The Role of Air–Sea Interaction in Controlling the Optimal Perturbations of Low-Frequency Tropical Coupled Ocean–Atmosphere Modes

ANDREW M. MOORE,* JÉRÔME VIALARD,+,& ANTHONY T. WEAVER,# DAVID L. T. ANDERSON,+ RICHARD KLEEMAN,@ AND JOLIE R. JOHNSON*

*Program in Atmospheric and Oceanic Sciences, and Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado
+European Centre for Medium-Range Weather Forecasts, Reading, Berkshire, United Kingdom
#Centre Europeén de Recherche et de Formation Avancée en Calcul Scientifique, Toulouse, France
@Courant Institute of Mathematical Sciences, New York University, New York, New York
&LODYC, Unité Mixte de Recherche CNRS, Université Paris, Paris, France

(Manuscript received 3 December 2001, in final form 2 October 2002)

ABSTRACT

In this paper the structure and dynamics of the optimal perturbations of tropical low-frequency coupled ocean–atmosphere oscillations relevant to El Niño–Southern Oscillation (ENSO) are explored. These optimal perturbations yield information about potential precursors for ENSO events, and about the fundamental dynamical processes that may control perturbation growth and limit the predictability of interannual variability. The present study uses a hierarchy of hybrid coupled models. Each model is configured for the tropical Pacific Ocean and shares a common ocean general circulation model. Three different atmospheric models are used: a statistical model, a dynamical model, and a combination of a dynamical model and boundary layer model. Each coupled model possesses a coupled ocean–atmosphere eigenmode oscillation with a period of the order of several years. The properties of these various eigenmodes and their corresponding adjoint eigenmodes are explored.

The optimal perturbations of each coupled model for two different perturbation growth norms are also examined, and their behavior can be understood in terms of the properties of the aforementioned eigenmode oscillations. It is found that the optimal perturbation spectrum of each coupled model is primarily dominated by one member. The dominant optimal perturbation evolves into the most unstable eigenmode of the system. The structure of the optimal perturbations of each model is found to be controlled by the dynamics of the atmospheric model and air–sea interaction processes. For the coupled model with a statistical atmosphere, the optimal perturbation center of action is spread across the entire tropical Pacific in the form of a dipole. For the coupled models that include deep atmospheric convection, the optimal perturbation center of action is primarily confined to the western Pacific warm pool. In addition, the degree of nonnormality of the eigenmodes is controlled by the atmospheric model dynamics. These findings are in general agreement with the results obtained from intermediate coupled models. In particular, the atmospheric models used here have also been used in intermediate coupled models that have been employed extensively in previous studies of the optimal perturbations of El Niño–Southern Oscillation. Thus, a direct comparison of the optimal perturbation behavior of those intermediate models and the optimal perturbations of the hybrid models used here can be made.

1. Introduction

The need to improve probabilistic forecasts of El Niño–Southern Oscillation (ENSO) has led to a flurry of interest in the optimal perturbations and optimal forcing patterns of the coupled ocean–atmosphere system. The optimal perturbations of a variety of coupled models of ENSO of varying complexity have been discussed in the literature, although there is considerable disagreement between them for a given perturbation growth norm. Some models suggest a pronounced ENSO-like response resulting from perturbations in the eastern tropical Pacific (Blumenthal 1991; Xue et al. 1994, 1997a,b; Chen et al. 1997; Thompson 1998; Fan et al. 2000), while other models yield the same response resulting from perturbations in the west Pacific (Moore and Kleeman 1996, 1997a,b; Eckert 1999). The latter support the popular idea that intraseasonal variability in the tropical atmosphere may act to enhance or even trigger ENSO episodes (Lau and Chan 1985, 1986, 1988; McPhaden 1999). However, the optimal perturbation structure can depend critically on the physics employed in a model and on the number of model degrees of freedom (Moore and Kleeman 2001). We address some of these important issues further in the present study.

Most previous studies of optimal perturbations have focussed on models of intermediate complexity, and as such, are often criticized for their simplicity. In this

Corresponding author address: Dr. Andrew M. Moore, PAOS, University of Colorado, Campus Box 311, Boulder, CO 80309-0311.
E-mail: andy@australis.colorado.edu

paper we attempt to address this latter issue by extending the ideas and findings of earlier studies to more complex coupled models in which one component is an ocean general circulation model (OGCM). We consider coupled models composed of the OGCM coupled to three different atmospheric models. These same three atmospheric models have also been coupled to a 1½-layer ocean model, and the optimal perturbations of the resulting so-called intermediate coupled models have been studied in detail in previous published works. The qualitative agreement between the behavior of the optimal perturbations explored in this study, and those from the intermediate coupled models using the same atmospheric models but a different ocean model suggests that the optimal perturbation structures of coupled models are primarily influenced by atmospheric physics and not ocean dynamics.

Our paper is laid out as follows. In section 2 we describe a hierarchy of coupled models that are used in the present study. For the benefit of readers unfamiliar with the ideas of generalized stability theory, we briefly review in section 3 the relationship between the eigenmodes of a dynamical system and the optimal perturbations. In section 4 we examine the structure of the most unstable coupled ocean–atmosphere eigenmodes of each member of the model hierarchy, and explore the impact of atmospheric physics on the nonnormality of the system. The optimal perturbations of each coupled model are examined in section 5 using various perturbation growth norms. We end with a discussion of our results and conclusions in section 6.

2. A hierarchy of coupled models

Of interest in this study are the optimal perturbations of the primary modes of coupled ocean–atmosphere variability on seasonal-to-interannual timescales, both as potential triggers for low-frequency variability, such as ENSO, and for generating ensemble forecasts. It could be argued that the most complete picture can only be obtained from coupled ocean–atmosphere GCMs (CGCMs). While this may be true, the practicalities of computing the optimal perturbations using a CGCM are prohibitive because of the disparity in timescales between atmospheric instabilities, ocean instabilities, and low-frequency coupled ocean–atmosphere modes. To isolate the optimal perturbations of the low-frequency coupled modes we have used simplified atmospheric models that describe only the low-frequency dynamics thought to be important for ENSO. This approach has been the foundation on which many previous studies of ENSO have been based. The models used here exclude the jets and fronts that yield very rapid perturbation growth at mid- and upper levels of the troposphere on fast timescales (Molteni et al. 1993; Buizza and Palmer 1995; Hartmann et al. 1995). Similar features in the ocean can also support optimal perturbation growth (Farrell and Moore 1992; Moore and Farrell 1993; Moore and Mariano 1999), but a judicious choice of perturbation growth norm can be used to focus attention on the lower frequency coupled modes of variability.

In this work we have used a hierarchy of three hybrid coupled models of ENSO. The hierarchical structure is determined by the complexity of the different atmospheric models used since the OGCM is common to each coupled model. Our choice of models was motivated by Moore and Kleeman (2001) who showed that atmospheric model physics have a significant impact on optimal perturbation structure in intermediate coupled models. The atmospheric models used in this study all emphasize a different aspect of air–sea interaction in the Tropics, and have been used in previous optimal perturbation studies where they were coupled to a 1½-layer ocean model. It is the impact of these different atmospheric processes on the optimal perturbations that we have explored here. Since all of the component models are described in detail elsewhere, only a short description is given here.

a. The ocean model

The primitive equation OGCM common to each model is Océan Parallélisé (OPA) 8.1 (Madec et al. 1998) configured for the tropical Pacific Ocean between 30°N and 30°S, and 120°E and 75°W, (see Vialard and Delecluse 1998a,b; Vialard et al. 2001). The model horizontal resolution is 1° zonal, and increases in the meridional direction from 0.5° between 5°N and 5°S, to 2° at 30°N and 30°S. There are 25 levels in the vertical with 10-m thickness in the upper 150 m, and realistic bathymetry and coastal geometry are included.

Prior to coupling with each atmospheric model, the OGCM was first spun up using climatological monthly mean, Florida State University (FSU) winds for 30 years. The model was also forced by a surface heat flux $Q = Q_h + \epsilon(T - T_r)$ where Q_h is the climatological seasonal cycle derived from the European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis (ERA), T is the model SST, T_r is the climatological mean observed seasonal cycle of SST from Reynolds and Smith (1994), and $\epsilon = -40$ W m^{-2} K^{-1}. A similar annual mean net freshwater flux forcing Q_s was used, composed of ERA net evaporation minus precipitation ($E - P$) and a relaxation to the annual mean sea surface salinity of Levitus (1982). At the end of year 30 there are no significant long-term trends in the upper ocean circulation within the 20°N–20°S band. Following this, the OGCM was run for a further 40 yr using the actual monthly mean FSU winds for the period 1960–99, and with Q and Q_s as above.

b. Coupled model 1

''Model 1'' consists of the OGCM coupled to a statistical model of the atmosphere. The latter was constructed from a singular value decomposition (SVD)

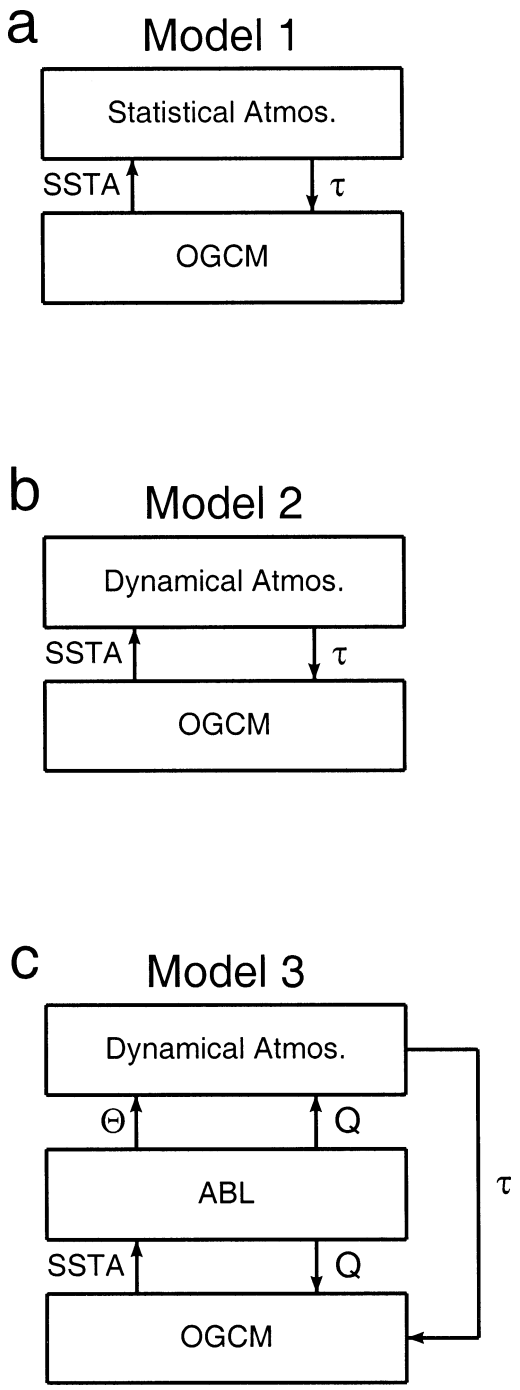

FIG. 1. A schematic of each hybrid coupled model: τ is the surface wind stress anomaly, Θ is the ABL potential air temperature, and Q is the anomalous surface heat flux.

analysis (Bretherton et al. 1992) of the monthly mean SST anomalies of the OGCM and the monthly mean FSU wind stress anomalies from the last 40 yr of the OGCM spinup. Similar atmospheric models have been used by others (e.g., Latif and Villwock 1990; Latif and Flügel 1991; Balmaseda et al. 1994; Syu et al. 1995; Macias et al. 1996; Kleeman et al. 1999).

In model 1 the OGCM was forced by the sum of the FSU monthly mean climatological winds and the wind anomalies computed from the statistical atmospheric model in response to OGCM SST anomalies as illustrated schematically in Fig. 1a. The OGCM SST anomalies were computed relative to the 1967–99 seasonal cycle of the OGCM spinup.

c. Coupled model 2

''Model 2'' consists of the OGCM coupled to the Kleeman (1991) two-level steady-state model of the atmosphere, an excellent approximation in the Tropics (Webster 1972; Gill 1980). The Kleeman model is a global anomaly model, and computes wind anomalies about the observed seasonal cycle. The atmosphere is heated by Newtonian cooling/relaxation to the SST anomaly (which mimics the effects of sensible heat exchange, surface radiation and shallow convection, processes that are not explicitly represented in the model), and by latent heating due to deep penetrative convection via a simple moist static energy dependent convection scheme. In model 2, OGCM SST anomalies were computed as in model 1 and passed to the atmospheric model as illustrated schematically in Fig. 1b. The OGCM was forced by the sum of the associated wind anomalies and the FSU monthly mean climatological winds.

d. Coupled model 3

''Model 3'' is composed of the OGCM coupled to the Kleeman atmospheric model via a model of the atmospheric boundary layer (ABL) developed by Kleeman and Power (1995; Fig. 1c). The ABL is a global anomaly model, and includes the effects of advection by the mean atmospheric circulation, turbulent heat transfer at the surface, and dissipation. The mean atmospheric circulation is prescribed from climatological observations, and the surface heat flux is computed using the normal bulk formula with stability-dependent transfer coefficients [see Kleeman and Power (1995) for complete details].

