracy of the derivations above.

3. Numerical results

Both nonlinear one–dimensional advection equation and two–dimensional inertial wave equation are used to test and compare the influence of different methods of the optimal step size calculations on the iteration steps, as well as the simulation results of 4DVAR processes.

3.1 *Nonlinear one–dimensional advection–diffusing equations*

The nonlinear one–dimensional advection–diffusion equation is given as follows:

$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}-\frac{\partial}{\partial x}(k(x)\frac{\partial u}{\partial x})=S(x),\quad 0\leqslant x\leqslant 2, \tag{15}$$

where the coefficient and the external forcing are given as $k(x)=2\times 10^{-2}$ and $S=0.025\cos(\pi x)$, respectively. The initial and boundary conditions of Equation (15) are expressed as

$$u(x,0)=\sin(2\pi x), \tag{16}$$

$$\frac{\partial}{\partial x}u(0,t)=\frac{\partial}{\partial x}u(2,t)=0. \tag{17}$$

And its linear adjoint equation can be written as

$$-\frac{\partial u^*}{\partial t}-\frac{\partial uu^*}{\partial x}+u^*\frac{\partial u}{\partial x}-\frac{\partial}{\partial x}(k(x)\frac{\partial u^*}{\partial x})=0. \tag{18}$$

To obtain the adjoint program code, there are usually two methods: direct method and indirect method (Wang et al., 1999). In our test, the direct method is used. That is to say, the adjoint program code is obtained by transposing the numerical model program code directly. The solutions of Equation (15) under the conditions (16) and (17) are expressed as $\tilde{u}_i^p=u(i\Delta x, p\Delta t)$, which are expressed in discrete space, where $i=0,1,\cdots,40$; $p=1,2,\cdots,60$; $\Delta x=0.05$; and $\Delta t=0.01$. It is worth noting that in such an ideal test, the solutions $\tilde{u}_i^p=u(i\Delta x, p\Delta t)$ are used as the prescribed observational data in corresponding integrating steps. The initial field at the first iteration is taken as zero value, which is of course far away from the true state of Equation (16). Then we want to see whether such misfit could be "repaired" by the assimilation process. In this paper, several selected methods of calculating the optimal step size are used in the assimilation process to obtain the optimal initial fields. By comparing the corresponding results with the different methods in determination of the optimal step size, the superiority of Equation (14) can be identified.

As shown in Fig. 1, the standardizing gradient norm $\|\nabla J\|$ (solid points) and the standardizing cost function J (hollow points) vary with the iteration steps: In Fig. 1a Equation (14) is used to calculate the optimal step size; in Fig. 1b the optimal step size is calculated by the method of Derber (1987); in Fig. 1c the one-dimensional searching method is used (after Gao and Chou, 1994).

It has been found that the descending of the standardizing gradient norm and the standardizing cost function in Fig. 1a are the fastest and that the iteration steps reaching the convergent threshold are also the fewest, compared with the profiles in Figs. 1b and 1c. Such comparison indicates that the proposed method (i.e., Equation (14)) is more effective than the method proposed by Derber (1987) and the one-dimensional searching method (Gao et al., 1994). Compared with Derber's method, though the computing time in every iteration step is the same, the iteration steps in our method are much fewer, and the total CPU cost in our method is less expensive, too. On the other hand, it may be hard to say that our method takes less computational expense in every iteration step. But one thing is obviously in certain that the total computing time in our method is much less than that in the one-dimensional searching method, simply because the less iteration steps are required in our method.

