Clim Dyn
DOI 10.1007/s00382-011-1034-0

# An empirical model of tropical ocean dynamics

**Matthew Newman · Michael A. Alexander · James D. Scott**

Received: 23 February 2010 / Accepted: 12 February 2011


**Abstract** To extend the linear stochastically forced paradigm of tropical sea surface temperature (SST) variability to the subsurface ocean, a linear inverse model (LIM) is constructed from the simultaneous and 3-month lag covariances of observed 3-month running mean anomalies of SST, thermocline depth, and zonal wind stress. This LIM is then used to identify the empirically-determined linear dynamics with physical processes to gauge their relative importance to ENSO evolution. Optimal growth of SST anomalies over several months is triggered by both an initial SST anomaly and a central equatorial Pacific thermocline anomaly that propagates slowly eastward while leading the amplifying SST anomaly. The initial SST and thermocline anomalies each produce roughly half the SST amplification. If interactions between the sea surface and the thermocline are removed in the linear dynamical operator, the SST anomaly undergoes less optimal growth but is also more persistent, and its location shifts from the eastern to central Pacific. Optimal growth is also found to be essentially the result of two stable eigenmodes with similar structure but differing 2- and 4-year periods evolving from initial destructive to constructive interference. Variations among ENSO events could then be a consequence not of changing stability characteristics but of random excitation of these two eigenmodes, which represent different balances between surface and subsurface coupled dynamics. As found in previous studies, the impact of the additional variables on LIM SST forecasts is relatively small for short time scales. Over time intervals greater than about 9 months, however, the additional variables both significantly enhance forecast skill and predict lag covariances and associated power spectra whose closer agreement with observations enhances the validation of the linear model. Moreover, a secondary type of optimal growth exists that is not present in a LIM constructed from SST alone, in which initial SST anomalies in the southwest tropical Pacific and Indian ocean play a larger role than on shorter time scales, apparently driving sustained off-equatorial wind stress anomalies in the eastern Pacific that result in a more persistent equatorial thermocline anomaly and a more protracted (and predictable) ENSO event.

**Keywords** Sea-surface temperatures · ENSO · Climate variability · Ocean-atmosphere interaction · Coupled models

M. Newman · J. D. Scott
CIRES Climate Diagnostics Center, University of Colorado, Boulder, CO, USA

M. Newman (✉) · M. A. Alexander · J. D. Scott
Physical Sciences Division,
NOAA Earth System Research Laboratory,
Boulder, CO, USA
e-mail: matt.newman@noaa.gov

## 1 Introduction

Penland and Sardeshmukh (1995; hereafter PS95) showed that observed tropical SST variability may be viewed as resulting from the linear stochastically-forced system

$$\frac{d\mathbf{T}_o}{dt} = \mathbf{L}\mathbf{T}_o + \mathbf{F}_s, \tag{1}$$

where $\mathbf{T}_o(t)$ are sea surface temperature (SST) anomalies throughout the tropical domain, $\mathbf{L}$ is a stable linear dynamical operator, and $\mathbf{F}_s$ is white noise (which can have spatial coherence and seasonal dependence). PS95 derived $\mathbf{L}$ empirically by constructing a Linear Inverse Model

Published online: 15 March 2011

(LIM); a similar result has also been suggested by several studies using intermediate coupled models (Flügel and Chang 1996; Moore and Kleeman 1997, 1999; Penland et al. 2000; Thompson and Battisti 2001). Equation 1 may be thought of as a multivariate extension to the univariate red noise null hypothesis for SST variability first proposed by Frankignoul and Hasselman (1977), in which the slowly evolving ocean integrates forcing by rapidly evolving weather noise. The "effectively linear" stochastic approximation (1) is similarly valid when nonlinear processes decorrelate much more rapidly than linear processes (e.g., Papanicolaou and Kohle 1974; Hasselmann 1976; see also Penland 1996), so that nonlinear terms may be parameterized as a (second) linear process plus unpredictable white noise. [This is in contrast to "linearization" of the prevailing equations, where it is assumed that the magnitude, rather than the time scale, of the nonlinear terms is small.]

Equation 1 is a good approximation of the observed statistics of ENSO evolution. For example, a LIM constructed from the zero- and one week lagged covariance statistics of weekly atmospheric and SST data more faithfully reproduces the entire power spectrum on seasonal to interannual time scales of monthly ENSO SST variability than do the corresponding spectra from virtually all ensemble members of the "twentieth-century" (20c3m) IPCC AR4 coupled GCMs (Newman et al. 2009). Also, forecasts made using the LIM are not only competitive with operational GCMs such as NCEP's coupled forecast system (CFS) but, particularly for forecast leads of 6 months and greater, generally have higher skill throughout most of the Pacific Basin (not shown), indicating that (1) is also a good approximation of the case-to-case evolution of SST anomalies.