In model 3, the OGCM SST anomalies were computed as in models 1 and 2. The resulting ABL temperature anomalies were computed relative to an ABL model seasonal cycle obtained by running the ABL for 30 yr using observed monthly mean climatological surface temperatures. In model 3, the net turbulent heat flux anomaly from the ABL was added to the climatological ERA fluxes described in section 2a and used to force the OGCM. The ABL temperature and specific humidity anomalies were used in the latent heat calculations for deep convection in the Kleeman atmospheric model. The wind stresses used to force the OGCM were computed as in model 2.

3. Eigenmodes and optimal perturbations

Before describing the model experiments and results, it is useful to briefly review some important ideas pertaining to optimal perturbations. Readers familiar with this material may proceed directly to section 4.

In sequel we will represent each coupled model symbolically as

$$\frac{\partial \mathbf{S}}{\partial t} = N(\mathbf{S}), \tag{1}$$

where $\mathbf{S}$ represents the state vector of the coupled system (i.e., the vector of model prognostic variables), and N collectively represents the operators that govern the time evolution of $\mathbf{S}$, some of which are nonlinear.

Our interest lies in the evolution of small perturbations $\mathbf{s}$ described by the tangent linear form of (1), namely $\partial \mathbf{s}/\partial t = \mathbf{A}\mathbf{s}$ where $\mathbf{A} = \partial N(\mathbf{S})/\partial \mathbf{S}|_{\mathbf{S}_0}$, $\mathbf{S}_0$ being a solution of (1) that is of particular interest. The vector $\mathbf{S}_0$ may be the actual time varying solution of (1), or it may be taken to represent the time mean trajectory. The latter is not strictly a solution of (1), but such approximations have been discussed and used extensively in the meteorological literature (e.g., Simmons and Hoskins 1976; Frederiksen 1982). In the present study, we will adopt this latter approach and concern ourselves with the case where $\mathbf{S}_0$ does not vary in time. In this case the tangent linear equation is autonomous (i.e., $\mathbf{A}$ is time independent).

The vector $\mathbf{S}_0$ will be referred to as the ''basic state'' and will represent the model annual mean conditions of the coupled system, which were computed as the annual average of each model component forced separately with climatological observed boundary conditions. The ocean basic state is therefore identical in all models. The basic-state SST and a vertical section of temperature along the equator are shown in Fig. 2 for future reference. A time-independent basic state was used for two reasons: (i) to eliminate the influence that differences in the ocean basic state of each coupled model, and its time evolution, may have on their eigenmode structures, and (ii) to simplify the analysis and dynamical interpretation of the resulting autonomous system. The eigenmodes and optimal perturbations of time-evolving basic states will be the subject of a future study.

Traditional linear stability analyses often focus on the eigenmodes of the tangent linear operator $\mathbf{A}$. The conventional view is that if $\mathbf{A}$ possesses at least one growing eigenmode, then the tangent linear dynamical system is considered to be asymptotically unstable (see Pedlosky 1979, his chapter 7). For autonomous systems, the eigenmodes of $\mathbf{A}$ evolve exponentially in time. However, it has long been known that this view of the stability of a dynamical system depends on the properties of $\mathbf{A}$. If $\mathbf{A}$ is normal (i.e., $[\mathbf{A}, \mathbf{A}^\dagger] \equiv \mathbf{A}\mathbf{A}^\dagger - \mathbf{A}^\dagger\mathbf{A} = 0$ where $\mathbf{A}^\dagger$ is the matrix adjoint of $\mathbf{A}$—not to be confused with the adjoint matrix) its eigenmodes are orthogonal, and no perturbation can grow faster than the most unstable eigenmode. On the other hand, if $\mathbf{A}$ is nonnormal ($[\mathbf{A}, \mathbf{A}^\dagger] \neq 0$) its eigenmodes are nonorthogonal. In this case the eigenmodes can linearly interfere and yield perturbations that can grow much more rapidly than the most unstable eigenmode (so-called super exponential growth). In fact perturbation growth is possible in nonnormal systems even when all of the eigenmodes are asymptotically stable (see Farrell and Ioannou 1996, for an excellent discussion of these ideas).

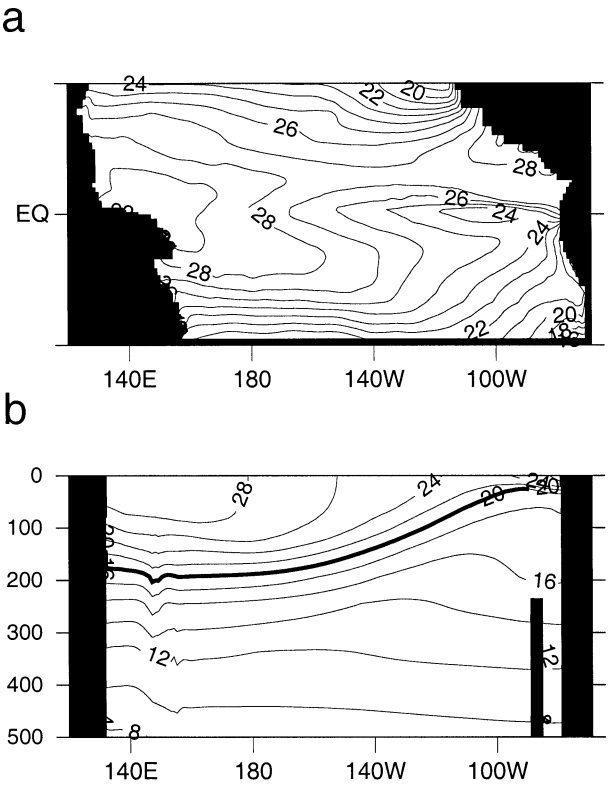

FIG. 2. (a) The ocean basic-state SST, and (b) a vertical section of the upper-ocean temperature used for the eigenmode and optimal perturbation calculations. The contour interval is 2°C.

The linear superposition of eigenmodes that yields the fastest growing perturbation with respect to a given norm and time interval is called an ''optimal perturbation.'' As a measure of perturbation amplitude, consider the norm $\mathcal{N}(t) = \mathbf{s}^{\mathrm{T}}(t)\mathbf{X}\mathbf{s}(t)$, where $\mathbf{X}$ is symmetric positive definite and defines the physical norm of interest (e.g., energy, enstrophy, etc.). A measure of perturbation growth or decay is the factor by which $\mathcal{N}$ changes over the time interval τ, which is given by $\mu = \mathcal{N}(\tau)/\mathcal{N}(0) = \mathbf{s}^{\mathrm{T}}(0)\mathbf{R}^{\mathrm{T}}(\tau)\mathbf{X}\mathbf{R}(\tau)\mathbf{s}(0)/\mathbf{s}^{\mathrm{T}}(0)\mathbf{X}\mathbf{s}(0)$, where $\mathbf{R}(\tau)$ is the propagator of the tangent linear equation, so that $\mathbf{s}(\tau) = \mathbf{R}(\tau)\mathbf{s}(0)$. The optimal perturbation is the perturbation that yields the largest value of μ subject to the constraint $\mathcal{N}(0) = 1$. According to the Rayleigh–Ritz method, the optimal perturbation is the eigenmode of $\mathbf{R}^{\mathrm{T}}(\tau)\mathbf{X}\mathbf{R}(\tau)$ with largest eigenvalue. If $\mathbf{X}$ is factorized as $\mathbf{X} = \mathbf{G}\mathbf{G}^{\mathrm{T}}$, the optimal perturbation is simply the fastest growing singular vector of $\mathbf{G}^{\mathrm{T}}\mathbf{R}(\tau)$.

For nonnormal systems, the eigenmodes of $\mathbf{A}$ and the singular vectors of $\mathbf{G}^{\mathrm{T}}\mathbf{R}(\tau)$ are different perturbations, the optimal perturbations being norm dependent, while the eigenmodes are not. The optimal perturbations are of general interest because they identify disturbances that grow most rapidly with respect to a given norm, and draw attention to dynamically interesting parts of the flow. Optimal perturbations have been studied extensively in meteorology and numerical weather prediction to understand the dynamics of rapid perturbation growth and atmospheric predictability. It is our belief that the optimal perturbations of the coupled ocean–atmosphere system will prove equally valuable for understanding ENSO dynamics and predictability.

As noted, the optimal perturbations represent a linear superposition of the eigenmodes of $\mathbf{A}$, and the projection of an optimal perturbation onto any particular eigenmode can be found by appealing to the bi-orthogonality between the eigenmodes of $\mathbf{A}$ and the corresponding eigenmodes of its adjoint. The adjoint operator $\mathbf{A}^{\dagger}$ is defined by $\langle \mathbf{v}, \mathbf{Au}\rangle = \langle \mathbf{u}, \mathbf{A}^{\dagger}\mathbf{v}\rangle$, where $\langle \mathbf{v}, \mathbf{u}\rangle$ denotes the inner product $\mathbf{v}^{\mathrm{T}}\mathbf{Xu}$ of two vectors $\mathbf{u}$ and $\mathbf{v}$ belonging to the same space (Courant and Hilbert 1953). Clearly $\mathbf{A}^{\dagger}$ and its eigenmodes (hereafter referred to as the adjoint eigenmodes) are norm dependent. An expression for $\mathbf{A}^{\dagger}$ can be found by noting that $\mathbf{X} = \mathbf{X}^{\mathrm{T}}$ and $\langle \mathbf{v}, \mathbf{Au}\rangle = \mathbf{v}^{\mathrm{T}}\mathbf{XAu} = \mathbf{u}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{v} = \mathbf{u}^{\mathrm{T}}\mathbf{X}(\mathbf{X}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{X})\mathbf{v} = \langle \mathbf{u}, (\mathbf{X}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{X})\mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{A}^{\dagger}\mathbf{v}\rangle$. Thus, the adjoint $\mathbf{A}^{\dagger}$ of $\mathbf{A}$ with respect to the chosen norm is $\mathbf{X}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{X}$. The adjoint of $\mathbf{A}$ with respect to the L2 norm ($\mathbf{X} = \mathbf{I}$) is therefore $\mathbf{A}^{\mathrm{T}}$. If $\mathbf{r}$ is an eigenvector of $\mathbf{A}^{\mathrm{T}}$, then $\mathbf{X}^{-1}\mathbf{r}$ is an eigenvector of the adjoint $\mathbf{X}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{X}$. In sequel the adjoint eigenmodes are defined relative to the L2 norm, so $\mathbf{A}^{\dagger} = \mathbf{A}^{\mathrm{T}}$.

The eigenvalue spectra of $\mathbf{A}$ and $\mathbf{A}^{\mathrm{T}}$ are identical, and eigenvalues occur in complex conjugate pairs. We will denote by $\{\sigma_n, \hat{\mathbf{s}}_n\}$ and $\{\lambda_n, \hat{\mathbf{r}}_n\}$ the {eigenvalue, eigenvector} sets of $\mathbf{A}$ and $\mathbf{A}^{\mathrm{T}}$ respectively, where $\lambda_n = \sigma_n^*$, and * denotes the complex conjugate. It is easy to show that

$$\hat{\mathbf{s}}_n^{\mathrm{H}}\hat{\mathbf{r}}_m(\sigma_n - \lambda_m^*) = \hat{\mathbf{s}}_n^{\mathrm{H}}\hat{\mathbf{r}}_m(\sigma_n - \sigma_m) = 0, \qquad (2)$$

where superscript H denotes the Hermitian transpose. For $n = m$, we have $\sigma_n = \sigma_m$, while for $n \neq m$, we require $\hat{\mathbf{s}}_m^{\mathrm{H}}\hat{\mathbf{r}}_n = 0$. Equation (2) is an expression of the bi-orthogonality between the eigenmodes of $\mathbf{A}$ and $\mathbf{A}^{\mathrm{T}}$.