3.2 *Two-dimensional inertial wave equations*

Taking the two-dimensional inertial wave equations as the test platform, the effects of the optimal step size in two selected methods are tested by numerical experiments. The inertial

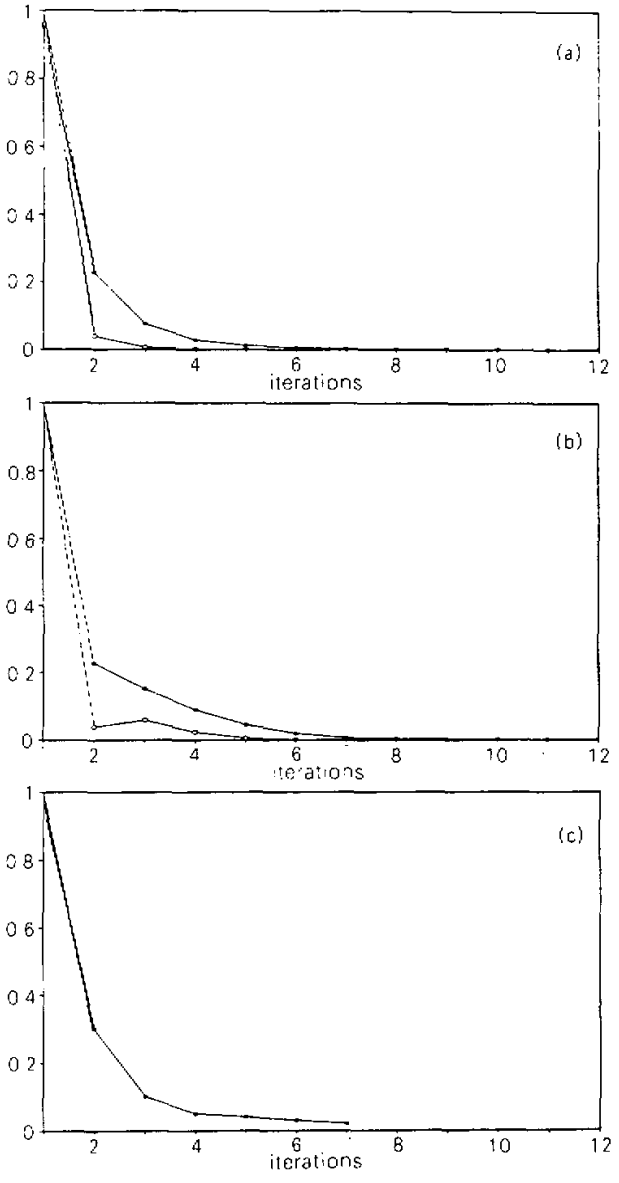

Fig. 1. The assimilation results with different methods in determination of optimal step size (a) When Equation (14) is used, the variations of the standardizing gradient norm $\| \nabla J \|$ (solid points) and the standardizing cost function J (hollow points) with the iteration steps; (b) When Derber's method is used, the variations of the standardizing gradient norm $\| \nabla J \|$ (solid points) and the standardizing cost function J (hollow points) with the iteration steps; (c) When one-dimensional searching method is used, the variations of standardizing gradient norm $\| \nabla J \|$ (solid points) with the iteration steps (after Gao et al., 1994).

wave equations are expressed as

$$\frac{\partial u}{\partial t} - fv = 0, \tag{19a}$$

$$\frac{\partial v}{\partial t} + fu = 0. \tag{19b}$$

If the initial wind fields are given by $u|_{t=t_0} = 0$ and $v|_{t=t_0} = v_0$, meanwhile the initial phase φ_0 is taken to be zero value, then the analytic solutions are

$$u^{\text{ana}} = v_0 \sin(ft + \varphi_0), \tag{20a}$$

$$v^{\text{ana}} = v_0 \cos(ft + \varphi_0). \tag{20b}$$

In the time-forward integrating processes, the discrete method of Equation (20) used here is followed from Pielke (1984),

$$(u_i^{\tau+1} - u_i^{\tau})/\Delta t = fv_i^{\tau}; (v_i^{\tau+1} - v_i^{\tau})/\Delta t = -fu_i^{\tau+1}. \tag{21}$$