While (1) has proven highly useful as a dynamical model for SSTs, it is not as satisfying as a dynamical model of the full ocean. Equation 1 assumes that SST is a proxy for the complete coupled atmosphere-ocean state vector, which PS95 notes is a plausible scenario so long as essentially instantaneous (on the seasonal time scale) SST-induced changes in surface winds drive an immediate thermocline response, and/or in the fast-wave limit (Neelin 1991; Neelin and Jin 1993) in which wave dynamics allow the deeper ocean to rapidly adjust to changes in SSTs. In this case, all other physically important ocean variables can be effectively regressed onto SST; that is, $\mathbf{X} = \mathbf{B}\,\mathbf{T}_o + \mathbf{r}$, where $\mathbf{X}$ is the remaining ocean state vector, $\mathbf{B}$ is a regression matrix and $\mathbf{r}$ is a white residual. (Note that this regression is generally not local.) Consequently, $\mathbf{L}$ in (1) includes all the *implicit* effects of $\mathbf{X}$ upon the evolution of $\mathbf{T}_o$. Clearly, this is at least partly true, evidenced both by a strong observed relationship between thermocline depth and SST (e.g., Zelle et al. 2004) and by the success of (1) as a model. Also, studies that constructed a LIM by including a variable corresponding to ocean heat content (Johnson et al. 2000; Xue et al. 2000) did not find statistically significant improvement in forecast skill up to a few seasons ahead.

However, for longer time scales (and thus forecast leads) we expect deeper ocean physics is too important for this simple relationship to hold, if for no other reason than $\mathbf{r}$ is unlikely to be white. For example, near-equatorial Kelvin and Rossby waves are an integral part of the El Niño phenomena, and within the delayed oscillator paradigm the latter set the time scale for ENSO (Schopf and Suarez 1988; Battisti and Hirst 1989). In the subtropics, thermocline variability clearly propagates west in accordance with Rossby wave theory and appears to influence the upper ocean along the equator via Kelvin waves along the western boundary (Capotondi and Alexander 2001). In the recharge-discharge paradigm for ENSO, zonally-averaged export of heat to and from the equator is the critical process for ENSO dynamics (Jin 1997). Indeed, heat content anomalies at 5°N, important in either paradigm, contribute to ENSO predictability at 6–12 month lead times (Latif and Graham 1992).

Moreover, if the linear stochastically forced paradigm is appropriate for tropical climate dynamics, then the LIM must also appropriately represent the correct balance of coupled air-sea physical processes, both linear and nonlinear. This makes LIM a potentially important diagnostic tool that could be used to understand, for example, how misrepresented physical processes might affect the simulation of ENSO in coupled models, so a LIM based on SST alone is necessarily insufficient and both atmospheric and subsurface information must be included in its state vector. As a first step, in this study we diagnose a LIM constructed from an extended state vector that includes SST, thermocline depth (as represented by the 20°C isotherm depth) and surface zonal wind stress, a combination of variables similar to that earlier employed by Johnson et al. (2000) and Xue et al. (2000) using more limited datasets. The resulting LIM, described in Sect. 2, improves both tropical SST predictions at longer leads and the simulation of power spectra and lag co-variance statistics relative to the SST-only LIM (shown in the "Appendix"), giving confidence in its diagnostic usefulness. Initial analysis of the different terms within the LIM and their effect upon the statistics of SST variability is made in Sect. 3. A discussion of how thermocline depth and surface wind stress variability interacts with the most rapidly amplifying SST anomalies and how they together evolve into ENSO events is then presented in Sect. 4, along with how two key eigenmodes of $\mathbf{L}$ that represent different balances of dynamical processes contribute to this evolution. Concluding remarks are made in Sect. 5.

