ADVANCES IN ATMOSPHERIC SCIENCES, VOL. 27, NO. 4, 2010, 741–749

A New Strategy for Solving a Class of Constrained Nonlinear Optimization Problems Related to Weather and Climate Predictability

DUAN Wansuo*[1] (段晚锁) and LUO Haiying[2] (骆海英)

[1] *State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing* 100029

[2] *Training Center, China Meteorological Administration, Beijing* 100081

(Received 26 August 2009; revised 10 October 2009)

ABSTRACT

There are three common types of predictability problems in weather and climate, which each involve different constrained nonlinear optimization problems: the lower bound of maximum predictable time, the upper bound of maximum prediction error, and the lower bound of maximum allowable initial error and parameter error. Highly efficient algorithms have been developed to solve the second optimization problem. And this optimization problem can be used in realistic models for weather and climate to study the upper bound of the maximum prediction error. Although a filtering strategy has been adopted to solve the other two problems, direct solutions are very time-consuming even for a very simple model, which therefore limits the applicability of these two predictability problems in realistic models. In this paper, a new strategy is designed to solve these problems, involving the use of the existing highly efficient algorithms for the second predictability problem in particular. Furthermore, a series of comparisons between the older filtering strategy and the new method are performed. It is demonstrated that the new strategy not only outputs the same results as the old one, but is also more computationally efficient. This would suggest that it is possible to study the predictability problems associated with these two nonlinear optimization problems in realistic forecast models of weather or climate.

Key words: constrained nonlinear optimization problems, predictability, algorithms

Citation: Duan, W. S., and H. Y. Luo, 2010: A new strategy for solving a class of constrained nonlinear optimization problems related to weather and climate predictability. *Adv. Atmos. Sci.*, **27**(4), 741–749, doi: 10.1007/s00376-009-9141-0.

1. Introduction

Optimization methods have been a useful tool in estimating the predictability of weather and climate (Lorenz, 1965; Fan and Chou, 1999; Smith et al., 1999; Mu, 2000; Mu et al., 2003). To quantifying the predictability, Lorenz (1965) applied the linear singular vector (LSV) approach and introduced this method to meteorology. The LSVs represent the optimal perturbations of a linearized model. One can use a Power Method (Packard et al., 1988) to obtain them. Furthermore, if one only needs the leading LSV that represents the fastest-growing initial perturbation of the linearized model, an unconstrained optimization algorithm can also be used to obtain it (Mu and Zhang, 2006), and there have been many highly efficient solvers developed for this task, such as the Limited memory Broyden-Fletcher-Goldfarb-Shanno method (L-BFGS; Liu and Nocedal, 1989), MIQN3 (Liu and Nocedal, 1989), etc.

To reveal the effect of nonlinearity on predictability, Mu et al. (2003) further proposed a novel technique to tackle the optimal initial perturbation in nonlinear models, i.e. conditional nonlinear optimal perturbation (CNOP). CNOPs represent a kind of initial perturbations that have the largest nonlinear evolution and are different from the leading LSV. A distinct advantage of the CNOP approach is that it con-

*Corresponding author: DUAN Wansuo, duanws@lasg.iap.ac.cn

siders the effect of nonlinearity (Mu et al., 2003; Mu and Zhang, 2006). Physically, CNOPs describe initial errors that satisfy a certain constraint and yield the largest prediction errors. Calculating CNOPs requires a constrained nonlinear optimization algorithm, where the sequential quadratic programming (SQP; Powell, 1982) algorithm is useful (Mu et al., 2003), but regardless, the available solvers can still only tackle the problems of small dimensionality (Mu and Zhang, 2006). In predictability studies for weather and climate, the associated numerical models are generally of high dimensionality and the resulting optimization problems are complex. For application in such situations, Mu and Zhang (2006) modified an existing SQP solver and efficiently obtained the CNOPs of a quasi-geostrophic model of 513. Other algorithms are also possible for computing the CNOPs of a complex model. For example, Mu et al. (2007) and Mu et al. (2009) adopted the spectral projected gradient 2 (SPG2; Birgin et al., 2000) solver to successfully calculate the CNOPs of the complex PSU/NCAR Meso-scale model (MM5). All the CNOPs computed in these studies provide useful information for exploring predictability limits for weather and climate predictions (Mu et al., 2003; Mu and Zhang, 2006; Mu et al., 2007; Mu et al., 2009; Duan and Mu, 2009).

Three predictability problems proposed by Mu et al. (2002) also highlight an important approach to studying these kinds of predictability problems. Namely, these problems are associated with the maximum predictability time, the maximum prediction error, and the maximum allowable initial error and parameter error. In that paper, these three problems were formulated into three constrained nonlinear optimization problems. Since the true state of weather and climate cannot be exactly known, the three predictability problems were further reduced into three others, i.e., solutions for the lower bound of the maximum predictability time, the upper bound of maximum prediction error, and the lower bound of the maximum allowable initial error and parameter error. By solving these three reduced nonlinear optimization problems, we can obtain estimations regarding the predictability limits for weather and climate predictions. For convenience, we hereafter refer to these three optimization problems as Problem-1, Problem-2, and Problem-3, respectively.

Problem-2 deals with the upper bound of the maximum prediction errors for a weather or climate event, in which the upper bound can be exactly reached. Problem-2 can also be understood as solving for the largest prediction error caused by the initial errors, subject to certain constraints (Duan and Mu, 2005). In this case, the CNOP is the initial error that causes the largest prediction error. Therefore, Problem-2 is related to the CNOP approach (Duan and Mu, 2005). Consequently, Problem-2 can be solved by an existing highly efficient solver for computing CNOPs.