Introducing the adjoint variables u^* and v^* for the corresponding variables u and v respectively, the adjoint equations of (21a) and (21b) can thus be written as

$$\frac{\partial u^*}{\partial t} + fv^* = 0, \tag{22a}$$

$$\frac{\partial v^*}{\partial t} - fu^* = 0. \tag{22b}$$

The adjoint program code is obtained through the same way as described in the advection case. In our numerical experiments, there are respectively 81 and 76 grid points along x and y directions with a grid size of 20 km in the horizontal and 9 grid points in the vertical, with a grid size of 1 km. The initial fields of u and v are uniformly distributed in the spatial space. The time step size is taken to be 187.0 min with the integrating number of 336. Then the total integrating time is 62832 min. The assimilation time interval is chosen as 10472 min, and there are total 7 assimilation time steps. For given initial fields, the forecasting solutions of u_i and v_i $(i=1,2,\cdots,7)$ at each assimilation time step are obtained by calculating Equation (21). At the same time, by calculating Equations (20a) and (20b), we can obtain the analytical solutions u_i^{ana} and v_i^{ana} $(i=1,2,\cdots,7)$ at the corresponding time step. Note that $v_0 = 5$ m/s and $\varphi_0 = 0$ are taken in Equation (20). Regarding the valid analytical solutions as the prescribed observational data, the cost function is thus defined as follows:

$$J = \frac{1}{2}\sum_{i=1}^{7}[(u_i - u_i^{\text{ana}})^2 + (v_i - v_i^{\text{ana}})^2], \tag{23}$$

where $i=1,2,\cdots,7$.

The numerical experiments with different methods of determining the optimal step size under three initial conditions are illustrated in Table 1. Firstly, the so-called standard initial fields in Table 1 are described by $u|_{t=t_0} = 0$ and $v|_{t=t_0} = v_0$, which indeed means that there exists no perturbation with the amplitude of v_0 and the phase of φ_0. Secondly, supposing that there exists an amplitude perturbation (v_{per}) in the standard initial fields, the original amplitude v_0 is added by the perturbation value of $v_{\text{per}} = 0.1v_0$, meanwhile the others remain unchanged. Thirdly, the phase perturbation (φ_{per}) is supposed in the initial fields by adding a

perturbation value of $\varphi_{per} = \frac{\pi}{6}$ to the original phase of zero. Similarly, the others remain unchanged.

Table 1. The numerical experiments with different methods of determining the optimal step size under three initial conditions

	Standard initial fields	A perturbation of amplitude existing in the initial fields	A perturbation of phase existing in the initial fields
$\rho = \frac{1}{J}$	Case1	Case3	Case5
$\rho^r = \rho^r \cdot \sum_{i=1}^{N} W_i \frac{<\hat{\Phi}(t_i), \Phi^r(t_i)>}{<\Phi^r(t_i), \Phi^r(t_i)>}$	Case2	Case4	Case6

In Table 1, total six cases are designed to test two methods of determining of optimal time step under three initial conditions. The variational assimilation results for Case 1 are shown in Fig. 2.

Since in Case 1 there are two control variables u and v, the optimizing processes of $u(t_0)$ and $v(t_0)$ should correspondingly require two components $\nabla_u J$ and $\nabla_v J$ of the gradient ∇J, where $\nabla_u J$ is the gradient of J according to the initial field $u(t_0)$, while $\nabla_v J$ is the gradient of J according to the initial field $v(t_0)$. A new variable, A is defined, which equals $\sqrt{u^2 + v^2}$. And its root-mean-square error (RMSE) is denoted by A_{rmse}. From Fig. 2 we can find that the optimal step size increases gradually with the varying of iteration steps in the range 0.0055 at the initial state to 0.011 at the final state. As a whole, the cost function J, the component gradient norms $\| \nabla_u J \|$ and $\| \nabla_v J \|$ and the root-mean-square error A_{rmse} all decrease gradually with the iteration steps. And their values at the final state are much less than those at the initial state. All this shows clearly that the assimilation process is convergent and it is feasible to calculate the optimal step size by using $\rho = \frac{1}{J}$. However, the optimal step size is very small at the beginning and even at the end. Moreover, the decreasing processes of J, $\| \nabla_u J \|$, $\| \nabla_v J \|$ and A_{rmse} are all slow with the iteration steps, unfortunately. Judging from this point, the method by $\rho = \frac{1}{J}$ is not very satisfactory.