To understand the physical relationship between the eigenmodes and adjoint eigenmodes, consider an initial value problem in which $\mathbf{s}(0)$ is composed of a linear superposition of the eigenmodes $\hat{\mathbf{s}}_n$, namely $\mathbf{s}(0) = \sum_{n=1}^{N} a_n\hat{\mathbf{s}}_n = \mathbf{Ea}$, where $\mathbf{E}$ is the matrix whose columns are the eigenvectors $\hat{\mathbf{s}}_n$, and $\mathbf{a}$ is the vector of amplitudes a_n. Furthermore, assume that the perturbation has unit norm, $\mathcal{N}(0) = \mathbf{s}(0)^{T}\mathbf{Xs}(0) = 1$. According to (2), the amplitude a_k of the kth eigenmode is $a_k = \hat{\mathbf{r}}_k^{\mathrm{H}}\mathbf{s}(0)/(\hat{\mathbf{r}}_k^{\mathrm{H}}\hat{\mathbf{s}}_k)$. The unit norm perturbation $\mathbf{s}(0)$ that maximizes the amplitude a_k of the kth eigenmode can be found by maximizing the Lagrange function $\mathcal{L}$ given by

$$\begin{aligned}\mathcal{L} &= a_k^* a_k + \chi[\mathbf{s}(0)^{\mathrm{H}}\mathbf{Xs}(0) - 1] \\ &= a_k^* a_k + \chi(\mathbf{a}^{\mathrm{H}}\mathbf{E}^{\mathrm{H}}\mathbf{XEa} - 1), \qquad (3)\end{aligned}$$

where χ is an unknown Lagrange multiplier. At the extrema of $\mathcal{L}$, we have $\partial\mathcal{L}/\partial\mathbf{a} = 0$ and $\partial\mathcal{L}/\partial\chi = 0$, which yields

$$\chi\mathbf{E}^{\mathrm{H}}\mathbf{X}(\mathbf{Ea}) = -a_k\boldsymbol{\epsilon}_k, \qquad (4)$$

where $\boldsymbol{\epsilon}_k$ is the unit vector with all elements zero except the kth element. Equation (4) shows that all elements of the vector on the right-hand side are zero except that corresponding to a_k. The left-hand side of (4) implies, then, that $\mathbf{XEa} = \mathbf{Xs}(0)$ must be orthogonal to all the columns of $\mathbf{E}$ (i.e., rows of $\mathbf{E}^{\mathrm{H}}$) except for column k. According to (2) this is a property of the adjoint eigenmodes, so $\mathbf{XEa} = \hat{\mathbf{r}}_k$, and thus $\mathbf{s}(0) = \mathbf{X}^{-1}\hat{\mathbf{r}}_k$, showing that a linear transformation of the kth adjoint eigenmode is the optimal excitation for the kth eigenmode. The coefficient $a_k = \hat{\mathbf{r}}_k^{\mathrm{H}}\mathbf{X}^{-1}\hat{\mathbf{r}}_k/(\hat{\mathbf{r}}_k^{\mathrm{H}}\hat{\mathbf{s}}_k)$ shows that the amplitude of the kth eigenmode depends on the projection of $\hat{\mathbf{s}}_k$ on $\hat{\mathbf{r}}_k$. Therefore, the more dissimilar the eigenmode and adjoint eigenmode are to one another, the larger a_k will be.

The ability of optimal perturbations to undergo super exponential growth depends on the nonnormality of the system. A useful quantitative measure of the degree of orthogonal projection (i.e., nonnormality) of a particular eigenmode $\mathbf{e}$ on the remaining eigenmodes of $\mathbf{A}$ is given by $\nu = |\mathbf{G}^{\mathrm{T}}\mathbf{b}\|\mathbf{G}^{\mathrm{T}}\mathbf{e}|/(\mathbf{b}^{\mathrm{T}}\mathbf{GG}^{\mathrm{T}}\mathbf{e})$ (Farrell and Ioannou 1999), where $\mathbf{b}$ is the corresponding eigenmode of $\mathbf{A}^{\dagger}$. The denominator of ν is simply $\mathbf{r}^{\mathrm{H}}\mathbf{e}$ where $\mathbf{r}$ is the eigenmode of $\mathbf{A}^{\mathrm{T}}$. As $\mathbf{r}$ and $\mathbf{e}$ become increasingly dissimilar, the denominator decreases. Therefore large values of ν are indicative of a high degree of nonnormality for a given eigenmode.

An important distinction between the optimal perturbations and adjoint eigenmodes is that an optimal perturbation will maximize the growth of the chosen norm without regard for its projection on the eigenmodes, while an adjoint eigenmode simultaneously maximises the growth of the chosen norm and the projection of the perturbation on the corresponding eigenmode. In the following, we will find that the most unstable, or least damped eigenmode of each coupled model emerges as the optimal perturbations evolve. However, there is no formal requirement for this to be the case. In some dynamical systems, it is the most nonnormal eigenmode that emerges (Aiken et al. 2002; Aiken et al. 2002, manuscript submitted to *Dyn. Atmos. Oceans*). In all cases considered here, the adjoint eigenmode is found to be the optimal excitation of the corresponding eigenmode, as expected, and in conjunction with the eigenmode provides a useful quantitative measure of the nonnormality of the system as described above. It is primarily for this latter reason that we study the adjoint eigenmodes of each model.

As the system becomes more nonnormal, the structures of the optimal perturbations and adjoint eigen-

TABLE 1. A summary of the characteristics of the most unstable eigenmodes and optimal perturbations of model 1 and model 2 for various values of the coupling strength γ. For each model the following are given: the eigenmode period and *e*-folding growth time in years, the value of $|\nu|$ defined in section 3 using the $\mathcal{T}$ norm, the growth factor μ of the fastest growing optimal perturbation for the $\mathcal{T}$ norm for an optimal growth time of $\tau = 6$ months, and the energy growth factor $e^{2\sigma\tau}$ of a perturbation with structure of the most unstable eigenmode, where σ is the eigenmode growth rate. Blank entries indicate values that were not computed.

	Model 1					Model 2				
γ	Period (yr)	*e*-folding time (yr)	$\lvert\nu\rvert$	μ	$e^{2\sigma\tau}$	Period (yr)	*e*-folding time (yr)	$\lvert\nu\rvert$	μ	$e^{2\sigma\tau}$
0.5									14	
0.8	3.6	0.7	24	127	4.6					
0.825						6.0	50	78	124	1.0
0.85						6.3	8.3	83		1.1
0.9						7.3	3.1	83		1.4
0.93						8.4	2.3	83		1.5
1.0	∞	0.3	14	618	28.0	23.8	1.4	111	366	2.0

modes become increasingly dissimilar to the eigenmodes they excite. For the climate system, models suggest that variability due to El Niño may be strongly related to a single coupled ocean–atmosphere eigenmode of the system. The optimal perturbations and adjoint eigenmodes therefore yield valuable information about how these eigenmodes are best excited either by isolated perturbations or by stochastic forcing (e.g., the Madden–Julian oscillation, westerly wind bursts, etc). Therefore, the dissimilarity between the structure of the ENSO oscillation and its optimal excitation is of fundamental importance from the point of view of both dynamics and prediction. Understanding the physical processes responsible for this behavior is a primary motivation for the present study.

4. Eigenmodes and nonnormality

a. Eigenmode structure

The eigenmodes of each tangent linear coupled model were computed iteratively using the tangent linear version of each coupled model and the Arnoldi package (ARPACK) library of Lehoucq et al. (1997). Details of the tangent linear components of each model are given in Weaver et al. (2003, hereafter WVA), Vialard et al. (2003, hereafter VWD) and Moore and Kleeman (1996), and the computational procedures are described in appendix B of Moore et al. (2002).

The most unstable eigenmode of each coupled model takes the form of a low-frequency coupled ocean–atmosphere oscillation with a period and exponential growth rate that depends on the strength of the coupling between the atmosphere and ocean. In each coupled model the wind stress anomalies (and heat flux anomalies in model 3) used to force the OGCM were multiplied by a factor γ, referred to as the ''coupling strength.'' Table 1 shows how the period and *e*-folding time of the most unstable eigenmode of models 1 and 2 vary with γ, and indicates that the period (*e*-folding time) increases (decreases) with increasing (decreasing) γ. No eigenmode solutions could be found using ARPACK for any of the models for values of γ below the primary bifurcation point (the value of γ below which the model ceases to produce self-sustaining oscillations). For model 2 this occurs around $\gamma = 0.8$. Model 3 exhibits qualitatively similar behavior to models 1 and 2, and the trends in period and *e*-folding time found in all the models generally agree with the behavior of the coupled ocean–atmosphere modes discussed by Jin and Neelin (1993a,b) and Neelin and Jin (1993). The following values of γ were used in all remaining calculations: model 1, $\gamma = 0.8$; model 2, $\gamma = 0.825$; model 3, $\gamma = 1.0$. These values of γ yield unstable eigenmodes with similar oscillation frequencies in each model, and as close as we could find to the observed frequency range of ENSO. The eigenmode of model 2 is the lowest-frequency unstable mode that could be identified in the present study. We will consider primarily unstable systems here because only in this case can we unambiguously identify the gravest eigenmode using ARPACK. However, we note that all of the results and conclusions presented here equally apply to stable systems.

The real and imaginary components of SST for the most unstable eigenmode of each coupled model are shown in Fig. 3. These fields represent the SST structure of the eigenmodes separated by one-quarter of a period. The associated period and exponential growth rate of each eigenmode is indicated, and the basic-state SST is shown in Figs. 2a for reference. The SST structure of the model 3 eigenmode is very similar to that of model 2 and is not shown. Instead, the ABL potential temperature anomaly θ of the model 3 eigenmode is shown in Figs. 3e,f. For each model, the ocean temperature structure of each eigenmode has a significant amplitude only down to the depth of the thermocline (not shown).

The eigenmode of model 1 (Figs. 3a,b) takes the form of an oscillation with a period of 3.61 yr. The eigenmode is ''ENSO-like'' in SST and takes the form of a standing wave in the east Pacific, which is more clearly illustrated

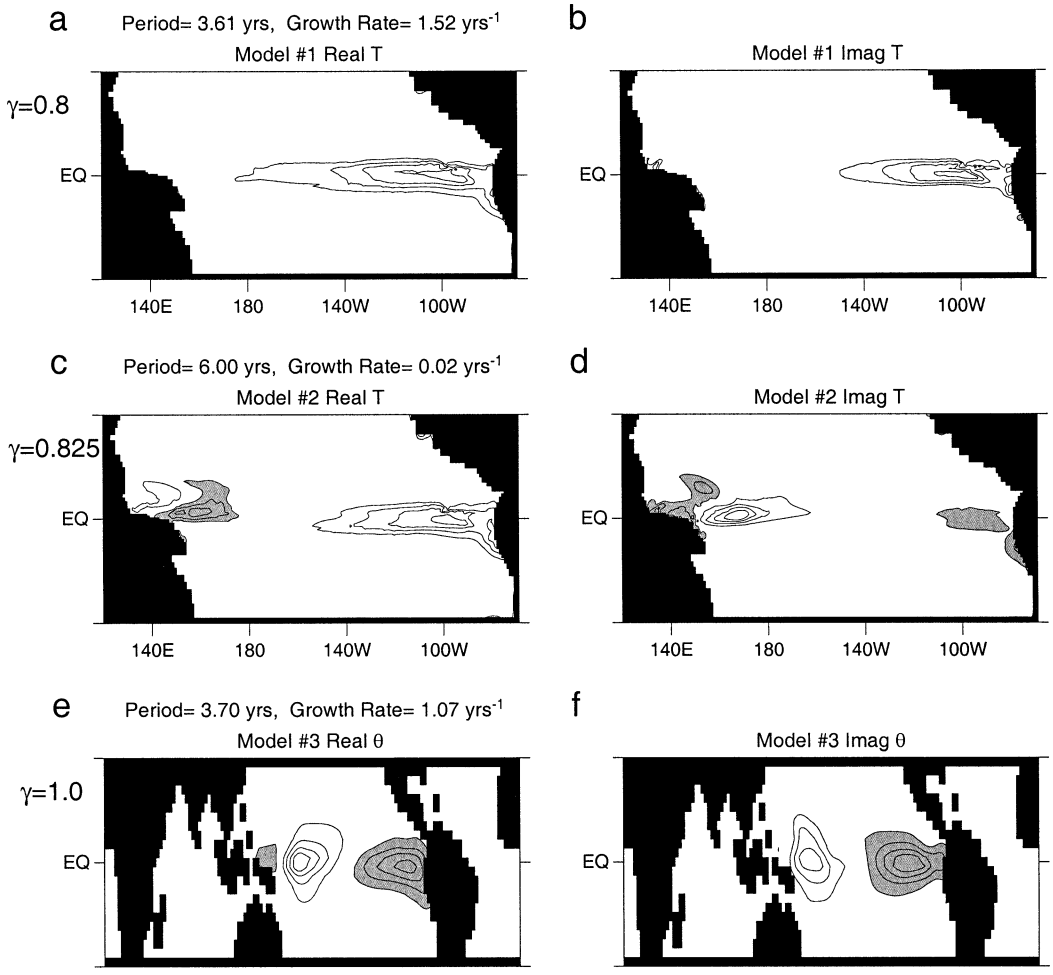

FIG. 3. The real and imaginary components of SST for the most unstable eigenmode of (a), (b) model 1, and (c), (d) model 2. (e), (f) The ABL temperature anomaly of model 3 is shown for the entire ABL model domain. The period and growth rate of the eigenmode is also indicated. Shaded and unshaded regions indicate perturbations of the opposite sign. The contour interval is arbitrary.

in Figs. 4a,b which shows a Hovmöller diagram of SST and zonal wind stress anomalies averaged between 2°N and 2°S.