For Problems 1 and 3, although Mu et al. (2002) and Duan and Mu (2005) have solved these by a filtering method, they did so for two very simple ordinary differential equation models of 2–3 dimensions. Even so, an expensive computation was still necessary. We concede that for a large-scale system of high dimensionality, it would be impossible for us to solve Problems 1 and 3 by the filtering method.

In this paper, a new strategy is designed to efficiently compute answers to Problems 2 and 3. The paper is organized as follows. In next section, the nonlinear optimization problems related to the three predictability problems proposed by Mu et al. (2002) are briefly introduced. In section 3, we design a new strategy for solving the Problems 2 and 3. In section 4, we use a simple model to compare the new and old strategies. Finally, a conclusion and discussion are presented in section 5.

2. Three nonlinear optimization problems related to predictability

In realistic predictions of weather and climate, a numerical model is often useful, in which the model parameters are predetermined, and these parameters each have uncertainties which cause model errors. Also, the initial observations are uncertain. In the three predictability problems proposed by Mu et al. (2002), both these two kinds of uncertainties are considered. To facilitate the description, we signify the initial observations as $\boldsymbol{u}_{0,\mathrm{obs}}$, the initial assumed values of the model parameters as $\boldsymbol{\mu}_{\mathrm{g}}$, the initial observational error (hereafter referred to as initial error) as $\boldsymbol{u}_0$ and the model parameter error as $\boldsymbol{\mu}$. We further assume that M_t and M_T stand for the propagators of a numerical model from the start time 0 stepping forward to prediction times t and T, respectively. Then the three predictability problems in Mu et al. (2002) can be described as follows.

Problem-1 Lower bound of maximum predictable time. Suppose that the maximum allowable prediction error of a weather or climate event is predetermined to be less than ε (the allowable prediction precision; a positive number), i.e.,

$$\|M_t(\boldsymbol{u}_0, \boldsymbol{\mu}) - M_t(\boldsymbol{u}_{0,\mathrm{obs}}, \boldsymbol{\mu}_{\mathrm{g}})\| \leqslant \varepsilon \,. \tag{1}$$

The information about the errors in the initial observations and the initial assumed values of the parameters are also known with the following levels of tolerance:

$$\|\boldsymbol{u}_{0,t} - \boldsymbol{u}_{0,\mathrm{obs}}\|_{\mathrm{A}} \leqslant \delta_1, \qquad \|\boldsymbol{\mu}_t - \boldsymbol{\mu}_{\mathrm{g}}\|_{\mathrm{B}} \leqslant \delta_2 \,, \tag{2}$$

where $\|\cdot\|_{\rm A}$ and $\|\cdot\|_{\rm B}$ are norms measuring the errors in initial conditions and parameters in the model, and $\boldsymbol{u}_{0,t}$ and $\boldsymbol{u}_t$ are the true values of the initial conditions and the model parameters. The lower bound of maximum predictable time can be estimated by the following nonlinear optimization problem

$$T_{\rm l} = \min_{\boldsymbol{u}_0 \in B_{\delta_1}, \boldsymbol{\mu} \in B_{\delta_2}} \{T_{\boldsymbol{u}_0,\boldsymbol{\mu}} | T_{\boldsymbol{u}_0,\boldsymbol{\mu}} = \max \tau , \\ \|M_t(\boldsymbol{u}_0, \boldsymbol{\mu}) - M_t(\boldsymbol{u}_{0,\rm obs}, \boldsymbol{\mu}_{\rm g})\| \leqslant \varepsilon, 0 \leqslant t \leqslant \tau\} , \tag{3}$$

where B_{δ_1} and B_{δ_2} are constraints regions with centers at $\boldsymbol{u}_{0,\rm obs}$, $\boldsymbol{\mu}_{\rm g}$, and radii δ_1, δ_2, respectively.

Problem-2 The upper bound of maximum prediction errors. In this case, information about the initial errors and the parameter errors is known. One hopes to estimate the upper bound of the maximum prediction error at a given prediction time T, which can be evaluated by the following nonlinear optimization problem

$$E_{\rm u} = \max_{\boldsymbol{u}_0 \in B_{\delta_1}, \boldsymbol{\mu} \in B_{\delta_2}} \|M_T(\boldsymbol{u}_0, \boldsymbol{\mu}) - M_T(\boldsymbol{u}_{0,\rm obs}, \boldsymbol{\mu}_{\rm g})\|_{\rm A} . \tag{4}$$

Problem-3 The lower bound of maximum allowable initial error and parameter error. This estimates the required accuracy of the initial conditions and for the model parameters when one forecasts the state at time T. In this case, the maximum allowable prediction error is known to be less than ε. Then, the following nonlinear optimization problem can be used to determine the lower bound of the maximum allowable initial error and parameter error

$$\bar{\delta}_{\max} = \max_{\delta} \{\delta | \ \|M_T(\boldsymbol{u}_{0,\rm obs}, \boldsymbol{\mu}_{\rm g}) - M_T(\boldsymbol{u}_0, \boldsymbol{\mu})\|_{\rm A} \leqslant \varepsilon , \\ \boldsymbol{u}_0 \in B_{\delta_1}, \boldsymbol{\mu} \in B_{\delta_2}, \delta_1 + \delta_2 = \delta\} . \tag{5}$$

Remarks In these three problems, if the errors in the parameter can be ignored, and furthermore the model can be assumed to be perfect, the problems correspond to the "first kind" of predictability introduced by Lorenz (1975); on the other hand, if the initial conditions are considered to be exact, then the corresponding problems become those of the "second kind" of predictability, concerning parameter errors (Lorenz, 1975).