The assimilation results in Case 2 are shown in Fig. 3, in which Equation (14) is used to determine the optimal time step.

As shown in Fig. 3, the optimal step size at first increases with the iteration steps, and ranges from 0.14 to 0.18. After the optimal time step turns to be of negative value at the 5th iteration step, it varies near around zero value for the succeeding iterations. It seems that somewhat self-adaptive capability exists in Equation (14) for calculation of the optimal step size. Compared with Fig. 2, the method in Case 2 is much better than the one in Case 1 and its iteration convergence is much more fast. The cost function J, the component gradient norm $\| \nabla_u J \|$ and $\| \nabla_v J \|$, as well as A_{rmse} in Case 2 are also decreasing rapidly with the adding of the iteration steps. The cost function of J and A_{rmse} quickly reaches the minimum just when the iteration steps approach to 3 about. Also for the component gradient norm $\| \nabla_u J \|$ and $\| \nabla_v J \|$, they almost reach their minima when the iteration steps approach to 5. All of them

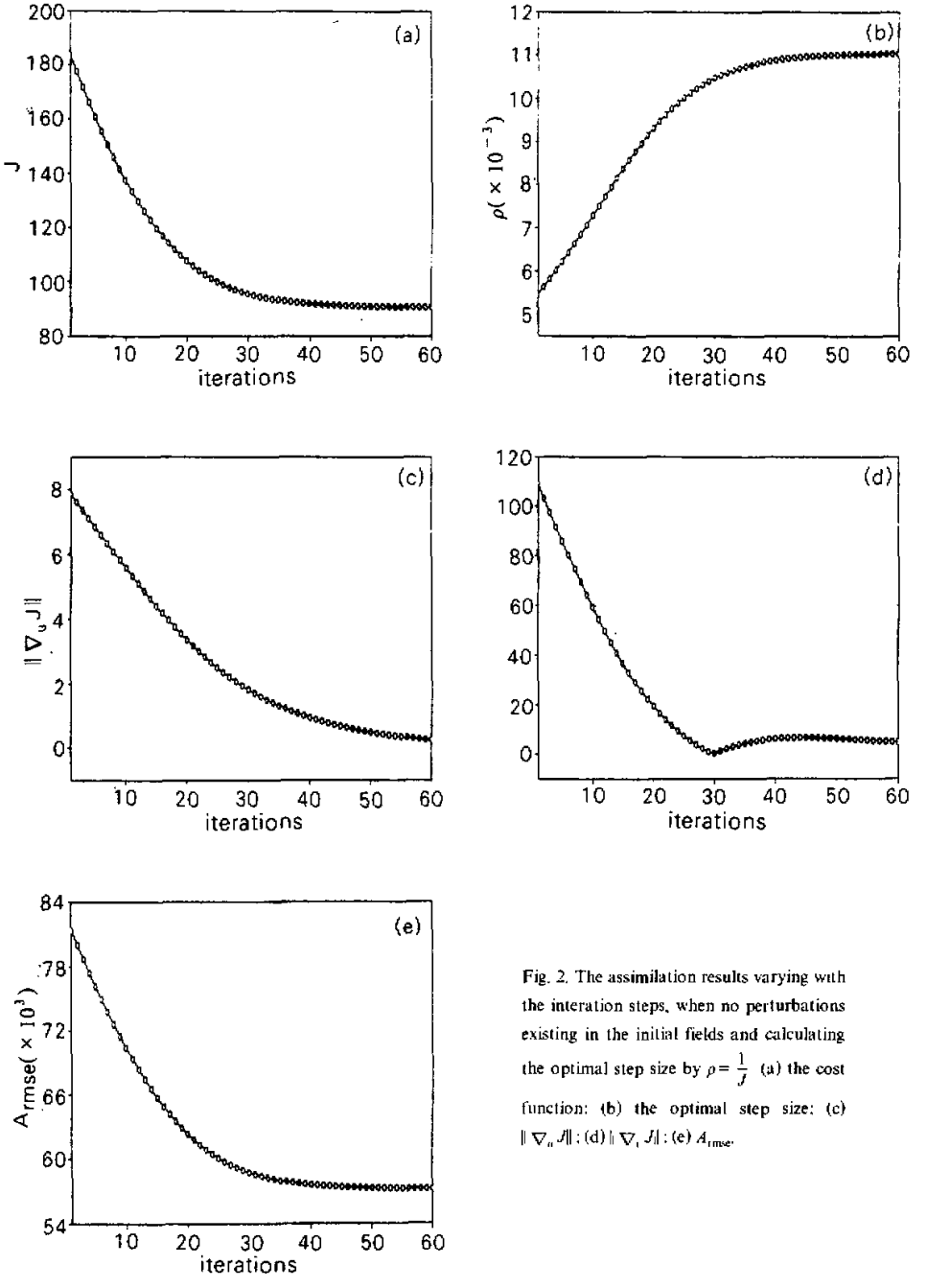

Fig. 2. The assimilation results varying with the interation steps, when no perturbations existing in the initial fields and calculating the optimal step size by $\rho = \frac{1}{J}$ (a) the cost function; (b) the optimal step size; (c) $\|\nabla_u J\|$; (d) $\|\nabla_v J\|$; (e) A_{rmse}.