The eigenmode of model 2 (Figs. 3c,d) takes the form of an oscillation with a period of 6 yr with a predominantly dipole structure in SST along the equator. The west Pacific node of the dipole over the warm pool is associated with deep convection in the Kleeman atmospheric model. This eigenmode takes the form of a standing wave as shown in Figs. 4c,d.

The ABL temperature fields for the model 3 eigenmode are shown in Figs. 3e,f, and reflect the SST anomalies. The eigenmode of model 3 takes the form of an oscillation with a period of 3.7 yr as shown in Figs. 4e,f.

The SST structure of the model 1 eigenmode is qualitatively similar to the first EOF of the observed SST (see Palmer and Anderson 1994, their Fig. 8), while the SST structures of the eigenmodes of models 2 and 3 are reminiscent of the second EOF. Thus, each coupled model captures some aspects of the observed modes of SST variability, and describes a subset of the atmospheric dynamics thought to be important for air–sea interaction in the tropical Pacific on seasonal to interannual timescales. The perturbation dynamics of each model may therefore provide valuable information about the potential influence of each atmospheric process on the nonnormality of the coupled system, and the potential impact on perturbation growth in nature and in complex CGCMs.

b. Adjoint eigenmode structure

The structure and properties of optimal perturbations depend on the degree of nonnormality of the eigenmodes of **A** (Farrell 1982, 1984, 1985, 1988a,b, 1989a,b). If **A** is normal, its eigenmodes are orthogonal and the optimal perturbations coincide with the eigenmodes. On the other hand, if **A** is nonnormal, which is usually the case, the eigenmodes are nonorthogonal. As

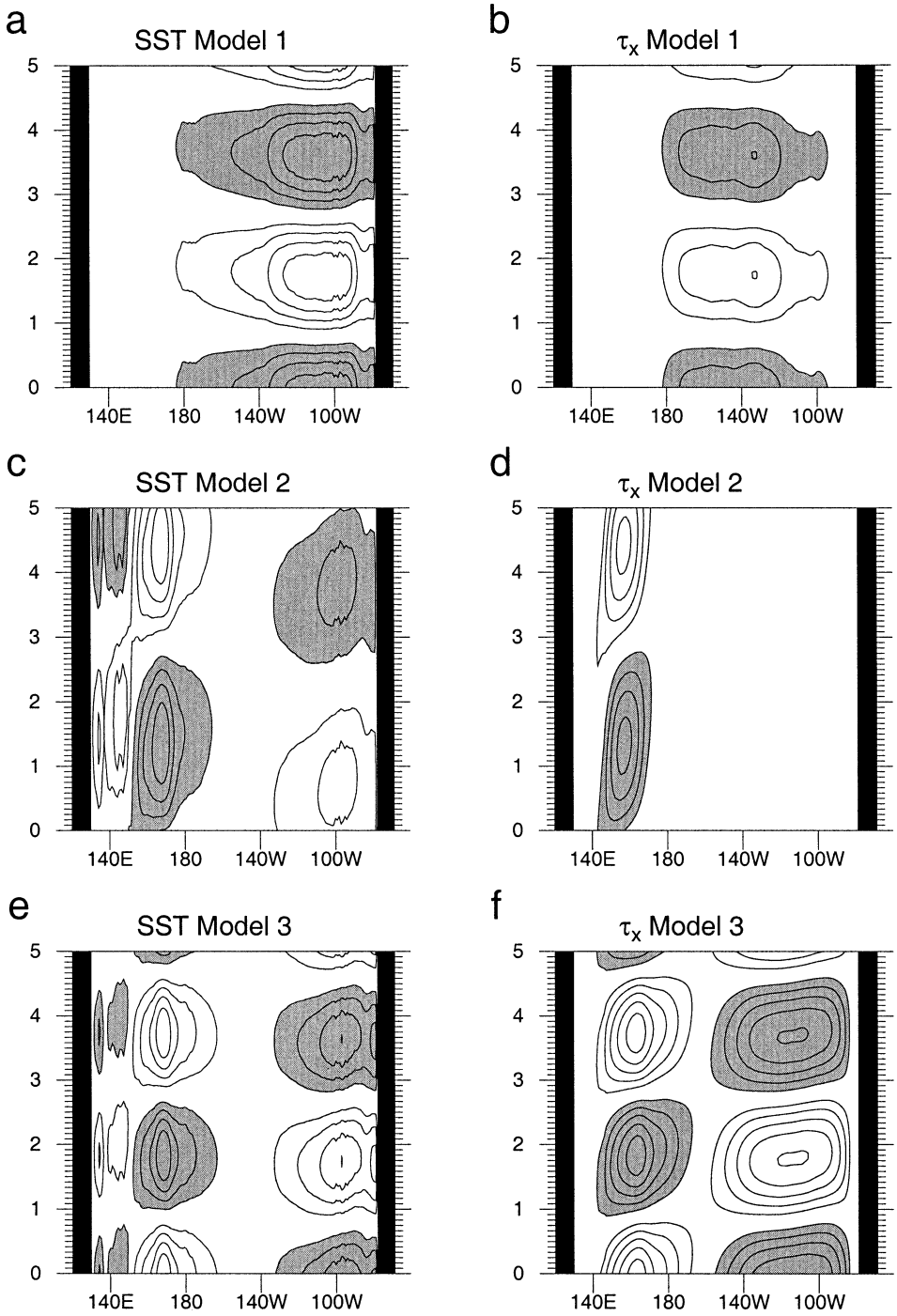

FIG. 4. Hovmöller diagrams of SST and zonal wind stress τ_x averaged between 2°N and 2°S for the most unstable eigenmode of (a), (b) model 1, (c), (d) model 2, and (e), (f) model 3. Shaded and unshaded regions indicate perturbations of opposite sign. The contour interval is arbitrary. In each case the exponential growth of the eigenmodes has been suppressed.

the degree of projection (i.e., nonnormality) of the eigenmodes on each other increases, it becomes easier to conceal a particular eigenmode with large amplitude by way of a linear superposition of the other members of the eigenmode spectrum. The structure of the resulting perturbation can be very different to that of the concealed eigenmode. As the perturbation evolves, dispersion of the eigenmodes occurs, and the concealed eigenmode emerges with large amplitude. The degree of nonorthogonality is therefore an important factor that

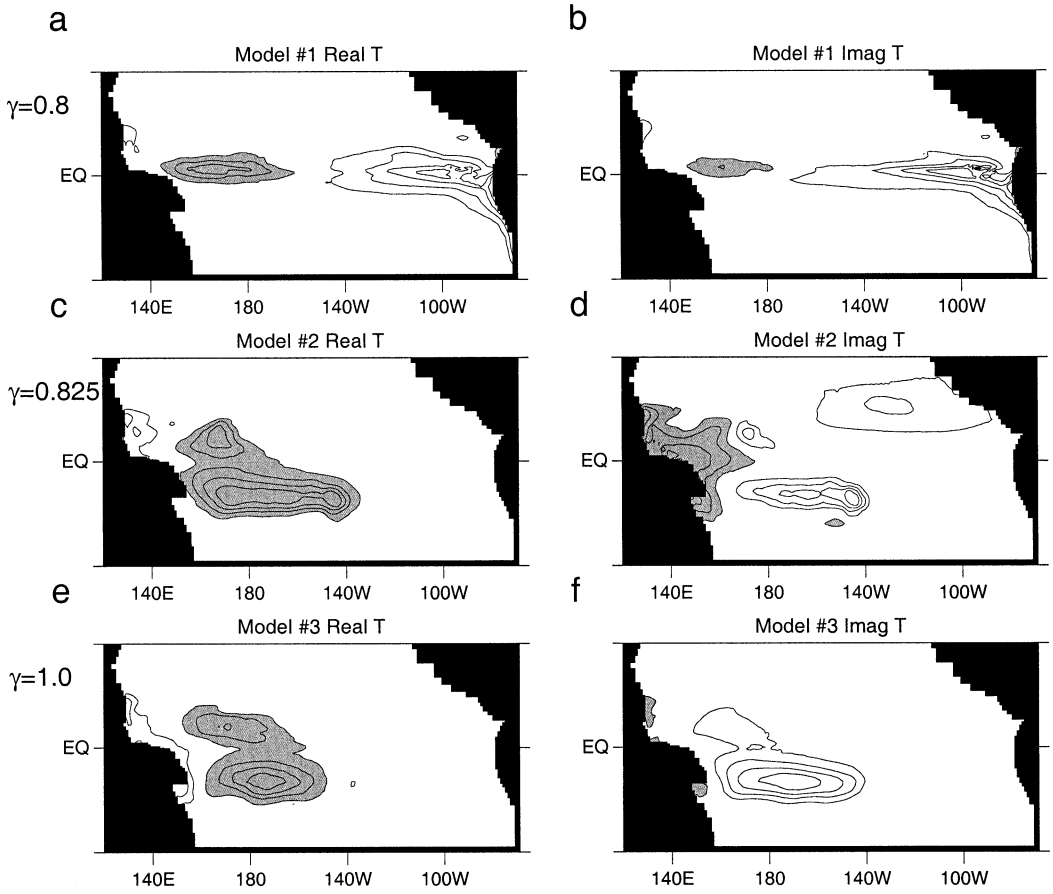

FIG. 5. The real and imaginary components of SST of the most unstable adjoint eigenmode for (a), (b) model 1, (c), (d) model 2, and (e), (f) model 3. Shaded and unshaded regions indicate perturbations of opposite sign. The contour interval is arbitrary.

determines the structure and growth of the resulting optimal perturbations (Farrell and Ioannou 1996).

Following section 3, we will denote the most unstable eigenmode of each coupled model as **e**, and the corresponding adjoint eigenmode as **b**. The most unstable adjoint eigenmodes **b** of each coupled model were computed using the adjoint versions of each tangent linear coupled model (WVA; Moore and Kleeman 1996). The SST structures of the most unstable adjoint eigenmodes for each coupled model are shown in Fig. 5.

For model 1, the adjoint eigenmode takes the form of a dipole spanning the entire equatorial Pacific (Figs. 5a,b). A Hovmöller diagram of the equatorial adjoint eigenmode SST is shown in Fig. 6a, confirming that the period is identical to that of the corresponding eigenmode (cf. Fig. 4a). The values of $|\nu|$ for the model 1 eigenmode for two norms relevant to later sections of the paper are shown in Table 2.

In the case of model 2, Figs. 5c,d shows that the adjoint eigenmode is confined mainly to the west Pacific warm pool. Figure 6b shows a Hovmöller plot of the adjoint eigenmode equatorial SST. The values of $|\nu|$ for the eigenmode of model 2 are shown in Table 2, and a comparison with those of model 1 reveals that the most unstable eigenmode of model 2 is more nonnormal than that of model 1. This is in agreement with the findings of Moore and Kleeman (2001) who showed that the deep convection over the warm pool in the Kleeman atmospheric model can significantly increase the nonnormality of the system. In this case inhomogeneities in the wind field and SST destroy the commutative properties of the tangent linear atmospheric latent heating operator with its adjoint, rendering the tangent linear operator nonnormal. These inhomogeneities create sources and sinks of $\mathcal{N}$ on which perturbations can feed and undergo super exponential growth. This is analogous to the more familiar role played by shear in a flow field that yield sources and sinks of energy and enstrophy due to barotropic and baroclinic processes (see Farrell 1984).