3. The new strategy for solving Problems 1 and 3

When only the initial errors (the parameter errors) are considered, the three predictability problems presented here consist of the first (second) kind of predictability. In Mu et al. (2002) and Duan and Mu (2005), the three problems presented there were related to the first kind of predictability. In those studies, Problem-2 was related to the CNOP approach and was solved efficiently by existing solvers such as SQP and L-BFGS, but Problems 1 and 3 were solved by a filtering method. As mentioned in the introduction, the filtering method is very time-consuming even for a simple model and cannot be used in a more complex model. In this section, we will suggest a new strategy to compute solutions to Problems 1 and 3 for the first kind of predictability, which may provide potential insights for applications in a more complex model.

3.1 *The new strategy for solving Problem-1*

We first consider Problem-1 for "first kind" predictability problems. In this case, Problem-1 becomes

$$T_{\rm l} = \min_{\boldsymbol{u}_0 \in B_{\delta_1}} \{T_{\boldsymbol{u}_0} | T_{\boldsymbol{u}_0} = \max \tau , \\ \|M_t(\boldsymbol{u}_0, \boldsymbol{\mu}_{\rm g}) - M_t(\boldsymbol{u}_{0,\rm obs}, \boldsymbol{\mu}_{\rm g})\| \leqslant \varepsilon, 0 \leqslant t \leqslant \tau\} , \tag{6}$$

and hence, Problem-1 solves for the lower bound of the maximum predictable time induced by the initial errors in B_{δ_1}. From Eq. (6), we notice that for a given initial error in B_{δ_1}, the maximum predictable time ($T_{\boldsymbol{u}_0}$) satisfying the criterion $\|M_t(\boldsymbol{u}_0, \boldsymbol{\mu}_{\rm g}) - M_t(\boldsymbol{u}_{0,\rm obs}, \boldsymbol{\mu}_{\rm g})\| \leqslant \varepsilon$ needs to be first determined. Furthermore, for each initial error in B_{δ_1}, we obtain a corresponding $T_{\boldsymbol{u}_0}$. Of all these $T_{\boldsymbol{u}_0}$, the smallest one is the lower bound of the maximum predictable time.

From the above discussion, it is known that the lower bound of the maximum predictable time, $T_{\rm l}$, corresponds to an initial error in B_{δ_1}, whose resultant prediction error at time $T_{\rm l}$ is the largest among those caused by the initial errors in B_{δ_1}. This motivates us to tackle Problem-1 by solving the problem of the largest prediction error. Furthermore, we have known that there are highly efficient solvers which can com-

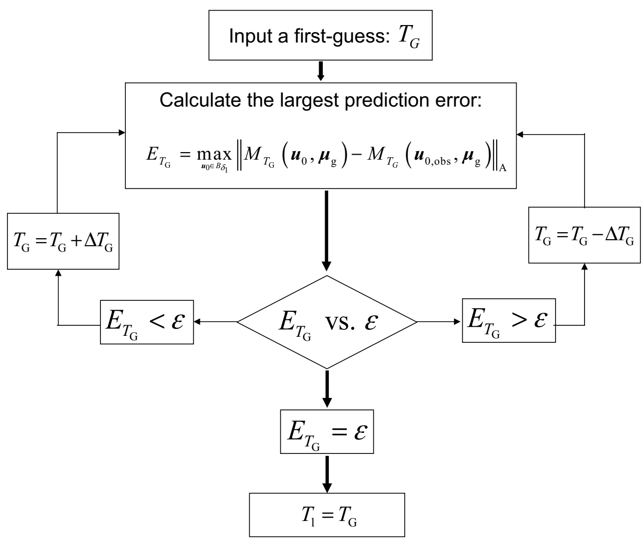

Fig. 1. Flow chart for solving Problem-1

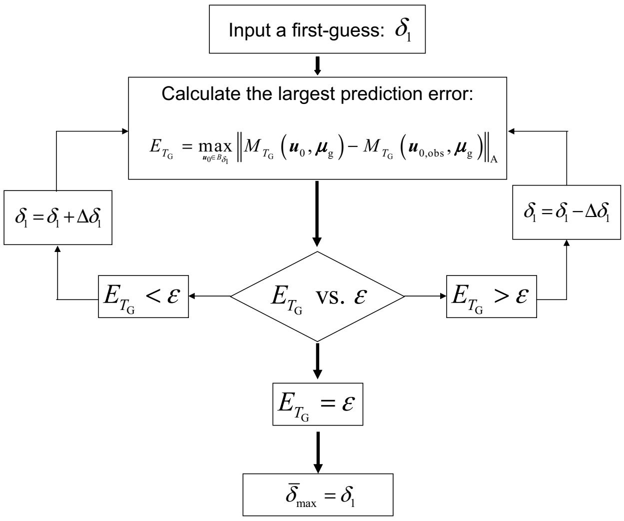

Fig. 2. Flow chart for solving Problem-3.

pute the largest prediction error for Problem-2 (see introduction).