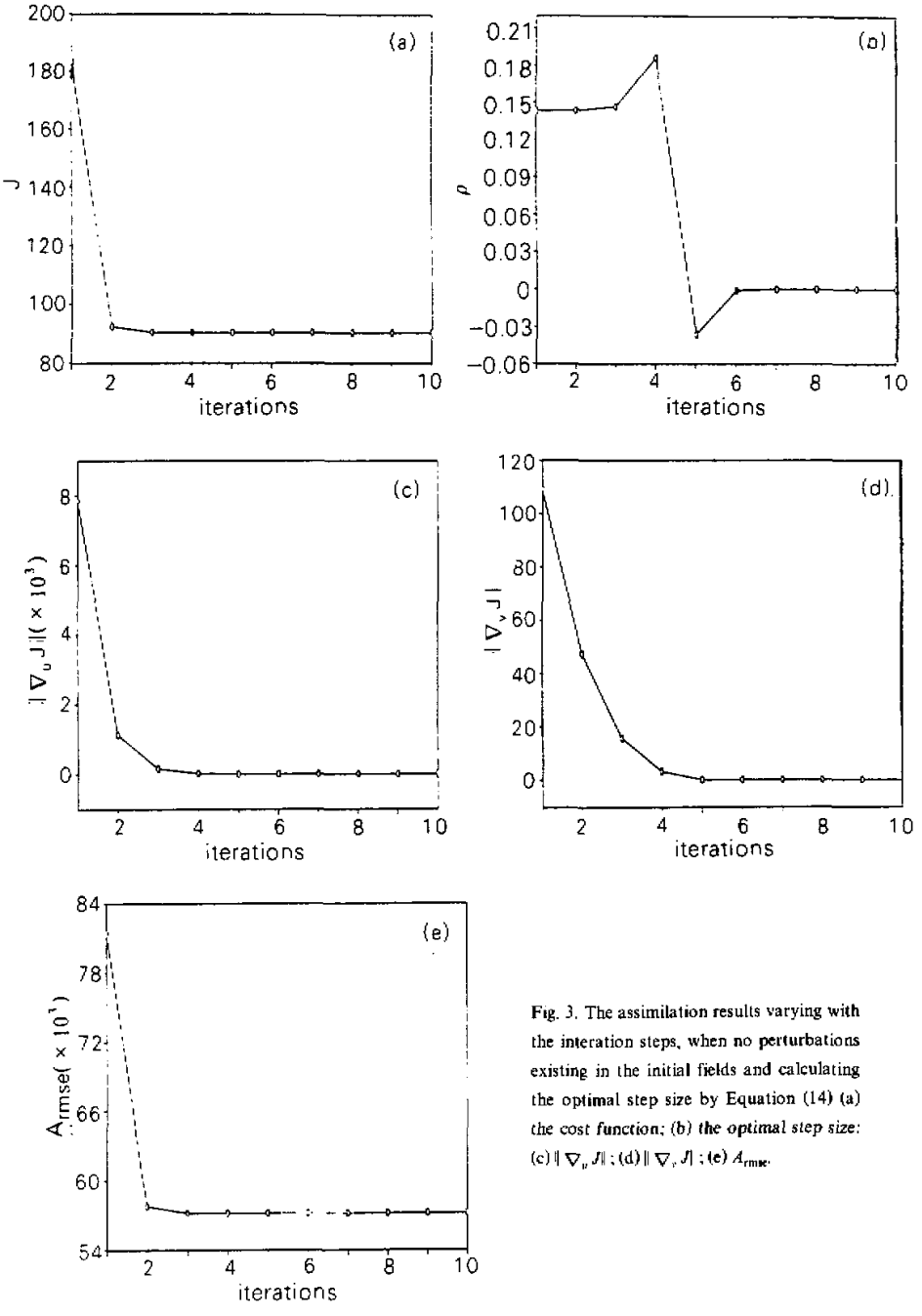

Fig. 3. The assimilation results varying with the interation steps, when no perturbations existing in the initial fields and calculating the optimal step size by Equation (14) (a) the cost function; (b) the optimal step size; (c) $\|\nabla_u J\|$; (d) $\|\nabla_v J\|$; (e) A_{rmse}.

have clearly indicated that Equation (14) is one of the very effective methods to the calculation of the optimal step size since it enlarges the assimilation iteration rate.

In fact, all of experiments listed in Table 1 have been done such as case3, case4, case5, and case6. More interesting, these cases have demonstrated when the perturbations exist in the initial fields, Equation (14) is a very effective way to properly calculate the optimal step size than other methods tested in the paper is.

Conclusively, the proper method of calculating the optimal step size is of great importance since it is deeply associated with the iteration convergence and its rate. And it is also directly relative to the feasibility of the algorithm.

5. Conclusion

Since the rapid development in weather numerical prediction, 4–dimensional variational data assimilation technique becomes very important, and while how to properly calculate the optimal step size is always a crucial issue. In this paper, a new method of calculating the optimal step size is proposed based on the property of 4DVAR. Taking advantage of the nonlinear one–dimensional advection–diffusing and the 2–dimensional inertial wave equations, the different methods of calculating the optimal step size are used to study their influence on the iteration steps and the simulation results of 4DVAR.

In the case when one–dimensional advection–diffusing equation is used, we find that the method adopted in this paper is more effective than that proposed by Derber (1987) or than that of the one–dimensional searching method (Gao, et al., 1994). Since both the standardizing gradient norm and the standardizing cost function are decreasing much more faster than those of other methods.

