Conditional nonlinear optimal perturbations as the optimal precursors for El Nino–Southern Oscillation events

Wansuo Duan and Mu Mu
LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China

Bin Wang
Department of Meteorology, School of Ocean and Earth Science and Technology, University of Hawaii, Honolulu, Hawaii, USA

Received 11 March 2004; revised 16 July 2004; accepted 10 September 2004; published 4 December 2004.

[1] We used the approach of conditional nonlinear optimal perturbation (CNOP) to investigate optimal precursors for El Nino–Southern Oscillation (ENSO) events with a theoretical coupled ocean-atmosphere model. The CNOPs of the annual cycle of the coupled system were computed for different time periods, and the derived CNOPs were compared with the linear singular vectors (LSVs). The results show the existence of the CNOPs of annual cycle and local CNOPs. These CNOPs have the robust optimal patterns, which have opposite polarities in sea surface temperature and thermocline depth anomalies in the eastern equatorial Pacific. We demonstrate that the CNOP (local CNOP), rather than LSVs, has the highest likelihood to develop into an El Nino (La Nina) event; thus the CNOPs (local CNOPs) can be regarded as the optimal precursors for El Nino (La Nina) events. These optimal precursors agree qualitatively well with the observations of period of 1980–2002. On the basis of the nonlinear oscillation described by the model, the physical mechanism of the optimal precursors for ENSO is discussed. 

INDEX TERMS: 3220 Mathematical Geophysics: Nonlinear dynamics; 4522 Oceanography: Physical: El Nino; 3299 Mathematical Geophysics: General or miscellaneous; KEYWORDS: ENSO model, nonlinear, optimal perturbations


1. Introduction

[2] Numerous models have been developed to simulate and to predict the El Nino–Southern Oscillation (ENSO). These models range from simple analytical [Wang and Fang, 1996; Jin, 1997a, 1997b; Wang et al., 1999] and intermediate coupled numerical models [Cane et al., 1986; Zebiak and Cane, 1987; Kleeman, 1993; Kleeman et al., 1995] to sophisticated coupled general circulation models (GCMs), which have simulated ENSO with varying degrees of success, but notable discrepancies between models and reality remain [Latif et al., 1994; Xue et al., 1997a]. Recent works have shown that it is of great significance for improving ENSO predictability to find out the precursors for ENSO events [Palmer et al., 1994; Moore and Kleeman, 1996; Xue et al., 1997a, 1997b; Thompson, 1998; Samelson and Tziperman, 2001].

[3] Theoretical studies have been devoted to searching for ENSO precursors. It is realized that the eigenmodes could not be the fastest growing perturbation in non-self-adjoint system. Xue et al. [1994], using a series of linear Markov models as an approximation to the Zebiak and Cane [1987] model, computed linear singular vectors (LSVs) and found that the fastest growing singular vector evolved into an ENSO event. Palmer et al. [1994] presented the fastest growing singular vector that evolves into a structure resembling ENSO with fast growth rate during April. Moore and Kleeman [1996] further investigated nonlinear evolution of singular vectors by use of the intermediate coupled model of Kleeman [1993]. It is demonstrated that their singular vectors have the potential to develop into ENSO events. Their SVs have the same structures as those described by Xue et al. [1994] and Palmer et al. [1994]. Recently, Thompson [1998] also used LSV to study the characteristic precursor to an ENSO warm event and found a number of interesting characteristics of the precursors lead to ENSO events.

[4] The aforementioned studies have attempted using the fastest growing perturbation to identify optimal growing initial patterns, which will be entitled the optimal precursors for ENSO in this paper. Note that the linear theory of SV assumes that the evolution of the initial perturbation can be described approximately by the tangent linear model (TLM). Owing to the absent of nonlinearity, TLM cannot always describe the maximum growth of the initial perturbation in nonlinear models [Oortwijn and Barkmeijer, 1995; Mu, 2000; Mu and Wang, 2001; Mu et al., 2003]. Thus it remains questionable as to whether the LSV can be regarded as the optimal precursors when nonlinearity appears in the ENSO model.
To reveal nonlinear characteristics of the coupled atmosphere and ocean system, Mu et al. [2003] proposed a novel concept, the conditional nonlinear optimal perturbation (CNOP). Their results suggest that CNOP is potentially more suitable than LSV in studying ENSO predictability. This paper reports a primary study on searching for ENSO precursor using CNOP. We will briefly review the ideas about CNOP in the next section and introduce the conceptual model for ENSO in section 3, then use them to investigate the optimal precursors for ENSO.