## 2 Model details and data

LIM may be broadly defined as extracting the dynamical evolution operator $\mathbf{L}$ of the system

$$\frac{d\mathbf{x}}{dt} = \mathbf{L}\mathbf{x} + \xi \tag{2}$$

from its observed statistics, as described for example in PS95 (see also Penland 1989, 1996; Winkler et al. 2001; Newman et al. 2003, 2009; Newman 2007; Alexander et al. 2008; Newman and Sardeshmukh 2008; Kallummal and Kirtman 2008). The statistics of the noise forcing $\mathbf{Q} = \langle \xi \xi^T \rangle dt$ are then determined by the corresponding Fluctuation–Dissipation relationship

$$d\mathbf{C}(0)/dT = \mathbf{L}\mathbf{C}(0) + \mathbf{C}(0)\mathbf{L}^T + \mathbf{Q}, \tag{3}$$

given the observed seasonally-varying (where $T$ is time of year) data covariance matrix $\mathbf{C}(0) = \langle \mathbf{x}\mathbf{x}^T \rangle$. The procedure and its strengths and pitfalls are discussed at length in the above papers, and are summarized in the "Appendix".

For the LIM presented in this paper, we chose the model state vector $\mathbf{x}$

$$\mathbf{x} = \begin{bmatrix} \mathbf{T}_O \\ \mathbf{Z}_{20} \\ \boldsymbol{\tau}_x \end{bmatrix}$$

where $\mathbf{T}_O$ is anomalous SST, $\mathbf{Z}_{20}$ is anomalous 20°C isotherm depth, and $\tau_x$ is anomalous surface zonal wind stress, determined from 42 years (1959–2000) of 3-month running mean data. Monthly means of $\mathbf{T}_O$ and $\mathbf{Z}_{20}$ can also be used to construct LIMs, but this time average is problematic when an atmospheric variable is included because of some aliasing of the Madden–Julian oscillation (MJO) signal. SST was obtained from the Hadley Sea Ice and Sea Surface Temperature analysis (HadISST; Rayner et al. 2003), 20°C isotherm depth from the SODA dataset (Carton and Giese 2008), and wind stress from the NCEP Reanalysis (Kalnay et al. 1996). Data were averaged into 2° latitude × 5° longitude gridboxes. Anomalies, determined by removing the long-term monthly mean, were projected onto their leading Empirical Orthogonal Functions (EOFs) determined for the region 25°S–25°N. Prior to computing EOFs, each field was normalized by its domain-averaged climatological root-mean-square amplitude. The leading 13/7/3 EOFs of $\mathbf{T}_o/\mathbf{Z}_{20}/\tau_x$ were retained, explaining about 83/36/32 percent of the variability of their respective fields. Locally, however, variance explained can be considerably higher (or lower), as seen in Fig. 1, which shows the (untruncated) variance of $\mathbf{T}_o$, $\mathbf{Z}_{20}$, and $\tau_x$, and the local fraction of variance explained by the truncated EOF bases. The time-varying coefficients of these EOFs, i.e., the principal components (PCs), define the 23-component state vector $\mathbf{x}$. For purposes of comparison, corresponding SST-only LIMs (or "SST-LIMs") were constructed by representing $\mathbf{x}$ with only $\mathbf{T}_O$, using either 13 or 23 PCs (SST13-LIM and SST23-LIM). Throughout the remainder of this paper,