On the other hand, in the situation of the two–dimensional inertial wave equation case, much more complicated sensitivities are conducted in order to illustrate the advantage of the proposed method. Firstly, it is considered that the optimal step size is calculated by the simplest way through $\rho=\frac{1}{J}$, i.e., Newtonian method. Moreover, supposing that no perturbations existing in the initial fields, we find that J, $\|\nabla_u J\|$ and $\|\nabla_v J\|$, as well as A_{rmse}(root–mean–square error of the numerical solutions u and v) are all gradually decreasing with the adding of the assimilation iteration steps. That is to say, the assimilation iteration process actually is in convergence. However, the iteration rate is very small, unfortunately. Furthermore, supposed that there were some perturbations of the amplitude or the phase existing in the initial fields, the iterative convergence of J, $\|\nabla_u J\|$, and $\|\nabla_v J\|$ became to be too slow. Clearly, the optimal step size determined by $\rho=\frac{1}{J}$ is not a satisfying method although it is of simple calculation. Nevertheless, if the optimal step size is calculated through multiplying $\frac{1}{J}$ with a proper coefficient, it can enlarge the iteration rate. The problem however is that to obtain the proper coefficient is a highly artificial and delicate job, there exists a great variation and uncertain in the multiplying factor in different cases and at different iteration steps. Such weakness is largely eliminated in the proposed method. Indeed, no matter perturbation exists in the initial fields or not, the cost function J, the gradient norm $\|\nabla_u J\|$, $\|\nabla_v J\|$, and A_{rmse}, all of them are decreased very fast with the adding of the assimilation iteration steps. For instance, when the iteration steps take about 4 or 5, they are almost reaching their minimum.