2. Conditional Nonlinear Optimal Perturbation

We write the evolution equations for the state vector \( w \), which may include surface current, thermocline depth, and sea surface temperature, etc., into an initial value problem:

\[
\begin{aligned}
\frac{\partial w}{\partial t} + F(w) &= 0, \quad \text{in } \Omega \times [0, T] \\
w_{t=0} &= w_0, 
\end{aligned}
\]

(1)

where \( w(x, t) = (w_1(x, t), w_2(x, t), \ldots, w_p(x, t)) \), \( F \) a nonlinear operator, and \( w_0 \) the initial state, \( (x, t) \in \Omega \times [0, T] \), \( \Omega \) a domain in \( \mathbb{R}^n \), and \( T < \infty \). \( x = (x_1, x_2, \ldots, x_n) \), \( t \) the time. Assume that the dynamical system equations and the initial state are known exactly, the future state can be determined by integrating equation (1) with the appropriate initial condition. The solution to equation (1) for the state vector \( w \) at time \( \tau \) is given by

\[
w(x, \tau) = M_\tau(w_0).
\]

(2)

Here, \( M_\tau \) is the propagator, which, as described by (2), "propagates" the initial value to the time \( \tau \) in the future. Let \( U(x, t) + u(x, t) \) be the solutions of problem (1) with initial value \( U_0 \) and \( U_0 + u_0 \) respectively, where \( u_0 \) is the initial perturbation. We have

\[
U(\tau) = M_\tau(U_0), \quad U(\tau) + u(\tau) = M_\tau(U_0 + u_0).
\]

(3)

So \( u(\tau) \) describes the evolution of the initial perturbation \( u_0 \).

For a chosen norm \( \| \cdot \| \), an initial perturbation \( u_{08} \) is called CNOP under the constraint \( \|u_0\| \leq \delta \), if and only if

\[
J(u_0) = \max_{\|u_0\| \leq \delta} J(u_0),
\]

where

\[
J(u_0) = G(M_\tau(U_0 + u_0) - M_\tau(U_0)),
\]

(4)

where \( G(\cdot) \) measures the evolution of the perturbation. Particularly it can be a norm (巴拉克 \& \| \cdot \|) of the state variable or the module of a variable (巴拉克 \| \cdot \|). We emphasize that CNOP \( u_{08} \) is the global maximum of \( J(u_0) \) in the ball \( \|u_0\| \leq \delta \). In case that there exists local maximum \( u_{00} \) of \( J(u_0) \), we call \( u_{00} \) a local CNOP.

CNOP can be computed by using sequential quadratic programming (SQP) solver, which is used to solve the nonlinear minimization problems with equality or/and in-equality constraint condition [Powell, 1982]. A brief description of the algorithms is given in Appendix A.

3. ENSO Model

The model we used in this paper is the theoretical coupled ocean-atmosphere model of Wang and Fung [1996] (hereinafter referred to as WF96), which has been used to investigate the predictability of ENSO by Mu and Duan [2003]. The ocean component of this coupled model is distilled from the ocean component of Zebiak and Cane [1987] model. The ocean model consists of two time-dependent equations that describe variations of thermocline depth anomaly and SST anomaly, respectively. The equation for thermocline depth anomaly was derived from linear shallow water dynamics of the upper ocean under long wave approximation; it describes adjustment of the thermocline by oceanic Kelvin and Rossby wave propagations and by local wind stress forcing. The SST equation was derived from the thermodynamic equation of surface layer, which describes how the thermocline fluctuation affects SST through nonlinear vertical advection of temperature by wind-induced upwelling. The atmospheric component is a steady version of the Lindzen and Nigam [1987] model, in which SST gradients drive the surface wind variations. Anomalies in SST cause anomalous winds, which not only generate oceanic waves to alter thermocline depth but also drive anomalous currents and upwelling, both in turn affecting SST. Of note is that the winds affect SST in nonlinear fashion through changing both the thermocline depth (thus the vertical gradient of temperature) and the strength of the upwelling. The steady atmosphere acts as a medium to connect SST and thermocline variations, enabling the SST and thermocline depth equations forming a closed set of governing equations for the coupled tropical ocean and atmosphere, the ENSO system as termed in WF96.

To the lowest order approximation, the spatial structure of ENSO is characterized by an equatorial symmetric, east-west standing oscillation. Such a seesaw-like structure has been observed in SST, sea level pressure [Trenberth and Shea, 1987], and thermocline fields [Wang et al., 1999]. This basin-scale spatial structure provides a physical basis for further simplifying the ENSO system. By assuming equatorial trapped meridional structure and east-west seesaw structure, application of Lorenz’s [1963] method of spectral truncation yields a nonlinear dynamic system model which consists of two ordinary differential equations that describe temporal variations of the SST and thermocline anomalies in the equatorial eastern Pacific. This dynamic system model captures the essence of the nonlinear coupling between the surface layer thermodynamics and the upper ocean dynamics. The two dimensionless equations describ-
ing, respectively, the evolutions of the anomalous SST $T$ and the anomalous thermocline depth $h$ in the equatorial eastern Pacific are:

\[
\begin{align*}
\frac{dT}{dt} &= a_1 T - a_2 h + \sqrt{\frac{2}{3} T(T - a_3 h)}, \\
\frac{dh}{dt} &= h(2h - T),
\end{align*}
\]

where

\[
\begin{align*}
a_1 &= (\bar{T}_z + \bar{T}_x - \bar{\alpha}_1')|_{x_0}, \\
a_2 &= (\mu + \delta_1)T_x'|_{x_0}, \\
a_3 &= \mu + \delta_1 \\
b &= \frac{2\alpha}{p(1 - 3\alpha^2)}.
\end{align*}
\]