In the filtering method, for each initial error in B_{δ_1} we have to evaluate the resultant prediction error at each prediction time to determine the corresponding maximum predictable time and then search for the smallest one of these maximum predictable times, a process which is very time-consuming. It has been analyzed that we can directly evaluate the largest prediction error by using the existing solvers employed for Problem-2, instead of computing the prediction error caused by each initial error in B_{δ_1} at each lead time. Moreover, highly efficient solvers such as SQP and L-BFGS are based on the fastest descent direction of the gradient of the corresponding objective function, and are therefore able to save considerable computational costs as compared to the filtering method. The main flow of this strategy is as follows.

For a reasonable first-guess, $T_{\rm G}$, of the T_l, we use the L-BFGS, the SPG2 or the modified SQP solver to calculate the largest prediction errors caused by the initial errors in B_{δ_1}, which is denoted by $E_{T_{\rm G}}$. If this largest prediction error is larger than the maximum allowable prediction error ε, we try a smaller $T_{\rm G} = T_{\rm G} - \Delta T_{\rm G}$, and do another computation of the largest prediction error at the update time $T_{\rm G}$; if the prediction error $E_{T_{\rm G}}$ is less than ε, we try a much larger $T_{\rm G} = T_{\rm G} + \Delta T_{\rm G}$. With several tries, if a $T_{\rm G}$ is found to satisfy $E_{T_{\rm G}} > \varepsilon$ for $T_{\rm G} + \Delta T_{\rm G}$ and $E_{T_{\rm G}} \leqslant \varepsilon$ for $T_{\rm G} - \Delta T_{\rm G}$, this $T_{\rm G}$ is approximately equal to the maximum allowable prediction errors ε and acts as the upper bound of the maximum predictable time. To make it clearer, we plot in Fig. 1 the flow chart of this strategy.

3.2 *The new strategy for solving Problem-3*

In the scenario of the first kind of predictability, Problem-3 becomes

$$\bar{\delta}_{\max} = \max_{\delta_1}\{\delta_1 |\, \|M_t(\boldsymbol{u}_0, \boldsymbol{\mu}_{\rm g}) - M_t(\boldsymbol{u}_{0,{\rm obs}}, \boldsymbol{\mu}_{\rm g})\| \leqslant \varepsilon, \boldsymbol{u}_0 \in B_{\delta_1}\}\,. \quad (7)$$

This nonlinear optimization problem addresses the lower bound of the maximum allowable initial error when one attempts to forecast successfully a weather or climate event. In the filtering method, in order to obtain $\bar{\delta}_{\max}$, one has to compute the prediction error caused by each initial error in B_{δ_1} for a first-guess δ_1 of $\bar{\delta}_{\max}$ to see if there exists any initial error in B_{δ_1} whose resultant prediction error does not satisfy the criterion $\|M_t(\boldsymbol{u}_0, \boldsymbol{\mu}_{\rm g}) - M_t(\boldsymbol{u}_{0,{\rm obs}}, \boldsymbol{\mu}_{\rm g})\| \leqslant \varepsilon$. However, it is this step that is particularly computationally costly. Therefore, in any new strategy, this step should be avoided if possible. In fact, from section 3.1 we have know that a similar problem can be solved by computing the largest prediction error among those caused by the initial errors in a given constraint. Thus, one does not need to compute the prediction error caused by each initial error in the constraint, and we only compute the largest prediction error and compare it with the maximum allowable prediction precision ε. In the new strategy for solving Problem-3, since we directly calculate the largest prediction error by using the algorithm used for solving Problem-2, the cost is largely reduced. The corresponding flow chart is illustrated in Fig. 2.

In this section, we have designed new strategies for solving Problems 1 and 3. These strategies were established for the computation of the largest prediction error related to Problem-2. In other words, the new strategies can directly calculate the largest prediction error and do not need to compute the prediction error caused by each initial error in the given constraint. This may result in the new strategies being highly efficiency. To verify this, two questions need to be addressed: Were the results obtained by the new strategies the same as those obtained by the old method? What are the computational costs of the new strategies?

4. Comparison between the new strategy and the old one

We use a simple El Niño-Southern Oscillation (ENSO) model (Wang and Fang, 1996, hereafter WF96) to show the high efficiency of the new strategies for solving Problems 1 and 3. The WF96 model describes the interannual variation of SSTA, $T_{\rm E}$, and thermocline anomaly, $h_{\rm E}$, in the Niño-3 region:

$$\begin{cases} \dfrac{dT_{\rm E}}{dt} = a_1 T_{\rm E} - a_2 h_{\rm E} + \sqrt{\dfrac{2}{3}} T_{\rm E}(T_{\rm E} - a_3 h_{\rm E}) \ , \\ \dfrac{dh_{\rm E}}{dt} = b(2h_{\rm E} - T_{\rm E}) \ , \end{cases}$$

where

$$\begin{aligned} a_1 &= (\Delta T_0' - \bar{u}_1' + \overline{T}_x' - \alpha_s')\big|_{x_{\rm E}} \ , \\ a_2 &= (\kappa + \sigma_1)\overline{T}_x'|_{x_{\rm E}} \ , \\ a_3 &= (\kappa + \sigma_1), \\ b &= \frac{2}{\sigma(1 - 3\varepsilon)} \ . \end{aligned}$$

The parameters, $\Delta T_0'$, $\bar{T}_x'$, and $\bar{u}_1'$, are time-dependent and determined by the climate mean state; κ, σ, ε, and σ_1 are all constant. For values of these parameters, readers are referred to WF96. In this paper, the WF96 model is discretized by a fourth-order Runge-Kutta scheme with $dt = 0.01$.