For model 3, Figs. 5e,f and 6c show that the adjoint eigenmode is similar to that of model 2. A Hovmöller diagram of the adjoint ABL θ is shown in Fig. 6d, and reveals that SST and θ are 180° out of phase over the warm pool. Recall from section 3 that the adjoint eigenmode is the optimal excitation for the corresponding

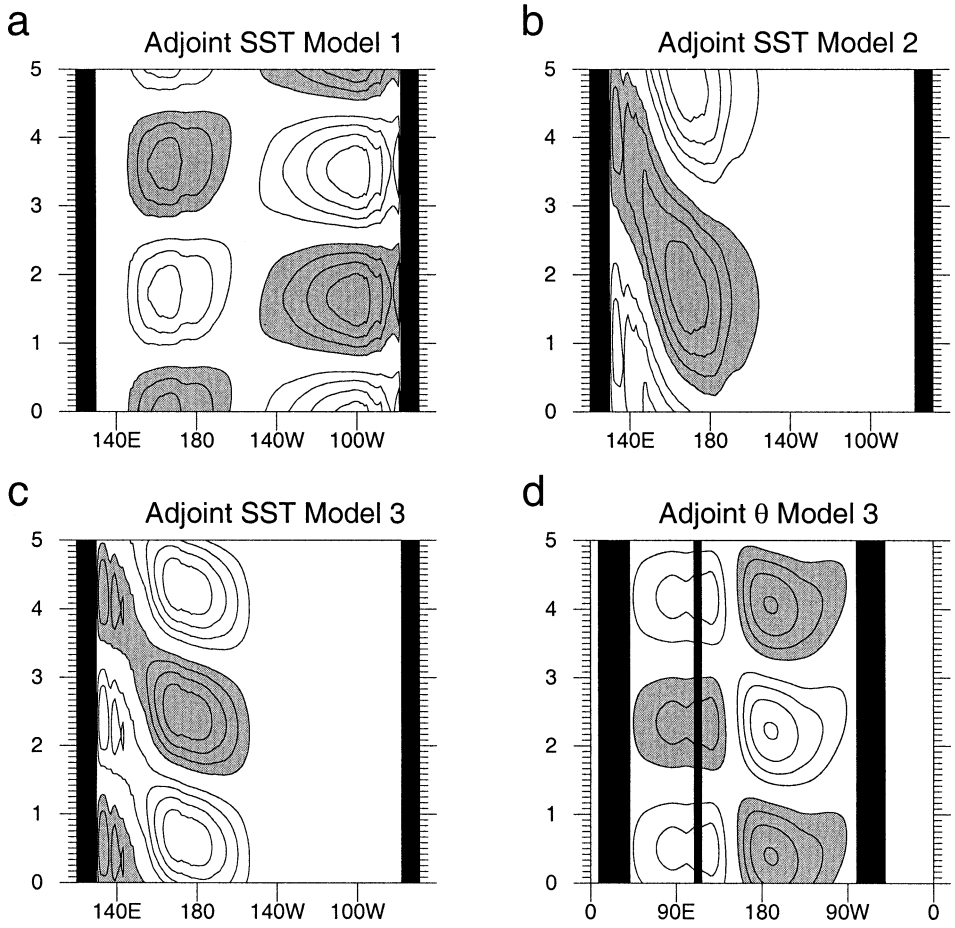

FIG. 6. Hovmöller diagrams of SST averaged between 2°N and 2°S for the most unstable adjoint eigenmode of (a) model 1, (b) model 2, and (c) model 3. (d) A Hovmöller diagram of ABL θ averaged between 2°N and 2°S for the most unstable adjoint eigenmode of model 3 for the entire longitudinal domain of the ABL. Shaded and unshaded regions indicate perturbations of opposite sign. The contour interval is arbitrary. In each case the exponential growth of the eigenmodes has been suppressed.

eigenmode, and yields super exponential growth of the chosen norm. Figures 6c,d reveals that this growth is partly achieved in model 3 via surface heat exchange induced by opposite-signed temperature perturbations of the sea surface and ABL.

The values of $|\nu|$ for the most unstable eigenmode of model 3 are shown in Table 2, which reveals that of all three models, model 3 is the most nonnormal. We will comment further on the structure of the adjoint eigenmodes in section 5 when we present the optimal perturbations.

The general structure of the most unstable eigenmodes and adjoint eigenmodes of each model are insensitive to variations in the coupling strength γ. On the other hand the values of $|\nu|$ for these eigenmodes do exhibit a sensitivity to γ as shown in Table 1. For model 1, $|\nu|$ decreases as γ increases, while for model 2, $|\nu|$ generally increases with γ.

TABLE 2. The values of $|\nu|$ where $\nu = |\mathbf{G}^T\mathbf{b}||\mathbf{G}^T\mathbf{e}|/(\mathbf{b}^T\mathbf{G}\mathbf{G}^T\mathbf{e})$ where **e** denotes the most unstable oscillatory eigenmode and **b** is the corresponding adjoint eigenmode. Values of $|\nu|$ are shown for the most unstable oscillatory eigenmodes of each coupled model using two different norms defined by **G**. Please see the main text for the appropriate definitions of **G** in each case.

	Model 1	Model 2	Model 3
L2 Norm	211	450	479
$\mathcal{T}$ Norm	24	78	83

5. Optimal perturbations of the annual mean state

From section 3, the optimal perturbations of each coupled model with respect to the norm $\mathcal{N} = \mathbf{s}^T\mathbf{G}\mathbf{G}^T\mathbf{s}$ are, by definition, the right singular vectors of $\mathbf{G}^T\mathbf{R}(\tau)$ where $\mathbf{R}(\tau)$ is the propagator of the tangent linear equation, τ is the optimal growth time, and **G** defines the norm used to measure perturbation growth. The squares of the corresponding singular values represent the growth factor μ of $\mathcal{N}$ for the optimal perturbations over the time in-

terval τ. In general, the norms used at initial and final time can be different as is often the case in numerical weather prediction where an estimate of the analysis error covariance is used to define **G** at initial time (Fan et al. 2000; Barkmeijer et al. 1998). As in section 4 we will confine our attention to autonomous systems where (1) was linearized about an annual mean state of each model, recalling that the ocean basic state is identical for all three models. As noted in section 3, this assumption greatly simplifies the analysis and interpretation of the optimal perturbations. For models 1 and 2, the state vector $\mathbf{s} = (T, S, u, v)^{\mathrm{T}}$, where T and S are the temperature and salinity perturbations, and (u, v) are the ocean current perturbations. For model 3, $\mathbf{s} = (T, S, u, v, \theta)^{\mathrm{T}}$, where θ is the boundary layer perturbation potential temperature.

For all calculations reported here, $\tau = 6$ months, which is a typical seasonal prediction forecast lead time, and the typical maturation time for ENSO events. The main limitation on the validity of the tangent linear assumption for the coupled model is nonlinearities in the OGCM. Using the same model, WVA and VWD found that for the large-scale ocean circulation of relevance to interannual climate variability, the tangent linear assumption is valid for realistic amplitude perturbations for periods of 6 months.

The singular vectors with respect to two different norms will be considered. The first norm is relevant to the problem of error growth in ENSO forecast models, while the second is of general dynamical interest.

a. A temperature norm

The first norm considered as a measure of perturbation growth is the basin-integrated squared temperature perturbation denoted $\mathcal{T}$ and defined as

$$\mathcal{T} = \int_{-H(\lambda,\varphi)}^{0} \int_{\lambda_w}^{\lambda_e} \int_{\varphi_s}^{\varphi_n} T^2 a \cos\varphi \, d\lambda \, d\varphi \, dz, \qquad (5)$$

where $H(\lambda, \varphi)$ is the spatially varying depth of the ocean; $\lambda_w \equiv \lambda_w(\varphi)$ and $\lambda_e \equiv \lambda_e(\varphi)$ define the longitude of the western and eastern boundaries, respectively; φ_n and φ_s are the northern and southern latitudes of the model domain; and a is the earth's radius. This norm is representative of the type of norm that would be of interest for ENSO prediction, and in this example would represent the temperature errors of a model forecast. The depth, latitude, and longitude ranges in (5) may be tailored for specific needs, but we consider the entire ocean domain. For $\mathcal{T}$, all elements of **G** are zero except the first N_o diagonal elements, where N_o is the number of ocean temperature grid points. At initial time $S = u = v = \theta = 0$ at all latitudes, longitudes, and depths.

The $\mathcal{T}$ growth factors of the first 10 members of the optimal perturbation spectrum of each coupled model are shown in Fig. 7. In each case, the spectrum is generally dominated by the first member, which is in agreement with the optimal perturbation spectra of the intermediate coupled models cited in section 1. For comparison, the optimal perturbation spectrum of the OGCM alone ($\gamma = 0$) using the same basic state is also shown in Fig. 7, and, in contrast to the coupled models, is very flat. The initial and final SST structures of the fastest growing optimal perturbations of models 1 and 2 are shown in Fig. 8. The SST structure for the model 3 optimal perturbation is very similar to that of model 2 and is not shown.

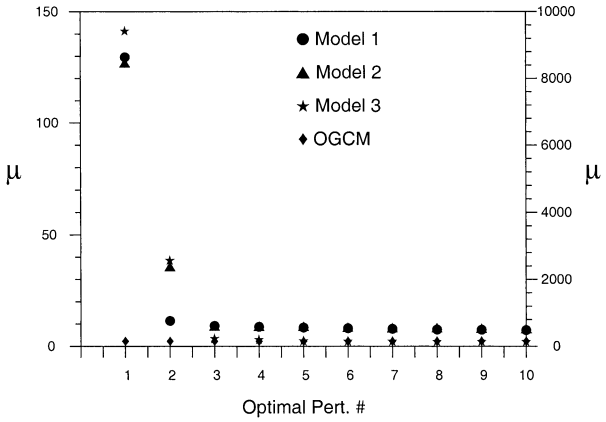

FIG. 7. The growth factor μ of the $\mathcal{T}$ norm for the first 10 members of the optimal perturbation spectrum of each coupled model and the OGCM ($\gamma = 0$) for $\tau = 6$ months. The left ordinate is for models 1 and 2 and the OGCM, and the right ordinate is for model 3.

For model 1, Fig. 8a shows that the fastest growing optimal perturbation (hereafter referred to as OP1) has an initial SST structure in the form of a dipole spanning the entire basin. The amplitude of the initial temperature perturbation is significant only down to the depth of the thermocline, as shown in Fig. 9a. The westward extension of the subsurface temperature perturbation along the thermocline evident in Fig. 9a is consistent with the recent adjoint sensitivity experiments of Galanti et al. (2002) using a similar hybrid coupled model.

The initial SST structure of OP1 is remarkably similar to the real phase of the most unstable adjoint eigenmode of model 1 shown in Fig. 5a. Figure 8b shows that OP1 evolves into an ENSO-like episode, which has a structure similar to the real phase of the most unstable eigenmode of model 1 (cf. Fig. 3a). Thus OP1 is optimal for exciting the most unstable coupled ocean–atmosphere eigenmode of the system. The similarity between OP1 and the adjoint eigenmode was anticipated because, as noted in section 3, the adjoint eigenmode is the optimal excitation for the corresponding eigenmode, and at the same time maximizes the growth of the norm associated with the inner-product that defines the adjoint operator. In this case the adjoint operator was defined with respect to the L2 norm, and the $\mathcal{T}$ norm is a subset of the L2 norm. The growth factor of $\mathcal{T}$ for OP1 is 127. However, a perturbation with the structure of the most unstable eigenmode would grow by a factor of $e^{2\sigma\tau}$,

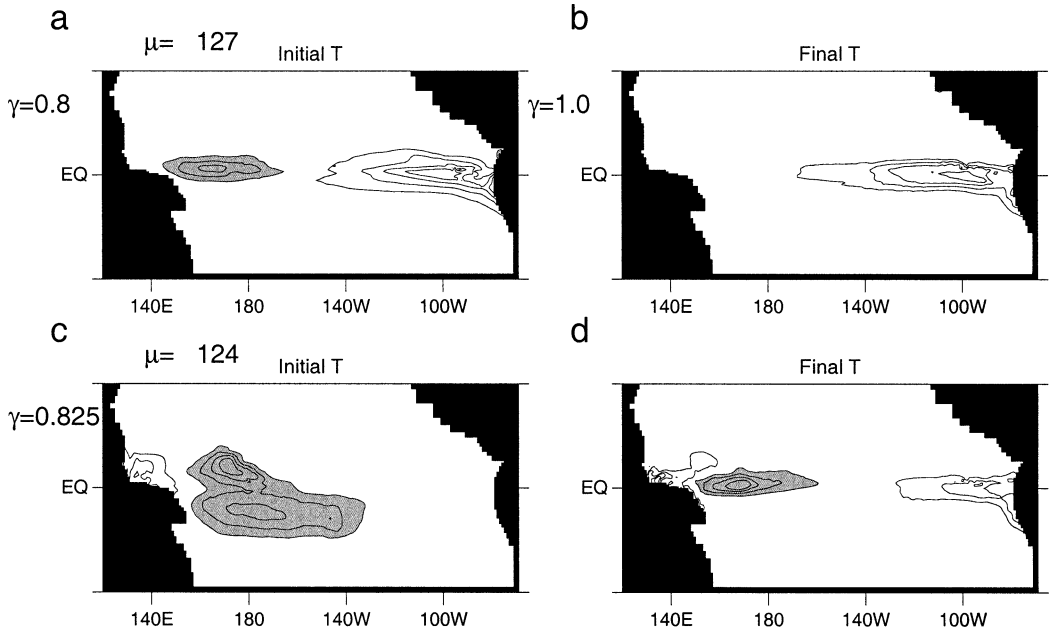

FIG. 8. The SST structure of the fastest growing optimal perturbations of (a), (b) model 1, and (c), (d) model 2 using the $\mathcal{T}$ norm for $\tau = 6$ months. The contour interval is arbitrary, and the growth factor μ of $\mathcal{T}$ is indicated.

where σ is the real part of its eigenvalue. According to Table 1, a perturbation with this structure would grow in model 1 by only a factor of 4.6 over a 6-month period. Thus the growth of OP1 over and above that expected for the most unstable eigenmode must be due to the interference of the nonnormal eigenmodes of the system.