The linear terms in the $T$-Eq describe the vertical advection by the anomalous upwelling of the mean SST ($\bar{T}_z T$) and the vertical advection by the mean upwelling of the anomalous SST ($\bar{T}_x(T - \mu h)$), and the linear damping ($-\alpha_1' T$). The linear coefficients are determined by the basic state. The coefficients $a_1$ and $a_2$ involve basic state parameters $\bar{T}_z$ and $\bar{T}_x$, which characterize, respectively, the mean temperature difference between the eastern and western basins and between the surface and subsurface water. Note that, these basic state parameters can be time-dependent, reflecting the climatological annual cycle of the basic state. The quadratic term in $T$-Eq comes from the nonlinear temperature advection by anomalous upwelling of the anomalous temperature. This term represents the nonlinear coupling between surface layer thermodynamics and upper ocean dynamics (thermocline depth fluctuations). The linear terms in $h$-Eq depict respectively, the effect of equatorial waves on thermocline adjustment ($2b(1+\mu)$) and the effect of the wind forcing ($-bh(1T)$). Two nondimensional coupling parameters are presented in this model. One is the air-sea coupling coefficient, $\alpha = (L_0 L_s)^2$, where $L_0$ is the oceanic Rossby radius of deformation and $L_s$ is the characteristic meridional length scale of the coupled ENSO mode. For $L_0 = 300 \text{ km}$, when $L_s$ varies from 1000 km to 400 km, $\alpha$ increases from 0.09 to 0.5625. Another coupling parameter is $\mu = \frac{\omega}{t_0}$, which measures the degree of coupling between thermocline fluctuation and SST. The meanings and typical values of the other parameters are listed in Table 1 of WF96 and Wang et al. [1999].

The steady solution $O(0, 0)$ represents the climatological mean equilibrium state or an ENSO “transitional” state (including annul cycle) in which both SST and the depth of thermocline are normal. In this paper, the model is integrated by fourth-order Runge-Kutta scheme with $dt = 0.01$, which represents one day.


WF96 model is an anomalous model about the annual cycle. Let $u_0$ be an initial anomaly. To describe the maximum evolution of SSTA at prediction time $\tau$, following (4) in section 2, we define the nonlinear optimization problem:

\[
J(u_{\delta}) = \max_{\|u_0\|\leq \delta} |T(\tau)|,
\]

where $T(\tau)$ is the evolution of model SSTA and obtained by integrating WF96 model from 0 to $\tau$. By solving this optimization problem, the optimal initial perturbation satisfying the constraint condition $\|u_0\| \leq \delta$, $u_{\delta}$, can be found.

In this paper, the norm $\|u_0\| = \sqrt{T_0^2 + h_0^2}$ is employed to define the constraint condition, where $T_0$ and $h_0$ represent the nondimensional initial SST and thermocline depth anomalies in the eastern tropical Pacific, respectively.

For the time interval $[0, \tau]$ with $\tau$ ranging from 1 to 12 months, we respectively take January, February, ..., December as the initial time and calculate the CNOPs of annual cycle $O$ with constraint condition $\|u_0\| \leq \delta$, $\delta \in [0.06, 0.30]$. It is shown that for the annual cycle, regardless of what the initial time is, there exist CNOP, $u_{\delta}$, and local CNOP, $u_{\delta L}$, at which the objective function $J(u_{\delta})$ attains local maximums in the phase space. These CNOPs and local CNOPs are all on the boundary of the corresponding disc $\|u_0\| \leq \delta$, $\delta \in [0.06, 0.30]$. For $\delta = 0.24$, the CNOP and local CNOP with initial time being January and $\tau = 12$ are respectively $(-0.1373, 0.1968)$ and $(0.1511, -0.1865)$, for $\delta = 0.30$, they are $(-0.0908, 0.2859)$ and $(0.1821, -0.2384)$. In addition, with $\tau$ increasing from 1 to 12, the CNOPs (local CNOPs) are always located in the II (IV)-quadrant and have the robust pattern of negative (positive) SST and positive (negative) thermocline depth anomalies qualitatively. For example, for initial time being January with $\tau = 3, 6, 9, 12$, the CNOPs with $\delta = 0.24$ are respectively $(-0.0099, 0.2398)$, $(-0.03732, 0.2370)$, $(-0.0909, 0.2221)$, and $(-0.1373, 0.1968)$, the corresponding local CNOPs are $(0.0190, -0.2392)$, $(0.0805, -0.2261)$, $(0.1334, -0.1995)$, and $(0.1511, -0.1865)$. [16] To compare CNOP (local CNOP) with LSV, we investigate the distribution of CNOP (local CNOP) and LSV in phase space, $u_{\delta L} = (-0.0218, 0.0206)$ is a LSV for initial time being January and $\tau = 12$ and located in II-quad- rant, which is the fastest growing perturbation of the TLM of WF96 model. To facilitate the discussion, we define two scaled LSVs,

\[
\begin{align*}
\bar{u}'_{\delta L} &= \|u_{\delta L}\|_{\|u_0\|_{\|u_0\|_{\|u_0\|_{\|u_0\|}}}}, \\
\bar{u}_L &= -\\|u_{\delta L}\|_{\|u_0\|_{\|u_0\|_{\|u_0\|}}},
\end{align*}
\]

thus

\[
\|\bar{u}_{\delta L}\| = \|\bar{u}'_{\delta L}\| = \delta, \quad \|\bar{u}_L\| = \|\bar{u}_{\delta L}\| = \delta.
\]