Conclusively, it is very important to calculate the optimal step size properly taken into account of convergence of the assimilation iteration as well as the convergent efficiency. Using the proposed method to calculate the optimal step size, it can greatly improve the algorithm efficiency and saving the calculating time, and thus has an especially practical value. In the view of theoretical point, our research provides a less-subjective and economical way to implement 4DVAR processes.

REFERENCES

Chao, W. C., and L. P. Chang, 1992: Development of a four-dimensional varitional analysis system using the adjoint method at GLA. Part I: Dynamics. *Mon. Wea. Rev.*, **120**, 1661–1673.

Chen Zitong, Shen Tongli, Ding Yihui, and Sun Linping, 1998: *Data assimilation system for mesoscale models part I: Layout of an adjoint model.* J. Nanjing Institute of Meteorology, **21(2)**, 165–172.

Chou Jifan, 1995: *The theory and new method of four-dimensional assimilation, some new techniques in numerical weather forecasting.* Meteorology Publishing Company, Beijing. 262–294.

Derber, J. C. 1987: Variational four-dimensional analysis using quasi-geostrophic constraints. *Mon. Wea. Rev.*, **115**, 998–1008.

Fillon, L., and R. Errico, 1997: Variational assimilation of precipitation data using moist convective parameterization schemes: A 1D-Var Study. *Mon. Wea. Rev.*, **125**, 2917–2942.

Gao Jidong, and Chou Jifan, 1994: Two kinds of inverse problem in NWP and a numerical method-ideal field experiment. *Acta Meteorologica Sinica*, **52(2)**, 129–137.

Gao Jidong, and Chou Jifan, 1995: The effects of the model sensitivity to initial condition upon the variational four-dimensional assimilation—the study based on Lorenz model. *Acta Meteorologica Sinica*, **53(4)**, 471–479.

Hamming, R. W., 1989: *Digital Filters*, Third Edition. Prentice-Hall, Englewood Cliffs, New Jersey, North America, 284pp.

Le Dimet F. X., and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorolagical observations: theoretical aspects. *Tellus*, **38A**, 97–110.

Lewis, J. M., and J. C. Derber, 1985: The use of adjoint equation to solve a variational adjustment problem with advective constraints. *Tellus*, **37A**, 309–322.

Navon, I. M., and D. M. Legler, 1987: Conjugate-gradient methods for large-scale minimization in meteorology. *Mon. Wea. Rev.*, **115**, 1479–1502.

Navon, I. M., X. Zou, J. Derber, and J. Sela, 1992: Variational data assimilation with an adiabatic version of the NMC spectral model. *Mon. Wea. Rev.*, **120**, 1433–1446.

Pielke, R. A., 1984: *Mesoscale Meteorological Modeling*. Orlando, Academic Press, 612 pp.

Talagrand O., and P. Courtier, 1987: Variational assimilation of meteorological observations with the adjoint vorticity equation. Part I: Theory. *Quart. J. Roy. Meteor. Soc.*, **113**, 1311–1328.

Wang Yunfeng, Wu Rongsheng, Wang Yuan, and Pang Yinong, 1999: Application of variational algorithms in semi-Lagrangian framework. *Advances in Atmospheric Sciences*, **16(3)**, 419–430.

Wu Rongsheng, 1990: *Atmospheric Dynamic*. Meteorology Publishing Company, Beijing. 296–304.

Zhu Jiang, 1995: Four-dimensional data quality control: variational method. *Acta Meteorologica Sinica*, **54(3)**, 480–487.

Zou X., Y.-H. Kuo, and Y.-R. Guo, 1995: Assimilation of atmospheric radio refractivity using a nonhydrostatic adjoint model. *Mon. Wea. Rev.*, **123**, 2229–2249.