Duan and Mu (2005) have used this model to demonstrate that reference states with initial values of $\boldsymbol{U}_0 = (T_{\rm E0}, h_{\rm E0})$ equal to $(-0.1, 0.1)$ and $(0.1, -0.1)$ respectively represent El Niño and La Niña events in the WF96 model. Furthermore, Duan and Mu (2005) regarded these as two theoretical hindcast modes to be predicted. In this paper, we also use these two events as basic states to solve the corresponding Problem-1 and Problem-3. For convenience, we denote these two events as $U_{\rm E}$ and $U_{\rm L}$, respectively. The norm $\|\boldsymbol{u}_0\| = \sqrt{T_{\rm E0}'^2 + h_{\rm E0}'^2}$ is used to measure the magnitude of the initial errors, where $T_{\rm E0}'$ and $h_{\rm E0}'$ represent the errors superimposed on initial values of $T_{\rm E0}$ and $h_{\rm E0}$, respectively. Then the constraint condition $\boldsymbol{u}_0 \in B_{\delta_1}$ in Problems 1 and 3 are equivalent to $\|\boldsymbol{u}_0\| \leqslant \delta_1$, i.e.,

$$\sqrt{T_{\rm E0}' 2 + h_{\rm E0}'^2} \leqslant \delta_1 \ .$$

With these prescripts, we estimate the lower bound of maximum predictable time and the maximum allowable initial errors of the above two events.

4.1 *Problem-1*

In the first kind of predictability studies, the WF96 model is assumed to be perfect. With the above notation and the given measurements, Problem-1 becomes

$$\begin{aligned} T_l = \min_{\|\boldsymbol{u}_0\| \leqslant \delta_1} \{ & T_{\boldsymbol{u}_0} | T_{\boldsymbol{u}_0} = \max \tau \ , \\ & \|M_t(\boldsymbol{U}_0 + \boldsymbol{u}_0) - M_t(\boldsymbol{U}_0)\| \leqslant \varepsilon, 0 \leqslant t \leqslant \tau\} \ . \end{aligned}$$

Now we compare the new strategy of computing a solution to Problem-1 with the old method used in Duan and Mu (2005). For convenience, we first briefly introduce here the old strategy. For a constraint disk

$$\sqrt{T_{\rm E0}'^2 + h_{\rm E0}'^2} \leqslant \delta_1 \ ,$$

its corresponding circumscribed square is considered. Foursquare-meshes of a certain size are used to discretize the circumscribed square. For any mesh point outside the disk, it is connected with the center of the disk, and the intersection point of this line with the boundary of the disk is obtained. Integrating the WF96 model from each of these intersection points and for the mesh points inside the disk and then comparing the prediction error caused by each initial error (represented by the mesh point in the constraint disk) with the prediction precision ε, the maximum predictable time $T_{\boldsymbol{u}_0}$ can be obtained. For all these maximum predictable times, the smallest one could be the lower bound of the maximum predictable time. It is conceivable that the lower bound of the maximum predictable time computed by this old strategy may depend on the size of the foursquare-meshes used for discretizing the constraint. Therefore, in the following numerical experiments, we perform comparisons of the new strategy to the old one with different sizes of foursquare-meshes.

In numerical experiments, the initial error bound δ_1 is chosen as 0.06, 0.07, 0.08, 0.09, and 0.10, and the maximum allowable prediction error ε is predetermined as 0.35, 0.40, 0.45, and 0.50, all of which are the same as those in Duan and Mu (2005). In the old strategy, several kinds of foursquare-meshes are used to discretize the circumscribed square of the constraint disk. By computing the lower bound of the maximum predictable time of the above El Niño and La Niña events based on the old strategy, it is found that when the size of the foursquare-mesh is decreased from 0.01 to 0.0001, the resultant lower bounds of the maximum predictable time become accordant for the foursquare-meshes whose size is smaller than 0.001. The lower bounds of the maximum predictable time obtained by the old strategy are related to the mesh sizes. It is believable that the smaller the foursquare-mesh, the more accurate the lower bound of the maximum predictable time computed by the old strategy. We would rather accept the results of the meshes with small sizes. However, in the new strategy, the results are independent of the mesh sizes. Then, we wish to know whether the predictabilities for Problem-1 estimated by the new strategy are the same as those computed by the old strategy.

According to the flow described in Fig. 1, we use the new strategy to obtain the lower bound of maximum predictable time for the chosen El Niño and La Niña events, which are listed in Tables 1 and 2, re-

Table 1. The lower bound of the maximum predictable time for the El Niño event $U_{\rm E}$ (obtained by the new strategy).

ε	$\delta_1 = 0.06$	$\delta_1 = 0.07$	$\delta_1 = 0.08$	$\delta_1 = 0.09$	$\delta_1 = 0.10$
0.35	267	204	173	154	139
0.40	317	216	183	161	146
0.45	403	229	193	169	152
0.50	420	246	204	177	159

spectively. The data shown in Tables 1 and 2 are the same as those obtained by the old strategy with the foursquare-mesh of sizes smaller than 0.001 [see also Tables 1 and 2 in Duan and Mu (2005)]. This indicates that the new strategy holds for solving Problem-1.

In addition, it has been mentioned that the new strategy does not use a foursquare-like mesh to discretize the constraint disk and, therefore, avoids the uncertainties induced by the choice of different mesh sizes in the old strategy. In this case, we are subsequently concerned with the computational costs of the new strategy.