Figure 1 shows the $u_{\delta L}(u_{\delta})$ and $u_{\delta L}'(u_{\delta}')$ with initial time being January for $\tau = 12$, where $A$ and $A'$ correspond to the $u_{\delta L}$ and $u_{\delta L}'$ with $\delta = 0.30$, and $B$ and $B'$ to the $u_{\delta L}$ and $u_{\delta L}'$ with $\delta = 0.30$ respectively. It is easily demonstrated that for the same value of $\delta$, when it is large, for instance, $\delta = 0.24$ or $0.30$, $u_{\delta L}$ and $u_{\delta L}'$ are respectively quite different from the scaled LSVs, $u_{\delta L}$ and $u_{\delta L}'$. The LSVs, with $\delta$ increasing from
0.06 to 0.30, show themselves a beeline, while the CNOPs and local CNOPs shape into a curve respectively. Consequently the differences between CNOP (local CNOP) and LSV become more and more considerable with the increasing $d$.

We further investigate the evolutions of CNOP and LSV. Integrating the nonlinear model and TLM with initial conditions $u_0^d$, we respectively obtain the $T$ components of the linear and nonlinear evolutions of $u_0^d$, denoted by $u_{LN}^d$ and $u_{LL}^d$ respectively. The results demonstrate that for the short time intervals, no matter what the initial time is, there are only trivial differences between the linear and nonlinear evolutions of the $T$ component of CNOP, and between those of LSVs respectively. However, for the long time intervals and large values of $d$, they have considerable differences respectively. Figure 2 plots $u_{LN}^d$ and $u_{LL}^d$ with $d = 0.24$ and $\tau = 12$ for initial time being January. It is shown that for $u_{LN}$, when the time exceeds July ($\tau \geq 7$), its nonlinear evolution begins to depart from the linear counterpart, and with the increasing time this departure becomes more and more serious. The investigation of the difference between the linear and nonlinear evolutions of $u_0^d$ also demonstrates the similar results.

For the model variable $h$, we also investigate the evolutions (Figure 3). It is illustrated that at the end of the optimization time interval, the nonlinear evolution of $h$ of the CNOP is not at maximum, which is even smaller than that of the corresponding LSV. In fact, the derived CNOP from (6) corresponds to the maximum of the objective function $J(u_0)$, which, however, only depicts the maximum evolution of the model SSTA. Therefore it is reasonable that the nonlinear evolution of $h$ of the CNOP may not attain the maximum.

For the model variable $h$, we also investigate the evolutions (Figure 3). It is illustrated that at the end of the optimization time interval, the nonlinear evolution of $h$ of the CNOP is not at maximum, which is even smaller than that of the corresponding LSV. In fact, the derived CNOP from (6) corresponds to the maximum of the objective function $J(u_0)$, which, however, only depicts the maximum evolution of the model SSTA. Therefore it is reasonable that the nonlinear evolution of $h$ of the CNOP may not attain the maximum.

The application of LSV theory to ENSO predictability studies.

Now we compare CNOP with LSV from another perspective, that is, the nonlinear evolutions of LSVs and CNOPs. For the CNOP of annual cycle, $u_0^d$ and $u_{LL}^d$ with initial time being January and $\tau = 12$, their nonlinear evolutions of $T$ component, $u_{LN}^d$ and $u_{LL}^d$ with $d = 0.24$, are also plotted in Figure 2. It is shown that, the $u_{LN}^d$ is significantly larger than $u_{LL}^d$, which indicates that for the initial perturbations in disk $||u_0|| \leq \delta$, CNOP is optimal compared to the LSV under the condition that they are of the same amplitude of norm. All these indicate that for the long time intervals and large initial perturbations, the TLM cannot be used to approximate nonlinear model to study ENSO predictability. Correspondingly the linearity limits
[20] Now we turn to investigate the local CNOP of annual cycle, \( u_{05} \). Let \( u_{LN} \) and \( u_{LL} \) be the nonlinear and linear evolutions of the SSTA of \( u_{05} \), \( u_{LN} \) and \( u_{LL} \) be those of the corresponding scaled LSV, \( u_{05} \). For the different initial time and \( \tau \), we explore the differences between the evolutions of the local CNOP and the scaled LSV with different values of \( \delta \) and find that there exist small differences between \( u_{LN} \) and \( u_{LN} \), and between \( u_{LL} \) and \( u_{LL} \). However, for the linear and nonlinear evolutions of local CNOP, and those of the scaled LSV, when the time intervals are long and the initial perturbations are large, they have considerable differences. Figure 4 describe the differences of the SSTA evolutions of \( u_{05} \) and \( u_{05} \) with \( \delta = 0.24 \) and \( \tau = 12 \) for initial time being January. It is illustrated that when \( \tau \geq 7 \), the linear evolution of \( u_{05} \) gradually departs from its nonlinear one with the increasing time. In addition, at the end of time interval, the difference become aggressively considerable. For the SSTA nonlinear evolutions of \( u_{06} \) and \( u_{06} \), they have not obvious difference at \( \tau = 12 \). The nonlinear evolution of SSTA corresponding to local CNOP is only a little larger than that of the scaled LSV, which can be easily understood by the local optimality of local CNOP. We also investigate the evolutions of the component \( h \) of local CNOP (Figure 5). The results demonstrate that the local CNOP derived from (6) only results in the locally maximum nonlinear evolution of SSTA of the initial perturbations, not in that of \( h \).