Figures 8c,d shows that the SST structure of OP1 for model 2 is confined primarily to the west Pacific warm pool. As in model 1 the amplitude of OP1 is significant

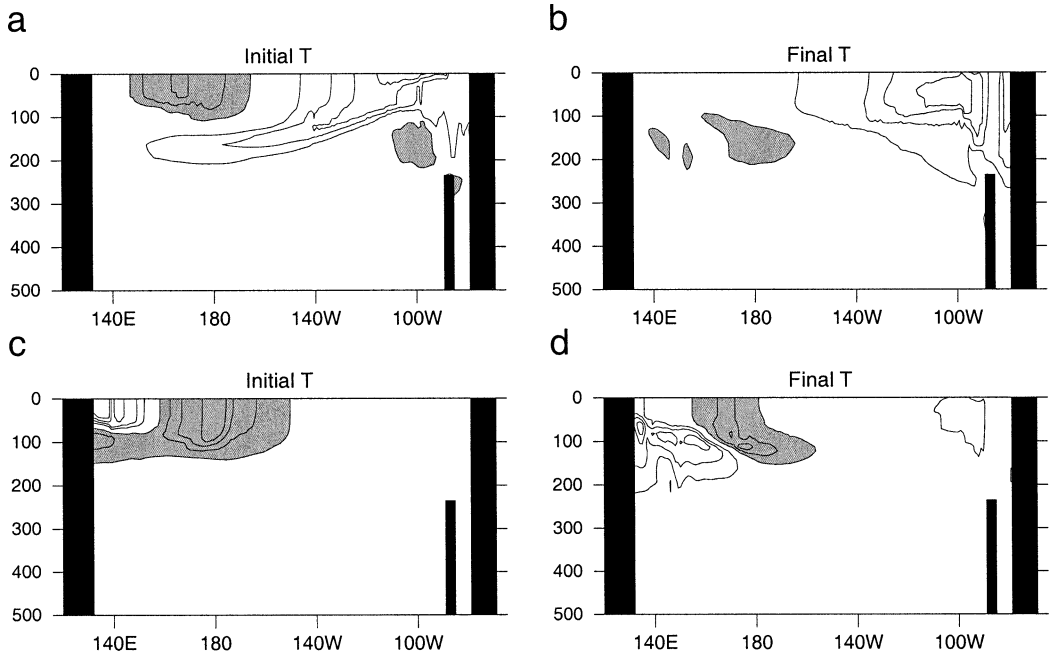

FIG. 9. Vertical sections of ocean temperature along the equator at initial and final time are shown for (a), (b) model 1, and (c), (d), model 2. Shaded and unshaded regions indicate perturbations of opposite sign. The contour interval is arbitrary.

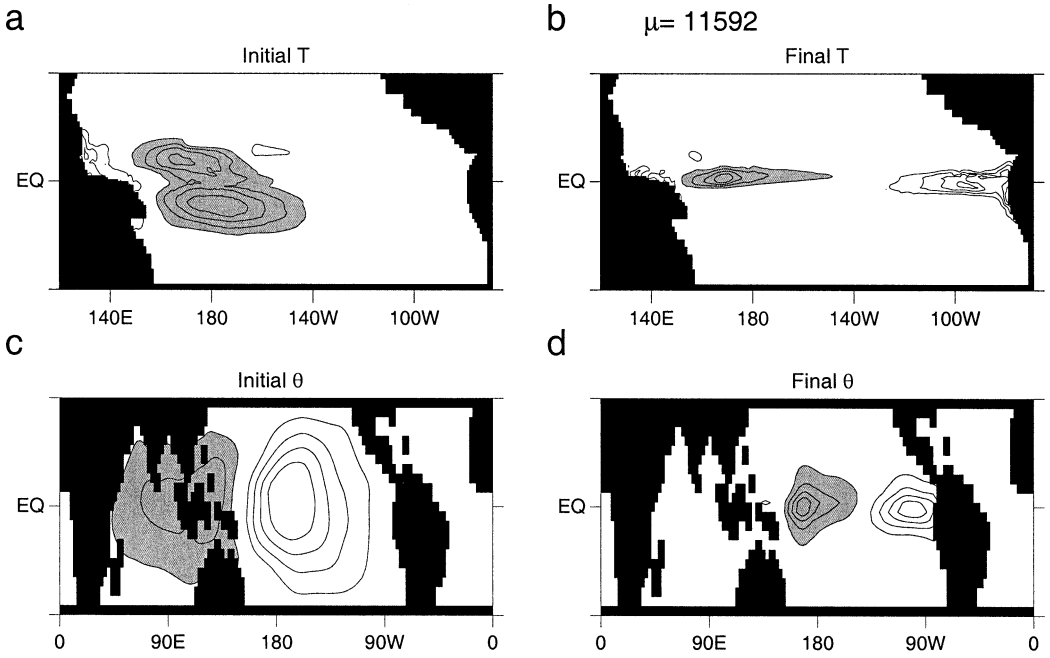

FIG. 10. The fastest growing optimal perturbation of model 3 using the $\mathcal{T}'$ norm and $\tau = 6$ months. The (a) initial and (b) final perturbation SST are shown, as are the (c) initial and (d) final ABL perturbation θ. The growth factor μ of $\mathcal{T}'$ is indicated. Shaded and unshaded regions indicate perturbations of opposite sign. The contour interval is arbitrary.

only above the thermocline (Figs. 9c,d), and OP1 bears a remarkable resemblance to the real phase of the most unstable adjoint eigenmode of model 2 shown in Fig. 5c. Figure 8d reveals that the model 2 OP1 evolves into a perturbation that is strikingly similar to the real phase of the most unstable eigenmode (cf. Fig. 3c). Thus, OP1 of model 2 is also optimal for exciting the most unstable coupled ocean–atmosphere eigenmode. The growth factor of $\mathcal{T}$ for OP1 is 124, but from Table 1, a perturbation with the structure of the most unstable eigenmode would grow by only a factor of 1.02 over a 6-month period. The growth of OP1 over and above that for the most unstable eigenmode is due to the interference of the nonorthogonal eigenmodes of model 2.

As noted, the SST structure of OP1 for model 3 is very similar to that of model 2. However, the growth factor of $\mathcal{T}$ for OP1 in model 3 is 9286, some 74 times larger than that of models 1 and 2. A perturbation with the structure of the most unstable eigenmode would grow by only a factor of 2.9. Thus, the greatly increased growth factor of the model 3 OP1 is due to the higher degree of nonnormality of the most unstable eigenmode compared to the corresponding eigenmodes of models 1 and 2 (see Table 2).

The same OP growth factor can be obtained for a range of eigenvalues and values of ν as suggested in Table 1, which shows evidence for this in models 1 and 2, which have similar OP growth factors for $\mathcal{T}$ despite differences in γ and the period, growth rate, and $|\nu|$ for their most unstable eigenmodes. Table 1 also reveals that a different choice of $\gamma = 1$ yields very disparate OP growth factors for $\mathcal{T}$ with that of model 1 being larger despite the fact that the model 2 eigenmode is more nonnormal than the model 1 eigenmode.

For unstable systems any perturbation will eventually evolve into the most unstable eigenmode. However, we note that OP growth is also possible in the coupled models when there are no growing eigenmodes. For example, Table 1 reveals that OP1 of model 2 still grows when $\gamma = 0.5$, which Table 1 suggests is well below the primary bifurcation point for this model. The structure of the optimal perturbation in this case is very similar to that in Figs. 8a,b. Therefore the results and conclusions presented here apply to both stable and unstable models. We have chosen to examine primarily unstable systems because, as noted above, only in this case can we unambiguously identify the gravest coupled eigenmodes using ARPACK.

For model 3, an additional temperature-based norm was considered:

$$\mathcal{T}' = \mathcal{T} + h_0 \int_0^{2\pi} \int_{\varphi_s}^{\varphi_n} \theta^2 a \cos\varphi \, d\lambda \, d\varphi, \qquad (6)$$

which represents $\mathcal{T}$ plus the domain-integrated squared ABL perturbation temperature, and h_0 is the boundary layer thickness. As before, $S = u = v = 0$ for the ocean at initial time, but now $\theta \neq 0$. The fastest growing optimal perturbation of model 3 for the $\mathcal{T}'$ norm is shown in Fig. 10 and is virtually identical to that for

the $\mathcal{T}$ norm. The growth factor of $\mathcal{T}'$, however, is larger than that of $\mathcal{T}$ for the same optimal growth time, and is achieved by initial temperature differences between the sea surface and ABL. Figure 10 shows that at initial time, perturbations in SST and θ are of opposite sign in the western Pacific, which enhances the air–sea heat exchange. This behavior is consistent with the 180° phase difference between the SST and θ anomalies of the most unstable adjoint eigenmode in Figs. 6c,d. An intermediate coupled model composed of the same atmosphere as model 3, but with a simplified ocean (Kleeman 1993), shows qualitatively similar behavior (not shown) confirming the controlling influence that the ABL dynamics have on the optimal perturbation structure.

b. Dynamics of $\mathcal{T}$ optimals

A detailed analysis reveals that the major contributor to the growth of $\mathcal{T}$ for OP1 is the influence of perturbation zonal heat fluxes primarily above and within the thermocline in the eastern Pacific. It is in the eastern Pacific that the model response to OP1 is largest as shown in Figs. 8b and 9b. The SST perturbations that develop in the east produce wind stress anomalies like those shown in Fig. 4b as the most unstable eigenmode emerges. These wind stress anomalies drive changes in the zonal velocity and Ekman transport in the upper ocean. The associated changes in Ekman-induced perturbation vertical velocity are primarily balanced by changes in the perturbation meridional velocity. It is the changes in perturbation zonal velocity that yield rapid growth of $\mathcal{T}$ in the thermocline since the basic-state zonal temperature gradient is relatively large in the eastern Pacific.

In model 2 it is the influence of perturbation vertical heat fluxes in the upper thermocline of the western Pacific that accounts for all of the growth of $\mathcal{T}$ for the optimal perturbation. As the most unstable eigenmode emerges, the corresponding wind stress perturbations (Fig. 4d) drive perturbation Ekman-induced vertical velocity perturbations. These are largest in the western Pacific in the upper thermocline where the basic-state vertical temperature gradient is relatively large also. The picture for the fastest growing optimal perturbation of model 3 for the $\mathcal{T}$ norm is qualitatively very similar to that for model 2. Detailed analyses for the $\mathcal{T}'$ norm (not shown), however, reveal that growth is augmented considerably by the surface exchange of heat between the ABL and ocean.

Clearly the dynamical balances that control the evolution of the perturbation norm change from model to model and are due to changes in atmospheric physics and dynamics since the ocean model is the same in each case. Different geographical regions are singled out for perturbation growth via air–sea interaction by different physical processes in the atmosphere and how these can feedback and conspire with the ocean circulation to produce rapid perturbation growth. This highlights the potentially rich dynamical behavior that may be present in CGCMs where some combination of all the above processes (and others as yet unidentified) may operate all at the same time.

c. The energy norm

The second norm considered was the basin-integrated ocean perturbation energy denoted $\mathcal{E}$ given by

$$\mathcal{E} = \frac{\rho_o}{2} \int_{-H(\lambda,\varphi)}^{0} \int_{\lambda_w}^{\lambda_e} \int_{\varphi_s}^{\varphi_n} \left[u^2 + v^2 + \left(\frac{g\rho}{\rho_o \overline{N}} \right)^2 \right] a \cos\varphi \, d\lambda \, d\varphi \, dz, \tag{7}$$

where ρ is the perturbation density, $\overline{N}(\lambda, \varphi, z)$ is the Brunt–Väisälä frequency of the basic state $\mathbf{S}_0$, and g is the acceleration due to gravity. For simplicity, the linear equation of state $\rho = \rho_o(\alpha T + \beta S)$ was used to derive (7) where $\alpha = -2.489 \times 10^{-4}$ kg m^{-3} K^{-1} and $\beta = 7.453 \times 10^{-4}$ kg m^{-3} psu^{-1} are the expansion coefficients for temperature and salinity, respectively, and $\rho_o = 1020$ kg m^{-3}. This norm is of obvious interest since it indicates how perturbation energy can be exchanged with the basic-state flow via barotropic and baroclinic processes. In model 3, $\theta = 0$ at initial time since θ does not contribute to $\mathcal{E}$. For $\mathcal{E}$ the matrix $\mathbf{X}$ of section 3 is singular and cannot be factorized. However, this problem may be circumvented if $S = 0$ at initial time.

The SST and upper-ocean temperature along the equator of OP1 for model 1 using the $\mathcal{E}$ norm is shown in Figs. 11a,b and 12a,b. At the surface, the initial structure of OP1 for the $\mathcal{E}$ norm is confined to the far eastern tropical Pacific (Fig. 11a) and follows the thermocline at depth (Fig. 12a). Like the $\mathcal{T}$ optimal, the $\mathcal{E}$ optimal evolves into the most unstable eigenmode (cf. Fig. 3b).