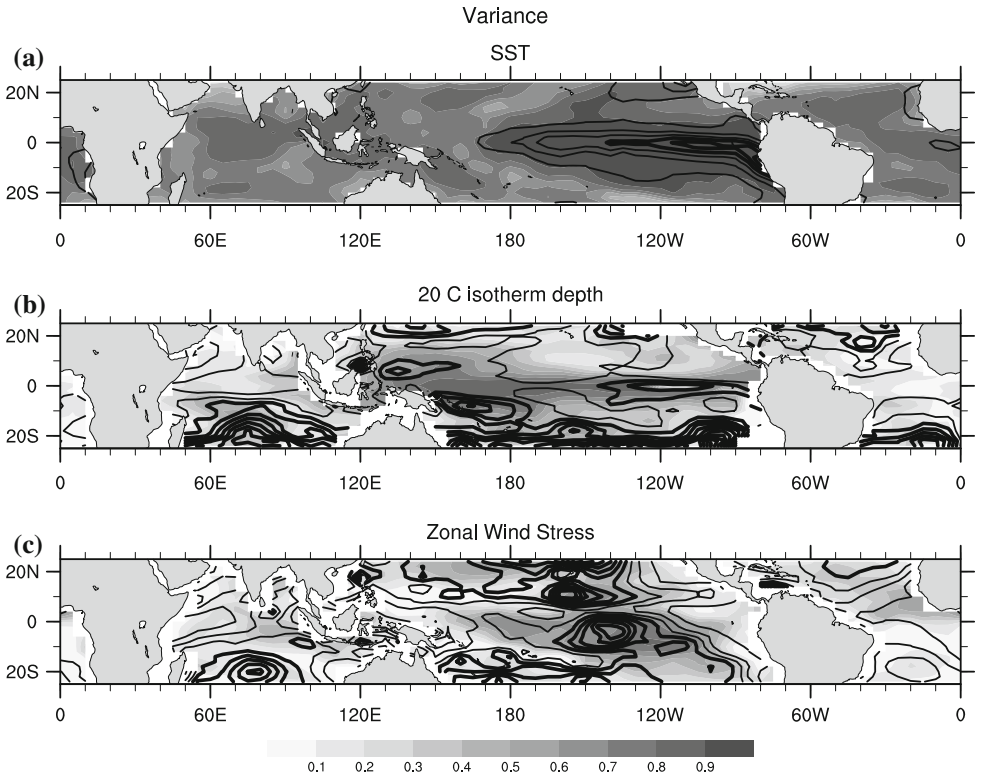


**Fig. 1** Variance (*contours*) and fraction of local variance explained by EOF truncation (*gray shading*) for variables used in the LIM. *Top* SST ($\mathbf{T}_O$); contour interval 0.25 $K^2$. *2nd row* 20°C isotherm depth ($\mathbf{Z}_{20}$); contour interval 100 $m^2$. *Bottom* zonal wind stress ($\boldsymbol{\tau}_x$); contour interval $5 \times 10^{-5}$ $N^2m^{-4}$. *Thicker contours* indicate larger values, starting at and increasing from 1 $K^2$, 300 $m^2$, and $20 \times 10^{-5}$ $N^2m^{-4}$, respectively

"ocean LIM" will denote the LIM constructed from the extended $\mathbf{T}_o$-$\mathbf{Z}_{20}$-$\tau_x$ state vector.

To use the LIM for diagnosis, we must first validate it by determining whether the effectively linear approximation of (2) holds, through a series of tests (PS95) that include a comparison of the observed evolution of individual events to their prediction by the LIM and a more general comparison of observed anomaly evolution statistics to those linearly predicted by (2). The results of these tests (see "Appendix") basically agree with both PS95 and subsequent studies that constructed linear empirical models from more limited ocean datasets (Johnson et al. 2000; Xue et al. 2000). The additional variables in the model state vector result in a LIM that passes the above tests with even greater fidelity than the SST-LIMs, although this improvement is not significant until the time scale is longer than a few seasons.

As an example, Fig. 2 shows that the SST23-LIM and the ocean LIM, constructed from the same number of degrees of freedom, have similar Niño 3.4 SST forecast skill for leads up to about 8 months. That is, replacing higher order SST PCs with subsurface information in the LIM does not improve skill for shorter leads, also earlier noted by C. Penland (personal communication) and Johnson et al. (2000). For longer leads, however, the ocean LIM has higher skill than SST-LIMs of any EOF truncation. [In fact, for forecast leads over 1 year, more than 13 EOFs degrades SST-LIM skill.] As a result, if a "good" forecast has anomaly correlation skill greater than or equal to 0.6, then the ocean LIM improves Niño 3.4 forecast skill by about 3–4 months. A similar result also obtains over the Indian Ocean (not shown).