[21] To further verify the optimality of CNOP for all the initial perturbations in the corresponding disk \( \|u_0\| \leq \delta \), we perform the following numerical experiments for \( u_{05} \) with \( \delta = 0.24 \). A large samples of initial perturbations in the disk \( \|u_0\| \leq 0.24 \), are chosen to find out the optimal initial pattern. We consider the circumscribed square of the disk and take square mesh of side 0.1. For any mesh points outside the disk, we connect this point with the center of the disk, and take the intersection point of this line with the bound of the disk, then integrate the model from each of these intersection points and the mesh points inside the disk. In calculation, we have tried several different meshes, and have found that there are similar results. So we use 0.1 in the calculation to show the results (Figure 6). It is demonstrated that the evolutions of model SSTA with other initial perturbations in the disk \( \|u_0\| \leq 0.24 \) (including LSV with the same amplitude of norm) are always less than that of SSTA with CNOP, which indicates that CNOP, rather than LSV, is indeed the nonlinear optimal initial perturbation with constraint condition \( \|u_0\| \leq \delta \). For \( u_{06} \), Figure 6...
demonstrates that its SSTA evolution has the opposite tendency and is reliably the local nonlinear optimal initial perturbation for the constraint condition.

[22] The other values of $\delta$, $\delta \in [0.06, 0.30]$, are also chosen to perform the above experiments, which also verified the optimality of CNOPs obtained by nonlinear optimization algorithm. For simplicity, only the case of $\delta = 0.24$ with initial time January and $\tau = 12$ is shown.

[23] In the following, we examine the physical characteristics CNOP and local CNOP bear in ENSO model. From Figure 2, we note that the CNOP of annual cycle, $u_{0N}$ with $\delta = 0.24$, evolves into the positive SSTA nonlinearly, which takes a striking resemblance to the development of El Nino. In fact, it acts as a precursors for El Nino event in WF96 model. Although the corresponding scaled LSV, $u_{0L}$ with $\delta = 0.24$, also develops into an El Nino (Figure 2), the amplitude is considerably smaller than that of $u_{0N}$. The results have shown that LSV is not the fastest growing perturbation in nonlinear WF96 model and cannot be considered to be the optimal precursor for ENSO. The CNOP of annual cycle, $u_{0N}$, due to its maximum evolution of model SSTA at prediction time $\tau$, acts as the conditional optimal initial perturbation of the nonlinear model. In addition, consequently it becomes the initial pattern that evolve into El Nino event most potentially. The intensity of the El Nino events caused by CNOPs depends on the amplitude of initial perturbations. The extensive numerical experiments demonstrate that the larger the CNOP, the stronger the El Nino event. It is readily derived that the sufficiently small CNOP cannot evolve into an El Nino. Therefore the optimal precursor of El Nino is the CNOP of annual cycle with reasonably large amplitude. For the local CNOP of annual cycle, the analysis demonstrates that $u_{0N}$ with moderately large value of $\delta$ is the optimal precursor of La Nina (Figure 4).

[24] In numerical experiments, we also investigate the cases of the initial time being April, July, October. The results also support the above conclusion. For simplicity, the details are not illustrated here.

[25] From the above discussion it is easily demonstrated that the CNOPs, $u_{0N}$, and the local CNOPs, $u_{0L}$, are the initial patterns that evolve into El Nino and La Nina events most probably, respectively. However, it is noticed that due to the property of the initial perturbations $u_{0N}$ and $u_{0L}$ being optimal and locally optimal respectively, the corresponding El Nino event could be stronger than La Nina event under the condition that the two initial perturbations are of the same amplitude of the norm, which is easily derived from the definition of the objective function. To compare them with the situation of LSVs, we perform the following numerical experiments.

[26] Taking the CNOP, $u_{0N}$, and the local CNOP, $u_{0L}$, with $\delta = 0.24$ as initial values, integrating the nonlinear model starting from January respectively, we obtain the evolutions of SSTA corresponding to CNOP and local CNOP, where SSTA evolution of CNOP, $u_{0N}$, is shown in Figure 7. For the T evolution of $u_{0N}$, we multiply it by $-1$ and obtain the evolution of ($-T$), $v_{0N}$, which is also shown in Figure 7. Comparisons between $u_{0N}$ and $v_{0N}$ demonstrate that El Nino is obviously stronger than La Nina on the condition that their initial patterns are of same amplitude, which supports the above view and is quite consistent with the fact of the ocean data analysis. However, when we use the two scaled LSVs, $u_{0L}$ and $u_{0L}$, to explore the intensity of El Nino and La Nina, since they correspond to the same singular value, their initial patterns are of same amplitude, which supports the above view and is quite consistent with the fact of the ocean data analysis.
La Niña. The reason is that the El Niño-La Niña asymmetry is caused by a nonlinear feedback of WF96 model.