From section 3, we have indicated that the main difference between the new strategy and the old one consists of the computation of the largest prediction error at the first guess $T_{\rm G}$ of $T_{\rm l}$,

$$E_{T_{\rm G}}(\boldsymbol{u}_{0\delta}) = \max_{\|\boldsymbol{u}_0\|\leqslant\delta_1} \|M_{T_{\rm G}}(\boldsymbol{U}_0 + \boldsymbol{u}_0) - M_{T_{\rm G}}(\boldsymbol{U}_0)\| ,$$

where $\boldsymbol{u}_{0\delta}$ is the initial error that causes the largest prediction error. In the filtering method, the solver searches for the lower bound of the maximum predictable time by comparing the prediction errors caused by the initial errors in the constraint disk with each other. It is well known that the lower bound of maximum predictable time corresponds to the largest prediction error at this time. This implies that the filtering method is in essence looking for the largest prediction error among those caused by the initial errors in the constraint disk. However, in order to obtain this result, the prediction error caused by each initial error (i.e., each at mesh point) in the disk must be calculated at each lead time. It is conceivable that for a complex model and a complicated constraint, the computation of the lower bound of maximum predictable time by the filtering method is almost impossible. However, in the new strategy, the computation of the largest prediction error is performed by a nonlinear optimization algorithm, which searches for the initial error that causes the largest prediction error along the descent direction of the gradient of the objective function. The algorithm to do this has been verified to be highly efficient. Furthermore, the new strategy does not use a foursquare-mesh to discretize the constraint and is independent of the chosen mesh size, although the prediction errors are also computed at each lead time. Despite this, we still perform a simple comparison between the new strategy and the old one to show the high efficiency of the new strategy. In Tables 3 and 4, we list the computational costs of computing the largest prediction errors of the old strategy on different foursquare-meshes. We also do the same for the new method given a first guess at $T_{\rm G} = 8$ months and an initial constraint bound of 0.08, where "Mesh-size" denotes the size of the foursquare-mesh discretizing the constraint, "$E_{T_{\rm G}}$" represents the largest prediction error, and "CPU-time" signifies the computational cost of the largest prediction error with one CPU. In the new strategy, the largest prediction error is computed by using the SQP solver, which is used to solve the optimization problem with equality and inequality constraints.

From Tables 3 and 4, it is demonstrated that the new strategy shows has negligible computational cost to calculate the largest prediction error related to Problem-1, but costs of the old method are proportional to the mesh sizes. While the results are accordant for the mesh sizes smaller than 0.001, computational costs are equivalent to at least 6 seconds of single-CPU time for the above first guess $T_{\rm G}$ and initial constraint 0.08. The computational cost of the old strategy is at least 260 times larger than for the new strategy and is quite intensive even for the simple model adopted here. Thus, the filtering method would be rather insufferable for a complex model; the negligible costs of the new strategy favors its application for a more complex weather/climate model. As mentioned

Table 2. The lower bound of the maximum predictable time for the La Niña event $U_{\rm L}$ (obtained by the new strategy).

ε	$\delta_1 = 0.06$	$\delta_1 = 0.07$	$\delta_1 = 0.08$	$\delta_1 = 0.09$	$\delta_1 = 0.10$
0.35	1032	1019	765	639	523
0.40	1036	1024	780	648	531
0.45	1040	1026	957	658	538
0.50	1045	1029	1021	671	546

Table 3. Comparison between the old and the new strategies for the El Niño event U_{E}.

		Mesh-size				
		0.0100	0.0050	0.001	0.0005	0.0001
$E_{T_{\mathrm{G}}}$	Old strategy	0.4919	0.4921	0.4925	0.4925	0.4925
	New strategy	0.4925	0.4925	0.4925	0.4925	0.4925
CPU-time (s)	Old strategy	0.0600	0.2550	6.0700	23.6000	599.5000
	New strategy	0.0240	0.0240	0.0240	0.0240	0.0240

above, the lower bound of the maximum predictable time obtained by the old strategy depends on the mesh size, but the new strategy is mesh size independent. Furthermore, with increasing mesh resolution in the constraint disk and then decreasing the mesh sizes, the lower bound of the maximum predictable time obtained by the old strategy gradually converges to that obtained by the new strategy. The new strategy is highly efficient, which suggests the possibility of studying Problem-1 in a realistic weather or climate model.

4.2 *Problem-3*

Regarding the scenario of the first kind of predictability, Problem-3 is as follows:

$$\bar{\delta}_{\max} = \max_{\delta_1}\{\delta_1 |\ \|M_T(\boldsymbol{U}_0+\boldsymbol{u}_0) - M_T(\boldsymbol{U}_0)\|_{\mathrm{A}} \leqslant \varepsilon, \|\boldsymbol{u}_0\| \leqslant \delta_1\}\ .$$

Duan and Mu (2005) adopted the filtering method to solve Problem-3 related to the El Niño event U_{E} and the La Niña event U_{L}. The filtering method in this case uses the following scheme: For a first guess δ_1, there exists a corresponding circumscribed square. A foursquare mesh of size 0.01 is used to partition the circumscribed square of the constraint disk $\|\boldsymbol{u}_0\| \leqslant \delta_1$ and its mesh-points are obtained. For the points outside the disk, we connect these points with the center of the disk and obtain the intersection points of the lines with the boundary of the disk. Integrating the WF96 model from these intersection points and for each point inside the disk, we obtain the prediction errors caused by the initial errors represented at the mesh-points. If all these prediction errors satisfy

$$\|M_T(\boldsymbol{U}_0+\boldsymbol{u}_0) - M_T(\boldsymbol{U}_0)\| \leqslant \varepsilon\ ,$$

we try another δ_1 larger than this one and do another computation of the prediction errors for the updated δ_1. Step by step, we finally find the maximum value, $\bar{\delta}_{\max}$.