OP1 of model 2 using the $\mathcal{E}$ norm is shown in Figs. 11c,d and Figs. 12c,d. A comparison of Fig. 11c with Fig. 8c reveals a number of similarities between the initial SST structures of the $\mathcal{E}$ and $\mathcal{T}$ optimals in the western Pacific. Like the $\mathcal{T}$ optimal, the $\mathcal{E}$ optimal evolves into the most unstable eigenmode of model 2 (see Fig. 3d). Figure 12c reveals a subsurface dipole in the vertical in the far western Pacific centered around 75-m depth.

The surface signature of the fastest growing $\mathcal{E}$ optimal of model 3 is similar to that of the $\mathcal{T}$ optimal as shown

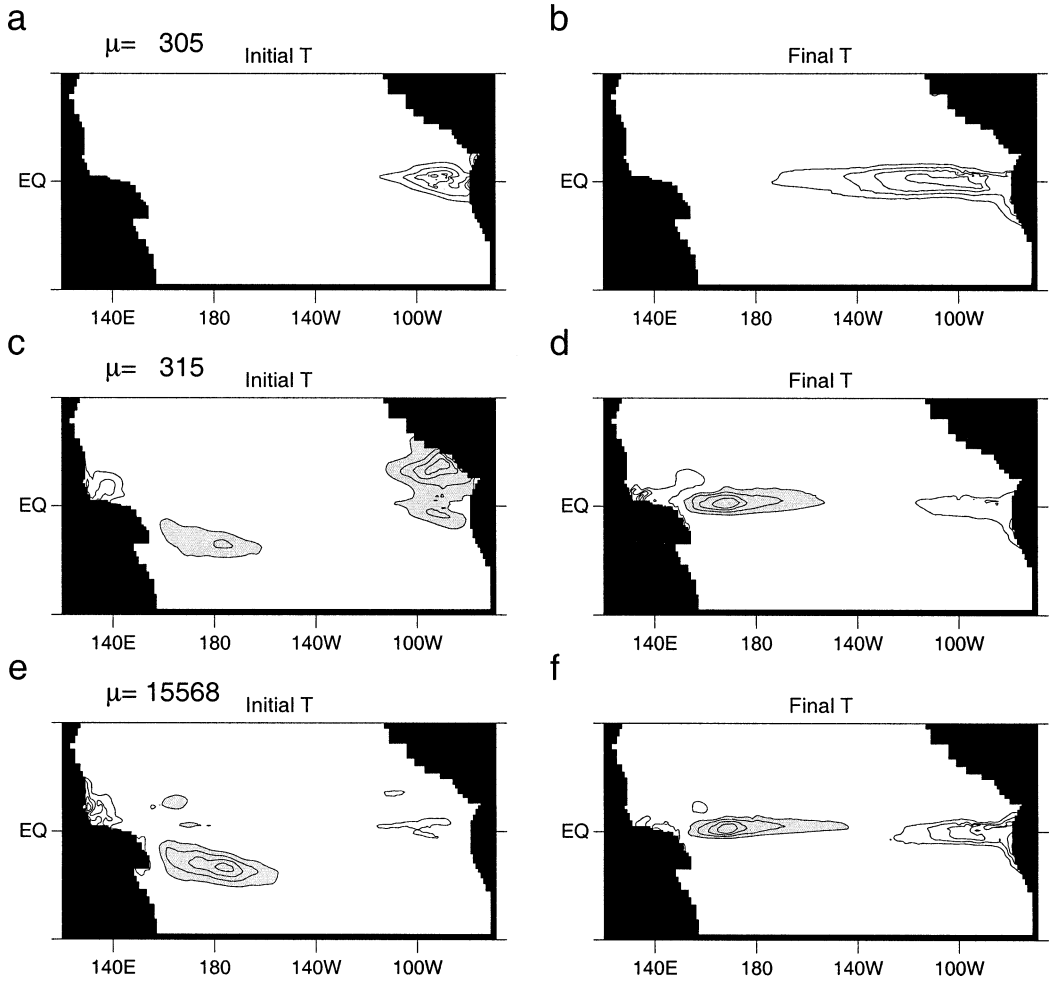

FIG. 11. Same as Fig. 8, but for the fastest growing optimal perturbation of the $\mathcal{E}$ norm. (e), (f) OP1 for model 3.

in Fig. 11e. The subsurface structure is similar to that of model 2 (Fig. 12e) and like the other optimal perturbations it evolves into the most unstable eigenmode. The growth factor $\mu = 15\,568$, is some 50 times larger than that of models 1 and 2.

d. Dynamics of $\mathcal{E}$ optimals

A detailed analysis of the perturbation energetics of each model reveals that, in all cases, the growth of $\mathcal{E}$ is dominated by baroclinic processes. In model 1, it is the east Pacific that is favored for the growth of $\mathcal{E}$ where the equatorial undercurrent approaches the surface. To understand this, it is useful to remember that the growth of perturbations by baroclinic processes is a form of thermal convection (Pedlosky 1979). For this to occur horizontal gradients of potential density are required. At middle latitudes these gradients can be maintained by the Coriolis force as described by the thermal wind relation. At the equator in the Pacific Ocean, where $f = 0$, the zonal density and pressure gradients are maintained by the easterly surface trade winds. The region where the undercurrent rises to the surface in the eastern Pacific is also a region of large horizontal temperature and density gradients (cf. Fig. 2b) since the undercurrent itself is driven by the zonal pressure gradient below the surface within the thermocline (Gill 1982). Thus, the necessary conditions for baroclinic energy release are present in this region.

In models 2 and 3, the western Pacific is favored for growth of $\mathcal{E}$ in regions where the basic state stratification is weak since it is here that thermal convection can occur most readily. In general, the geographic locations of maximum growth of $\mathcal{T}$ and $\mathcal{E}$ largely coincide, indicating that the same regions of the ocean are susceptible to nonnormal perturbation growth irrespective of the choice of perturbation growth norm.

6. Summary and conclusions

We have explored the properties and dynamics of the optimal perturbations of a hierarchy of hybrid coupled ocean–atmosphere models of the tropical Pacific. The coupled model hierarchy was constructed so as to yield

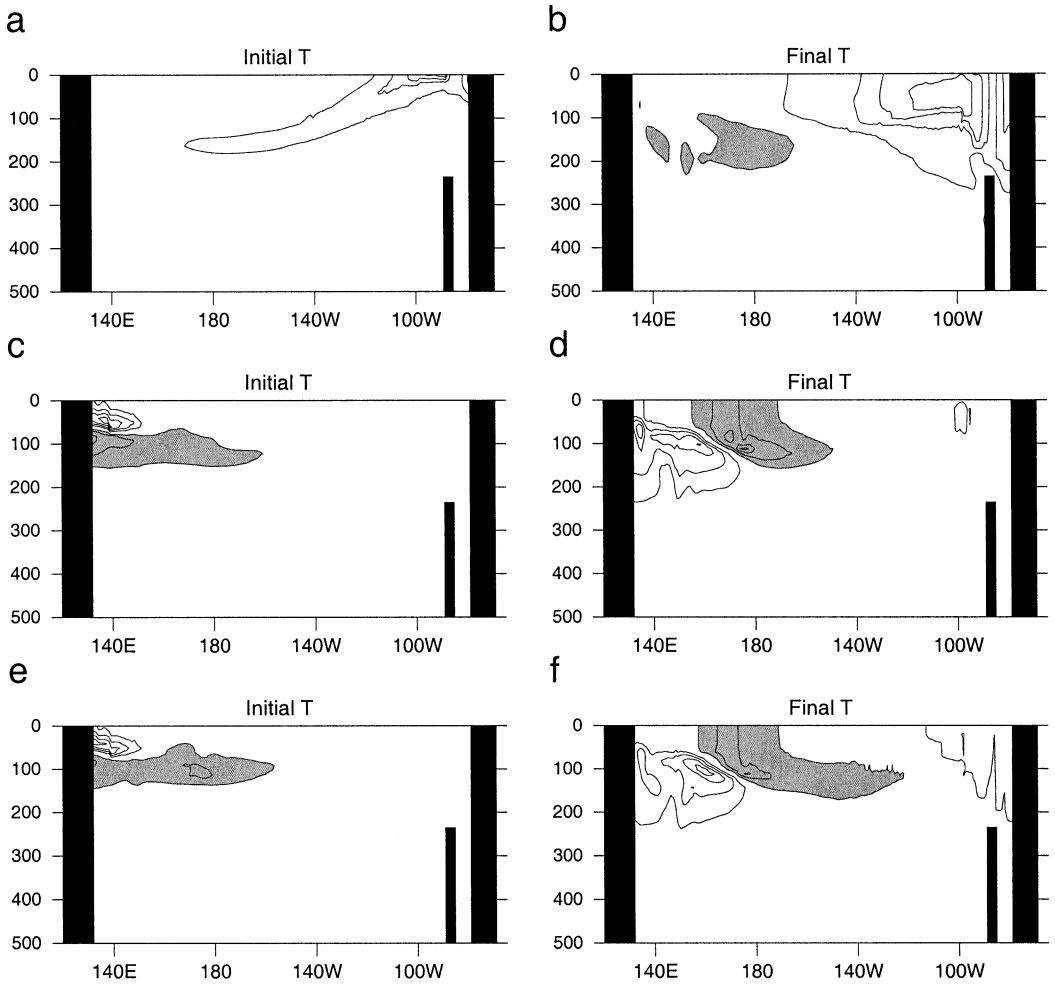

FIG. 12. Same as Fig. 9, but for the fastest growing optimal perturbation of the $\mathcal{E}$ norm. (e), (f) OP1 for model 3.

information about the dependence of optimal perturbation structure and growth rate on atmospheric dynamics and air–sea interaction processes. Each hybrid coupled model shares the same OGCM while three different atmospheric models of varying complexity were employed. These same three atmospheric models have also been coupled to a 1½-layer ocean model, and the resulting intermediate coupled models employed in a series of studies by some of the authors to explore the optimal perturbations of ENSO (Moore and Kleeman 1996, 1997a,b). The commonality of the atmospheric models employed in these two sets of models means that a direct comparison of the results of this and previous studies is possible and meaningful. The optimal perturbation studies using the aforementioned intermediate coupled models, and others cited in section 1, have revealed the following.

1) The optimal perturbation spectrum is typically dominated by one member.
2) The dominant optimal perturbation evolves into the gravest coupled ocean–atmosphere eigenmode of the system.
3) The optimal perturbations of coupled models that utilize statistical models of the atmosphere emphasize the sensitivity of the central and eastern Pacific to perturbation growth.
4) The optimal perturbations of coupled models that include a parameterization for deep atmospheric convection emphasize the sensitivity of the western Pacific warm pool to perturbation growth.
5) The degree of nonnormality of the gravest coupled ocean–atmosphere eigenmode is strongly controlled by the atmospheric component of the coupled system.
6) The structure of the optimal perturbations for models with a dynamical atmosphere are relatively insensitive to the choice of perturbation growth norm.

These same conclusions can, by and large, be drawn from the experiments presented here. Therefore, the results of previous studies of the optimal perturbations for

ENSO using intermediate coupled models are supported by the hybrid coupled models used here and appear to be robust.

There are a few noteworthy differences between the results of the intermediate model and hybrid model experiments. First, with regard to conclusion 1, Fig. 7 shows that the second member of the optimal perturbation spectra of models 2 and 3 has a growth factor that is significantly larger than that of the uncoupled OGCM. For both models these OPs take the form of SST dipoles centered near the dateline, which, after 6 months, evolve into large-scale SST anomalies confined to the equatorial eastern and central Pacific. These OPs may be associated with stable coupled modes of the system that we were unable to isolate using ARPACK. In model 1, OP2 exhibits only modest growth. In this case the number of degrees of freedom of the system is limited by the fact that only two SVD SST and wind stress patterns were retained in the atmospheric model. Therefore only two unique combinations of these SVD patterns are possible; OP1 represents one of these while OP2 represents the other.

Second, with regard to conclusion 6, the optimal perturbations of models 1 and 2 do show some sensitivity to the choice of norm. However, the geographical location of maximum growth is the same for both norms considered, both in stable and unstable cases.

The results from the hybrid coupled models used here and those from the intermediate coupled models (summarized above) suggest that the structure of the optimal perturbations of low-frequency tropical coupled ocean–atmosphere models such as ENSO is influenced more by atmospheric dynamics and the details of the air–sea interactions processes than by ocean dynamics. While this may appear to be an obvious statement, it is nonetheless an important conclusion. Until now, the relative role of the atmosphere and ocean in controlling optimal perturbation structure has been unclear, and optimal perturbation calculations performed using a wide range of intermediate coupled models have made this a controversial issue. This study, in conjunction with those that utilize intermediate coupled models using the same atmospheric models is the first attempt to address this important issue in moderately complex coupled models.