5. Data Analysis

[27] The theoretical results obtained in section 4 suggest that initially negative (positive) SST and positive (negative) thermocline depth anomalies over the eastern tropical Pacific will most likely evolve into an El Niño (a La Niña) event. To verify these theoretical results, in the following we analyze available ocean data over the equatorial eastern Pacific (5°S–5°N, 150°–90°W) region to examine the initial patterns of realistic ENSO events.

[28] According to the theoretical model, in the present study we primarily focuses on investigating the evolutions of monthly mean SSTA, and thermocline depth anomaly, which is estimated by using the equatorial 20°C isotherm depth anomaly (h20A) as a surrogate. In this paper, SSTA over Niño-3 region are obtained by the observed SST, and h20A are derived from NCEP reanalysis equatorial h20 data.

[29] During the period of 1980–2002, four major El Niño and three La Niña events occurred, which are illustrated in Figures 9 and 10 respectively. By examining each ENSO event, it is found that before the occurrence of El Niño (La Niña) event, the preferred transition of thermocline depth is about 4 months ahead of SST transition (see the time period marked by rectangles plotted in Figures 9 and 10), that is to say, the positive (negative) h20A always appear earlier than positive (negative) SSTA before the onset of El Niño (La Niña), which agrees qualitatively with the optimal El Niño (La Niña) precursor obtained in this paper, the pattern with negative (positive) SSTA and positive (negative) h20A.

6. Physical Mechanism Responsible for the Growth of the Optimal Precursory Disturbances

[30] It has been shown that for the long time interval and large amplitudes of initial perturbations (value of δ is large), the nonlinear evolutions of SSTA of CNOPs (local CNOPs) are larger (smaller) than their linear counterparts. For the same magnitude of the given norm, the SSTA nonlinear evolutions of CNOPs with large magnitude of norm are substantially larger than those of the scaled LSVs. For the local CNOPs, since they are locally optimal for the constraint condition, their nonlinear evolutions of model SSTA are very close to those of the corresponding scaled LSVs. Consequently, CNOP of annual cycle acts as the optimal precursor of El Niño, and local CNOP serves as that of La Niña event. Further investigation demonstrates that the optimal precursor of El Niño (La Niña), for different initial times, is characterized by negative (positive) SSTA and positive (negative) thermocline depth anomaly. In the following, we will discuss the physical mechanism, which is responsible for the growth of the optimal precursory disturbance.

[31] The ENSO model (5) has two characteristic lines given by \( \frac{dT}{dt} = 0 \) and \( \frac{dh}{dt} = 0 \), which are, respectively,

\[
\begin{align*}
  h &= \frac{\left(a_1 - \sqrt{\frac{a_2}{T}}\right) T}{a_2 + a_3 \sqrt{\frac{a_2}{T}}} = f(T), \quad \text{or} \quad T = f^{-1}(h), \\
  h &= \frac{1}{2} T.
\end{align*}
\]

The two characteristic lines partition the phase plane or the SST-h oscillation cycle into four phases (Figure 11) [Wang and Fang, 1996]. Figure 11 sketches an ENSO cycle in the phase plane. We rewrite the first equation of (5),

\[
\frac{dT}{dt} = a_1 T - a_2 h + \eta(T, h),
\]

where \( \eta(T, h) = \sqrt{\frac{a_2}{T}}(T - a_3 h). \) \( \eta(T, h) \) represents the nonlinear term of the model and emphasizes the nonlinear effect of the temperature advection to SST-h oscillation. Let \( \eta(T, h) = 0 \), and we obtain the straight line \( T - a_3 h = 0 \). To facilitate the discussion, it is also plotted in Figure 11.

[32] It follows from the above sections that the nonlinear evolution of the model SST of CNOP, for the long time interval and large amplitude of initial perturbations, is significantly larger than the corresponding linear counterpart. This suggests that the nonlinearity plays an important role in model ENSO. The explanation is as follows. For the nonlinear evolution of CNOP with annual cycle, when the component \( T \) approaches to be positive and gradually becomes large, the model SSTA, \( T \), will satisfy \( a_3 h < T < f^{-1}(h) \) (\( h > 0 \)) during the period of \( \delta \) in Figure 11, and further makes the nonlinear term \( \eta(T, h) > 0 \). From (7), it is readily derived that the larger the nonlinear term, the larger the positive SSTA. This is a positive feedback between positive SSTA and the nonlinear term. If the nonlinear term is omitted in (7), the \( \frac{dT}{dt} \) is dominated by the linear part. The positive SSTA will become smaller than that with positive nonlinear term. Therefore the nonlinear term increases the instability of El Niño event, which therefore results in the aggressively large nonlinear evolution of CNOP compared to the linear one. As for the nonlinear evolution of local CNOP, which evolves into a La Niña when the amplitude of perturbation is large, the developing negative \( T \) and \( h \) will make \( f^{-1}(h) < T < a_3 h \) (\( h < 0 \)) during the period of \( \delta \) in Figure 11. At this time the nonlinear term \( \eta(T, h) \) is still larger than zero. It can be shown that the positive nonlinear term suppresses the negative SSTA being more negative, which forms a negative feedback between positive nonlinear term and SSTA during La Niña event. As a result, the nonlinear evolution of local CNOP is considerably smaller than its linear evolution, where the nonlinear term \( \eta(T, h) \), increases the stability of La Niña event. Hence if the TLM is used to investigate El Niño (La Niña), it cannot capture the effect of nonlinearity to its evolution. That is to say, the linearized model of WF96 model does not reflect the nonlinear effect of the temperature advection to SST-h oscillation.