Although the filtering method has been used to tackle Problem-3 by Duan and Mu (2005), it is very time-consuming, and so we cannot apply it in a complex model. In fact, in the filtering method, for the first guess δ_1 and the subsequent updated values, we must calculate the prediction error caused by each initial error in $\|\boldsymbol{u}_0\| \leqslant \delta_1$ to see if it satisfies the criterion

$$\|M_T(\boldsymbol{U}_0+\boldsymbol{u}_0) - M_T(\boldsymbol{U}_0)\| \leqslant \varepsilon\ .$$

It is this step that costs a large part of the computation time of solving for $\bar{\delta}_{\max}$. However, in the new strategy suggested in section 3, we do not need to calculate the prediction errors caused by all the initial errors in the constraint, and instead directly evaluate the largest prediction error by using a ready-made nonlinear optimization algorithm such as SQP (or SPG2) or L-BFGS. In the latter case, we only judge if the largest prediction error is less than the maximum allowable prediction error ε.

In order to verify the effectiveness of the new strategy for solving Problem-3, a series of experiments are performed. For a given start time January, we choose $\varepsilon = 0.35,\ 0.40,\ 0.45,\ 0.50$ as the maximum allowable prediction errors to constrain the accuracy of the predictions, where the prediction times are for August, October, and December of the year. In Table 5, we list the lower bound of the maximum allowable initial error for the El Niño event U_{E} by using the new strategy. The results demonstrate that the lower bound of the maximum allowable initial error determined by the new strategy are almost the same as those determined by the old strategy [see Table 3 in Duan and Mu

Table 4. Comparison between the old and the new strategies for the La Niña event U_{L}.

		Mesh-size				
		0.0100	0.0050	0.0010	0.0005	0.0001
$E_{T_{\mathrm{G}}}$	Old strategy	0.3538	0.3540	0.3542	0.3542	0.3542
	New strategy	0.3542	0.3542	0.3542	0.3542	0.3542
CPU-time (s)	Old strategy	0.0700	0.2600	6.0100	23.9000	600.2000
	New strategy	0.0230	0.0230	0.0230	0.0230	0.0230

Table 5. The lower bound of the maximum allowable initial error ($\bar{\delta}_{\max}$) for the E1 Niño evert $U_{\rm E}$ (obtained by the new strategy).

Time period	$\varepsilon = 0.35$	$\varepsilon = 0.40$	$\varepsilon = 0.45$	$\varepsilon = 0.50$
Jan–Aug	0.0633	0.0654	0.0686	0.0704
Jan–Oct	0.0586	0.0606	0.0644	0.0674
Jan–Dec	0.0516	0.0551	0.0614	0.0641

Table 6. Comparison between the old and the new strategies for the El Niño event $U_{\rm E}$ with the time period January–December and a first guess $\delta_1 = 0.07$.

		Mesh-size				
		0.0100	0.0050	0.0010	0.0005	0.0001
$E_{T_{\rm G}}$	Old strategy	0.4249	0.4253	0.4255	0.4255	0.4255
	New strategy	0.4255	0.4255	0.4255	0.4255	0.4255
CPU-time (s)	Old strategy	0.0700	0.2700	6.4500	24.8000	628.3200
	New strategy	0.0700	0.0700	0.0700	0.0700	0.0700

(2005)], where the foursquare-meshes of different sizes were also used to discretize the constraint in the old strategy and the results obtained by foursquare mesh with sizes smaller than 0.001 are coherent.

As mentioned above, in the computation for Problem-3, differences between the new strategy and the old one are also evident with respect to the computation of the largest prediction error. In numerical experiments, we also compare the new strategy and the old one based on the two aspects of the accuracy of the largest prediction error and the costs. Similar results are obtained. In fact, we demonstrate that in computing the lower bound of the maximum allowable initial error, the largest prediction error obtained by the old strategy depends on the mesh sizes. Actually, only when the mesh sizes are larger than 0.001 do the resultant largest prediction errors remain unchanged with increasing resolution of the mesh in the constraint disk. Nevertheless, the time costs becomes more and more expensive. For the new strategy, the computation of the largest prediction error is independent of the mesh and has negligible cost. Furthermore, its resultant largest prediction error is the same as that of the old strategy with the mesh sizes smaller than 0.001. It is believable that, in the old strategy, the smaller the mesh size, the more accurate the resultant largest prediction error. That is to say, the unchanged largest prediction error obtained with mesh sizes smaller than 0.001 may be the exact largest prediction error. Thus, the new strategy not only yields the same results as the old strategy with small mesh sizes, but also has negligible costs for the simple model adopted here. In Table 6, we show the results of the comparison between the old strategy and the new one for the El Niño event $U_{\rm E}$ with time period January–October and a first guess $\delta_1 = 0.07$. It is obvious that the new strategy may be more applicable than the old one when applied in a complex model.

The maximum allowable initial errors of the La Niña event $U_{\rm L}$ are also computed by the new strategy and the same results were obtained as by the old strategy. Furthermore, the new strategy has negligible computational cost for the largest prediction error compared to the old strategy. For simplicity, the details are omitted here.