Given the dependence of optimal perturbation structure on atmospheric physics and dynamics, the hybrid coupled models used here suggest a number of potentially important perturbation growth mechanisms that may influence the growth of disturbances in complex CGCMs and limit the predictability of ENSO. Using hybrid coupled models to compute optimal perturbations for generating CGCM ensemble forecasts may therefore be one way of circumventing the challenging problem (discussed in section 2) of computing such perturbations directly using a CGCM. The ability of CGCMs to support perturbation growth via the mechanisms suggested by the hybrid coupled models is currently under investigation.

Acknowledgments. AMM was supported by a research grant from the NSF Division of Climate (ATM-9809790). We are indebted to Dan Sorensen for making ARPACK so freely available.

REFERENCES

Aiken, C. M., A. M. Moore, and J. H. Middleton, 2002: The nonnormality of coastal ocean flows around obstacles, and their response to stochastic forcing. *J. Phys. Oceanogr.*, **32,** 2955–2974.

Balmaseda, M. A., M. K. Davey, and D. L. T. Anderson, 1995: Decadal and seasonal dependence of ENSO predictive skill. *J. Climate*, **8,** 2705–2715.

Barkmeijer, J., M. van Gijzen, and F. Bouttier, 1998: Singular vectors and estimates of the analysis-error covariance metric. *Quart. J. Roy. Meteor. Soc.*, **124,** 1695–1713.

Blumenthal, M. B., 1991: Predictability of a coupled ocean–atmosphere model. *J. Climate*, **4,** 766–784.

Bretherton, C. S., C. Smith, and J. M. Wallace, 1992: An intercomparison of methods for finding coupled patterns in climate data. *J. Climate*, **5,** 541–560.

Buizza, R., and T. N. Palmer, 1995: The singular-vector structure of the atmospheric global circulation. *J. Atmos. Sci.*, **52,** 1434–1456.

Chen, Y.-Q., D. S. Battisti, T. N. Palmer, J. Barsugli, and E. S. Sarachik, 1997: A study of the predictability of tropical Pacific SST in a coupled atmosphere/ocean model using singular vector analysis: The role of the annual cycle and the ENSO cycle. *Mon. Wea. Rev.*, **125,** 831–845.

Courant, R., and D. Hilbert, 1953: *Methods of Mathematical Physics.* Vol. I. Wiley Interscience, 559 pp.

Eckert, C., 1999: On predictability limits of ENSO. A study performed with a simplified model of the tropical Pacific ocean–atmosphere system. Rep. 55, 76 pp. [Available from Max Planck Institut für Meteorologie, Bundesstrasse 55 D-20146, Hamburg, Germany.]

Fan, Y., M. R. Allen, D. L. T. Anderson, and M. A. Balmaseda, 2000: How predictability depends on the nature of uncertainty in initial conditions in a coupled model of ENSO. *J. Climate*, **13,** 3298–3313.

Farrell, B. F., 1982: The initial growth of disturbances in baroclinic flow. *J. Atmos. Sci.*, **39,** 1663–1686.

——, 1984: Modal and nonmodal baroclinic waves. *J. Atmos. Sci.*, **41,** 668–673.

——, 1985: Transient growth of damped baroclinic waves. *J. Atmos. Sci.*, **42,** 2718–2727.

——, 1988a: Optimal excitation of neutral Rossby waves. *J. Atmos. Sci.*, **45,** 163–172.

——, 1988b: Optimal excitation of perturbations in viscous shear flow. *Phys. Fluids*, **31,** 2093–2102.

——, 1989a: Optimal excitation of baroclinic waves. *J. Atmos. Sci.*, **46,** 1193–1206.

——, 1989b: Transient development in confluent and diffluent flow. *J. Atmos. Sci.*, **46,** 3279–3288.

——, and A. M. Moore, 1992: An adjoint method for obtaining the most rapidly growing perturbation to oceanic flows. *J. Phys. Oceanogr.*, **22,** 338–349.

——, and P. J. Ioannou, 1996: Generalized stability theory. Part I: Autonomous operators. *J. Atmos. Sci.*, **53,** 2025–2040.

——, and ——, 1999: Perturbation growth and structure in time-dependent flows. *J. Atmos. Sci.*, **56,** 3622–3639.

Frederiksen, J. S., 1982: A unified three-dimensional instability theory for the onset of blocking and cyclogenesis. *J. Atmos. Sci.*, **39,** 969–982.

Galanti, E., E. Tziperman, M. Harrison, A. Rosati, R. Giering, and Z. Sirkes, 2002: The equatorial thermocline outcropping—A seasonal control on the tropical Pacific ocean–atmosphere instability strength. *J. Climate*, **15,** 2721–2739.

Gill, A. E., 1980: Simple solutions for heat-induced tropical circulation. *Quart. J. Roy. Meteor. Soc.,* **106,** 447–462.
——, 1982: *Atmosphere–Ocean Dynamics.* Academic Press, 662 pp.
Hartmann, D. L., R. Buizza, and T. N. Palmer, 1995: Singular vectors: the effect of spatial scale on linear growth of disturbances. *J. Atmos. Sci.,* **52,** 3885–3894.
Jin, F.-F., and J. D. Neelin, 1993a: Modes of interannual tropical ocean–atmosphere interaction—A unified view. Part I: Numerical results. *J. Atmos. Sci.,* **50,** 3477–3503.
——, and ——, 1993b: Modes of interannual tropical ocean–atmosphere interaction—A unified view. Part III: Analytical results in fully coupled cases. *J. Atmos. Sci.,* **50,** 3523–3540.
Kleeman, R., 1991: A simple model of the atmospheric response to ENSO sea surface temperature anomalies. *J. Atmos. Sci.,* **48,** 3–18.
——, 1993: On the dependence of hindcast skill on ocean thermodynamics in a coupled ocean–atmosphere model. *J. Climate,* **6,** 2012–2033.
——, and S. B. Power, 1995: A simple atmospheric model for the surface heat flux of the ocean. *J. Phys. Oceanogr.,* **25,** 92–105.
——, J. P. McCreary, and B. A. Klinger, 1999: A mechanism for the decadal variation of ENSO. *Geophys. Res. Lett.,* **26,** 1743.
Latif, M., and A. Villwock, 1990: Interannual variability in the tropical Pacific as simulated in coupled ocean–atmosphere models. *J. Mar. Systems,* **1,** 51–60.
——, and M. Flügel, 1991: An investigation of short range climate predictability in the tropical Pacific. *J. Geophys. Res.,* **96,** 2661–2673.
Lau, K.-M., and P. H. Chan, 1985: Aspects of the 40–50 day oscillation during the northern winter as inferred from outgoing longwave radiation. *Mon. Wea. Rev.,* **113,** 1889–1909.
——, and ——, 1986: The 40–50 day oscillation and the El Nino/Southern Oscilation: A new perspective. *Bull. Amer. Met. Soc.,* **67,** 533–534.
——, and ——, 1988: Intraseasonal and interannual variations of tropical convection: A possible link between 40–50 day oscillation and ENSO? *J. Atmos. Sci.,* **45,** 506–521.
Lehoucq, R. B., D. C. Sorensen, and C. Yaug, 1997: *ARPACK Users' Guide: Solution of Large Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods.* Rice University, 140 pp.
Levitus, S., 1982: Climatological atlas of the world ocean. NOAA Prof. Paper 13, U.S. Government Printing Office, 173 pp.
Macías, J., D. Stephenson, L. Terray, M. Balmaseda, and D. L. T. Anderson, 1996: ENSO and seasonal variability in a hybrid coupled model of the tropical Pacific. *Annales Geophysicae,* **14,** C647.
Madec, G., P. Delecluse, M. Imbard, and C. Levy, 1998: *OPA 8.1 Ocean General Circulation Model Reference Manual.* Institut Pierre Simon Laplace des Sciences l'Environnement Global, LODYC, Universite Pierre et Marie Curie, 91 pp.
McPhaden, M. J., 1999: Genesis and evolution of the 1997–98 El Niño. *Science,* **283,** 950–954.
Molteni, F., R. Mureau, and T. N. Palmer, 1993: Predictability and finite time instability of the northern winter circulation. *Quart. J. Roy. Meteor. Soc.,* **119,** 269–298.
Moore, A. M., and B. F. Farrell, 1993: Rapid perturbation growth on spatially and temporally varying oceanic flows determined using an adjoint method: Application to the Gulf Stream. *J. Phys. Oceanogr.,* **23,** 1682–1702.
——, and R. Kleeman, 1996: The dynamics of error growth and predictability in a coupled model of ENSO. *Quart. J. Roy. Meteor. Soc.,* **122,** 1405–1446.
——, and ——, 1997a: The singular vectors of a coupled ocean–atmosphere model of ENSO. Part I: Thermodynamics, energetics and error growth. *Quart. J. Roy. Meteor. Soc.,* **123,** 953–981.
——, and ——, 1997b: The singular vectors of a coupled ocean–atmosphere model of ENSO. Part II: Sensitivity studies and dynamical significance. *Quart. J. Roy. Meteor. Soc.,* **123,** 983–1006.
——, and A. J. Mariano, 1999: The dynamics of error growth and predictability in a model of the Gulf Stream. Part I: Singular vector analysis. *J. Phys. Oceanogr.,* **29,** 158–176.
——, and ——, 2001: The differences between the optimal perturbations of coupled models of ENSO. *J. Climate,* **14,** 138–163.
——, C. L. Perez, and J. Zavala-Garay, 2002: A non-normal view of the wind-driven ocean circulation. *J. Phys. Oceanogr.,* **32,** 2681–2705.
Neelin, J. D., and F.-F. Jin, 1993: Modes of interannual tropical ocean–atmosphere interaction—A unified view. Part II: Analytical results in the weak coupling limit. *J. Atmos. Sci.,* **50,** 3504–3522.
Palmer, T. N., and D. L. T. Anderson, 1994: Prospects for seasonal forecasting. *Quart. J. Roy. Meteor. Soc.,* **120,** 755–793.
Pedlosky, J., 1979: *Geophysical Fluid Dynamics.* Springer-Verlag, 710 pp.
Reynolds, R. W., and T. M. Smith, 1994: Improved global sea surface temperature analysis. *J. Climate,* **7,** 929–948.
Simmons, A. J., and B. J. Hoskins, 1976: Baroclinic instability on the sphere: Normal modes of the primitive and quasigeostrophic equations. *J. Atmos. Sci.,* **33,** 1454–1477.
Syu, H.-H., J. D. Neelin, and D. S. Gutzler, 1995: Seasonal and interannual variability in a hybrid coupled GCM. *J. Climate,* **8,** 2121–2143.
Thompson, C. J., 1998: Initial conditions for optimal growth in a coupled ocean–atmosphere model of ENSO. *J. Atmos. Sci.,* **55,** 537–557.
Vialard, J., and P. Delecluse, 1998a: An OGCM study for the TOGA decade. Part I: Role of salinity in the physics of the western Pacific fresh pool. *J. Phys. Oceanogr.,* **28,** 1071–1088.
——, and ——, 1998b: An OGCM study for the TOGA decade. Part II: Barrier layer formation and variability. *J. Phys. Oceanogr.,* **28,** 1089–1106.
——, C. Menkes, J.-P. Boulanger, P. Delecluse, E. Guilyardi, and M. J. McPhaden, 2001: A model study of oceanic mechanisms affecting equatorial Pacific sea surface temperature during the 1997–98 El Niño. *J. Phys. Oceanogr.,* **31,** 1649–1675.
——, A. T. Weaver, and P. Delecluse, 2003: Three- and four-dimensional variational assimilation with a general circulation model of the tropical Pacific Ocean. Part II: Physical validation. *Mon. Wea. Rev.,* in press.
Weaver, A. T., J. Vialard, and D. L. T. Anderson, 2003: Three- and four-dimensional variational assimilation with a general circulation model of the tropical Pacific Ocean. Part I: Formulation, internal diagnostics, and consistency checks. *Mon. Wea. Rev.,* in press.
Webster, P. J., 1972: Response of the tropical atmosphere to local steady forcing. *Mon. Wea. Rev.,* **100,** 518–541.
Xue, Y., M. A. Cane, S. E. Zebiak, and M. B. Blumenthal, 1994: On the prediction of ENSO: A study with a low order Markov model. *Tellus,* **46A,** 512–528.
——, ——, and ——, 1997a: Predictability of a coupled model of ENSO using singular vector analysis. Part I: Optimal growth in seasonal background and ENSO cycles. *Mon. Wea. Rev.,* **125,** 2043–2056.
——, ——, and ——, 1997b: Predictability of a coupled model of ENSO using singular vector analysis. Part II: Optimal growth and forecast skill. *Mon. Wea. Rev.,* **125,** 2057–2073.