[33] The above discussion illustrates that \( \eta(T, h) \) enhances the instability of El Niño events, and suppresses that of La Niña events. This indicates that for the same amplitude of initial perturbations, an El Niño is stronger than a La Niña event. Thus this explains the asymmetry of El Niño and La Niña in WF96 model. This also shed lights on the fact why the two scaled LSVs, \( \tilde{u}_{0L} \) and \( \tilde{u}_{0L} \), in the absence of nonlinearity, evolve respectively into an El Niño and a La Niña event with the same amplitude in the TLM. Therefore CNOP unravel the effect of nonlinearity to its evolutions and is superior over to the LSVs in exploring the optimal precursor for ENSO.
Figure 9. Evolution of anomalous SST (SSTA) and the depth of the equatorial 20°C isotherm (h20A) during four ENSO warm events: (a) 1982/1983; (b) 1986/1987; (c) 1991/1992; (d) 1997/1998. The SSTA, monthly mean averaged over the Nino-3 region (5S–5N, 90–160W), is obtained by the observational SST, and equatorial h20A is derived from NCEP ocean reanalysis data, which is monthly mean averaged along equatorial eastern Pacific (90–160W). The rectangles mark the time period when the pattern with negative SSTA and positive h20A emerges.
Figure 10. Evolution of anomalous SST (SSTA) and the depth of the equatorial 20°C isotherm (h20A) during three ENSO cold events: (a) 1984/1985, (b) 1988/1989, and (c) 1999/2000. The SSTA and equatorial h20A are obtained as in Figure 9. The rectangles mark the time periods when the pattern with positive SSTA and negative h20A arises.
The robust optimal precursors of ENSO have simple physical interpretation. In the model ENSO cycle (Figure 11), $T$ and $h$ evolve into the phase II (IV), in which the variation in $h$ and SST have an opposite tendency. The phase leading of the thermocline displacement to SST variation provides a negative feedback that turns the coupled system from warming to cooling or vice versa. When an ENSO phase turns from warming (cooling) to cooling (warming), the pattern of the positive (negative) SST and negative (positive) thermocline depth anomaly emerges robustly, which acts as an optimal precursor of ENSO.

How do the optimal precursors obtained in this study evolve into ENSO events? To answer this question, we take El Niño event as an example to explain the growth mechanism of the optimal precursors. During the period of a cold event, when the cooling in the eastern tropical Pacific is sufficiently strong, two negative feedback processes act against the cooling. First, the cooling decreases anomalous vertical temperature difference across the mixed layer base, which offsets the contribution of anomalous upwelling to SST cooling. Second, when the cooling is sufficiently strong so that $T < 2h$ (nondimensional) and $\frac{dT}{dh} > 0$ ($b > 0$) (see equation (1)), the thermocline depth in the equatorial eastern Pacific will increase and take the lead in transiting to positive anomaly. Then the sinking thermocline would amplify the temperature of the water upwelled into the mixed layer and suppress cooling, so that the SST over the eastern tropical Pacific region increases gradually. Consequently a positive SST anomaly occur in the eastern equatorial Pacific. This will reduce the zonal SST gradient along the equator and weaken the climatological mean easterlies to the west of the SST anomaly. The weakening of the easterlies reduces equatorward Ekman convergence in the surface layer and suppresses the mean equatorial upwelling, leading to further warming in the eastern Pacific. This positive feedback was first visualized by Bjerknes [1969] and demonstrated by previous coupled stability analysis [Philander et al., 1984; Hirst, 1986; Neelin, 1991]. This instability causes the rapid warming of SST in eastern tropical Pacific, then further results in an El Niño event.

7. Summary and Discussion

The properties of the optimal growth for a theoretical coupled model for ENSO with an annual cycle were explored by the approach of conditional nonlinear optimal perturbation (CNOP). The results demonstrate that the CNOPs of annual cycle are quite different from the linear singular vectors (LSVs) in phase space for the long optimization time intervals and large amplitude of perturbations. In addition, their nonlinear and linear evolutions also remain significant differences. Owing to the presence of nonlinearity of WF96 model, the nonlinear evolutions of CNOPs (local CNOPs) are notably larger (smaller) than their linear counterparts for the large initial perturbations. Further studies find out that the CNOPs (local CNOPs) of annual cycle are of the robust pattern with negative (positive) SST and positive (negative) thermocline depth anomalies qualitatively. These patterns tend to evolve into El Nino (La Nina) event most potentially, which can therefore be regarded as the optimal precursor of El Nino (La Nina). This theoretical result is verified against the 22-year NCEP reanalysis data qualitatively.