In summary, we demonstrate that new strategies can effectively solve Problems 1 and 3 for the first kind of predictability. These strategies similarly extend to the corresponding problems of the second kind of predictability. In this case, we only need to consider errors superimposed on the model parameters. For simplicity, we do not describe the details here.

5. Conclusion and discussion

In this paper, we propose a strategy for solving for the lower bounds of the maximum predictable time and the maximum allowable initial errors of the first kind of predictability, which can also be extended to solve problems of the second kind of predictability. This strategy has been compared to an existing filtering method. A series of comparisons show that the results of new strategy are almost the same as those of the old method; furthermore, the methods presented here can save a large amount of computation time. The most advantageous aspect of the new strategy is the computation of the largest prediction errors caused by initial errors in a given constraint. The largest prediction error in the new strategy is calculated by an existing optimization solver, which has been verified to be highly efficient, whereas in the old method, the largest prediction error is filtered out by comparing the prediction errors caused by the initial errors in the given constraint. It is conceivable that using the filtering method is impossible for solving the above two predictability problems in a complex and realistic model. But with the new strategy, since a highly efficient optimization solver is used, this provides the possibility of studying these two predictability problems for a realistic weather and climate model.

The new strategy is proposed to solve for the lower bound of maximum predictable time and the maximum allowable initial error and parameter error and offers considerable computation savings as compared to the old strategy. Still, we should realize its limitations. In choosing a reasonable first-guess value of the new optimization strategy, we have to try several values due to imperfect initialization. Despite thus, the new strategy of computing the lower bound of maximum predictable time still represents progress as compared to the old strategy because the filtering method must try each lead time and each mesh point to determine the lower bound of maximum predictable time, which may require far more computational cost. Therefore, our next work will be to study a feasible initialization for the new strategy. In a word, there is still much more work to be done to improve the strategy. It is expected that an even more highly efficient algorithm can be achieved with the help of computational mathematician.

Acknowledgements. This work was jointly sponsored by the Key Knowledge Innovation Program of the Chinese Academy of Sciences (Grant. No. KZCX2-YW-QN203), the National Basic Research Program of China (2007CB411800), the GYHY200906009 of China Meteorological Administration.

REFERENCES

Birgin, E. G., J. M. Martınez, and M. Raydan, 2000: Nonmonotone spectral projected gradient methods on convex sets. *SIAM Journal on Optimization*, **10**, 1196–1211.

Duan, W. S., and M. Mu, 2005: Applications of nonlinear optimization methods to quantifying the predictability of a numerical model for El Niño-Southern Oscillation. *Progress in Natural Science*, **15**, 915–921.

Duan, W. S., and M. Mu, 2009: Conditional nonlinear optimal perturbation: Applications to stability, sensitivity, and predictability. *Science in China* (*D*), **52**(7) 884–906.

Fan, X. G., and J. F. Chou, 1999: Methods and experiments of numerical prediction raised as inverse problem. Part I: Three kinds of inverse problems and numerical solutions. *Chinese J. Atmos. Sci.,* **23**, 543–550. (in Chinese)

Liu, D. C., and J. Nocedal, 1989: On the limited memory BFGS method for large scale optimization. *Mathematical Programming*, **45**, 503–528.

Lorenz, E. N., 1975: Climate predictability, Appendix 2.1. Global Atmospheric Research Programme Publication Series, No. 16, World Meteorology Organization Geneva, 265pp.

Lorenz, N., 1965: A study of the predictability of a 28-variable atmosphere model. *Tellus*, **17,** 321–333.

Mu, M., 2000: Nonlinear singualr vectors and nonlinear singular values. *Science in China* (*D*), **43**, 375–385.

Mu, M., and Z. Y. Zhang, 2006: Conditonal nonlinear optimal perturbation of a barptrophic model. *J. Atmos. Sci.*, **63**, 1587–1604.

Mu, M, W. S. Duan, and J. C. Wang, 2002: The predictability problems in numerical weather and climate prediction. *Adv. Atmos. Sci.*, **19**, 191–204.

Mu, M., W. S. Duan, and B. Wang, 2003: Conditional nonlinear optimal perturbation and its applications. *Nonlinear Processes in Geophysics*, **10**, 493–501.

Mu, M., H. L. Wang, and F. F. Zhou, 2007: A preliminary application of conditional nonlinear optimal perturbation to adaptive observation. *Chinese J. Atmos. Sci.*, **31**, 1102–1112. (in Chinese)

Mu, M., F. F. Zhou, H. L. Wang, 2009: A method to identify the sensitive areas in targeting for tropical cyclone prediction: conditional nonlinear optimal perturbation. *Mon. Wea. Rev.*, **137,** 1623–1639.

Packard, A., M. K. H. Fan, and J. Doyle, 1988: A power method for the structured singular value. Decisions and Control. *Proc. 27th IEEE Conference*, **3,** 2172–2137.

Powell, M. J. D. 1982: VMCWD: A FORTRAN subroutine for constrained optimization. DAMTP Report 1982/NA4, University of Cambridge, England.

Smith, L. A., C. Ziehmann, and K. Fraedrich, 1999: Uncertainty dynamics and predictability in chaotic systems. *Quart. J. Roy. Meteor. Soc.*, **125**, 2855–2886.

Wang, B., and Z. Fang, 1996: Chaotic oscillation of tropical climate: A dynamic system theory for ENSO. *J. Atmos. Sci.*, **53**, 2786–2802.