The robust optimal precursors of ENSO are explained physically. It is shown that the optimal precursors occur in the ENSO transition phase variation robustly, in which the thermocline depth displacement takes a phase lead to SST variation and provides a negative feedback that turns the coupled system from one state to another state. In addition, the nonlinear behaviors of them reveal the nonlinear effect of temperature advection to ENSO oscillation. Practically, the nonlinear temperature advection enhances the instability of El Nino, and suppresses that of La Nina, which results in the asymmetry of them about the SST normal state.

Li and Mu [2002] also explored the precursor of ENSO by the method of data analysis. Their attentions were mainly paid to the western equatorial Pacific (warm pool). The results of them argued that before the onset of El Nino, the initial positive ocean temperature anomaly First arose in the subsurface layer of warm pool about one year in advance. In this paper, due to the limitation of the model adopted, the interest has to be confined to the equatorial eastern Pacific. As a result, the situation of the warm pool before ENSO event cannot be investigated theoretically. To further understand the ENSO cycle, the more complicated model should be adopted to disclose theoretically the ENSO precursor from the equatorial western Pacific.

In view of the simplicity of the model, the results of this study are qualitatively indicative. The numerical experiments performed here are of exploratory nature. However, we are greatly encouraged by these results. It is expected that for a more realistic model, there will be more significant findings by using the approach of CNOP, which are the works in the future.

Appendix A

The constrained nonlinear optimization problem considered in this paper, after discretization and proper trans-
formation of the objective function, can be written in the form

$$\min_{x \in \mathbb{R}^n} F(x),$$

subject to

$$h(x) \leq 0,$$

where $h = (h_1, h_2, \ldots, h_m)^T$ is a vector of nonlinear functions. It is assumed that at any point $x$ the gradient $\nabla F(x)$ of the objective function and the Jacobian $J(x) = \frac{\partial h}{\partial x} = (\nabla h_1, \ldots, \nabla h_m)$ of constraint function can be computed. Then the CNOP can be obtained by the following steps.

[41] Step 0. Set iteration $i = 0$, a solution guess $x^0$, a Hessian Lagrangian estimate $H^0 = I$, which is the identity matrix, and an initial given value of Lagrange multiplier, $\lambda^0$.

[42] Step 1. Compute $d^i$ by the following quadratic programme (QP) subproblem,

$$\min_{d^i} \left( \nabla F(x^i)^T d^i + \frac{1}{2} d^i^T H^i d^i \right),$$

subject to

$$h(x^i) + \nabla h(x^i)^T d^i \leq 0,$$

where $d^i$ is a direction of descent for the objective function. Then using $d^i$, we determine the Lagrange multiplier $\lambda^{i+1}$ corresponding to the QP subproblem [Barclay et al., 1997].

[43] Step 2. Check convergence. If $x^{i+1}$, $\lambda^{i+1}$ satisfy $\|\nabla F(x^{i+1}, \lambda^{i+1})\| \leq \epsilon$, where $\nabla F = \nabla F + \nabla h\lambda$, and $\epsilon$ is a given positive number to guarantee the convergence, then $x^*$ is the point at which the objective function $F(x)$ is minimal. Otherwise, let $x^{i+1} = x^i + \alpha d^i$, $\alpha \leq 1$, and then go to step 3.

[44] Step 3. Update Hessian Lagrangian. Let $s^i = x^{i+1} - x^i$, and $y^i = \nabla L(x^i, \lambda^i) - \nabla L(x^i, \lambda^i)$. The new Hessian Lagrangian, $H^{i+1}$, can be obtained by calculating

$$H^{i+1} = H^i - \frac{H^i s^i y^i s^i}{s^i H^i s^i} + \frac{y^i y^i}{y^i y^i}.$$

Then go to step 1.

[45] In SQP algorithm, the definition of the QP Hessian Lagrangian $H^i$ is crucial to the success of an SQP solver. In the work of Gill et al. [1997], $H^i$ is a limited-memory quasi-Newton approximation to $G = \nabla^2 F$, the Hessian of the modified Lagrangian. Another possibility is to define $H^i$ as a positive-definite matrix related to a finite difference approximation to $G$ [Barclay et al., 1997]. In this paper, we adopt the former one.

[46] The SQP algorithms have proved reliable and efficient for constraint nonlinear optimization problems. Many ready-made solvers, for example, NLPLSOL [Schittkowski, 1985], NPSOL [Gill et al., 1986], etc., have been developed to solve the general purpose constraint nonlinear optimization problems. Therefore, in this paper, we directly use the above solvers to compute their minima.

[47] Acknowledgments. The authors are grateful to the anonymous reviewers for useful suggestions. This work was jointly supported by the National Natural Science Foundation of China (40223029, 40221503), and KZCX2-208 of the Chinese Academy of Sciences.

References


W. Duan and M. Mu, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China. (duanws@lasg.iap.ac.cn)
B. Wang, Department of Meteorology, School of Ocean and Earth Science and Technology, University of Hawaii, Honolulu, HI 96822-2219, USA.