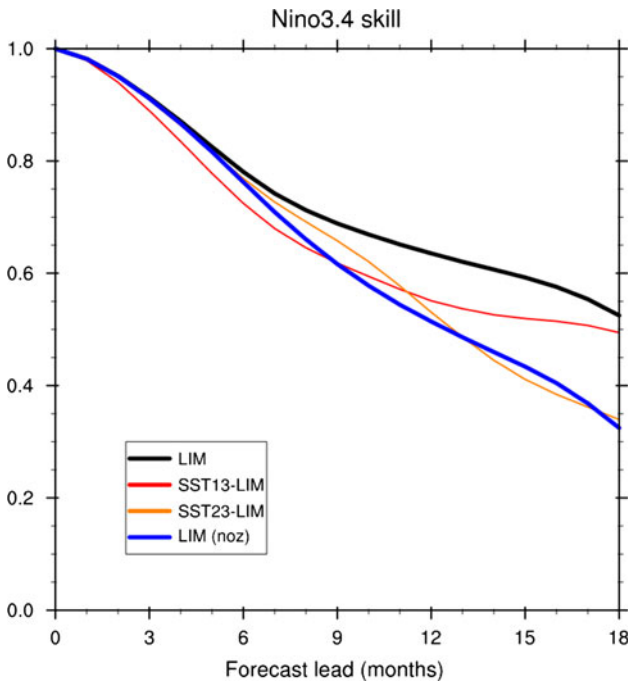


**Fig. 2** Cross-validated forecast skill (1959–2000) of Niño3.4, defined as the area averaged SST in the region 5°S–5°N, 170°W–120°W, for forecasts made by (*black*) the LIM, (*blue*) the version of the LIM where thermocline interactions are removed (by setting $\mathbf{L}_{ZT} = \mathbf{L}_{TZ} = \mathbf{L}_{\tau Z} = \mathbf{L}_{Z\tau} = 0$ in $\mathbf{L}$), $\mathbf{L}_{\text{noz}}$, (*red*) the SST13-LIM, and (*orange*) the SST23-LIM

## 3 Processes within the dynamical operator

The dynamics of different tropical oceanic processes may be investigated by rewriting (2) as

$$\frac{d}{dt}\begin{bmatrix}\mathbf{T}_o\\ \mathbf{Z}_{20}\\ \boldsymbol{\tau}_x\end{bmatrix} = \begin{bmatrix}\mathbf{L}_{TT} & \mathbf{L}_{TZ} & \mathbf{L}_{T\tau}\\ \mathbf{L}_{ZT} & \mathbf{L}_{ZZ} & \mathbf{L}_{Z\tau}\\ \mathbf{L}_{\tau T} & \mathbf{L}_{\tau Z} & \mathbf{L}_{\tau\tau}\end{bmatrix}\begin{bmatrix}\mathbf{T}_o\\ \mathbf{Z}_{20}\\ \boldsymbol{\tau}_x\end{bmatrix} + \begin{bmatrix}\boldsymbol{\xi}_T\\ \boldsymbol{\xi}_Z\\ \boldsymbol{\xi}_\tau\end{bmatrix}. \quad (4)$$

Note that $\mathbf{L}_{TT}$ is distinct from the SST-LIM linear operator, which implicitly includes linear diagnostic relationships between $\mathbf{Z}_{20}$ and $\boldsymbol{\tau}_x$ and $\mathbf{T}_O$. By explicitly separating out the effects of the other two variables on $\mathbf{T}_O$ and vice versa, we can use (4) to identify $\mathbf{L}_{TT}$ with surface ocean processes, $\mathbf{L}_{ZZ}$ with internal ocean processes, $\mathbf{L}_{\tau\tau}$ with surface atmospheric processes, and the off-diagonal submatrices with coupling. Of course, these operators each implicitly retain the influence of variables not included in $\mathbf{x}$, especially considering the more severe EOF truncation of $\mathbf{Z}_{20}$ and $\boldsymbol{\tau}_x$, and to the extent that the terms are related to the same unspecified variables they may not be entirely independent. Also, processes with time scales much shorter than the 3-month averaging period used to construct $\mathbf{x}$ (notably, Kelvin waves) are still represented by a diagnostic relationship plus some noise.

With these important caveats in mind, we "remove" effects of thermocline variability by constructing a new operator $\mathbf{L}_{\text{noz}}$, where we set $\mathbf{L}_{ZT} = \mathbf{L}_{TZ} = \mathbf{L}_{\tau Z} = \mathbf{L}_{Z\tau} = 0$ in $\mathbf{L}$. First, to evaluate the impact of thermocline anomalies upon SST forecasts we repeated all forecasts using $\mathbf{L}_{\text{noz}}$ as the forecast operator, finding that while thermocline interaction with SSTs has a small effect on SST skill for forecast leads shorter than about 6 months, it has an increasing impact for longer-lead forecasts, reaching the .05 significance level for leads greater than a year (Fig. 2). How thermocline interactions affect SST variability statistics can likewise be seen by determining a new covariance $\mathbf{C}_{\text{noz}}(0)$ from (3) using $\mathbf{L}_{\text{noz}}$ and the original noise covariance $\mathbf{Q}$, and predicted $\tau$-lag covariance from $\mathbf{C}_{\text{noz}}(\tau) = exp(\mathbf{L}_{\text{noz}}\tau)\mathbf{C}_{\text{noz}}(0)$. The results are compared to the corresponding fields determined from the full LIM in Fig. 3; note that SST variance from the LIM is correct by construction (apart from a difference due to the EOF truncation), and that the 9-month lag-covariance predicted by the LIM is an excellent match to observations (see

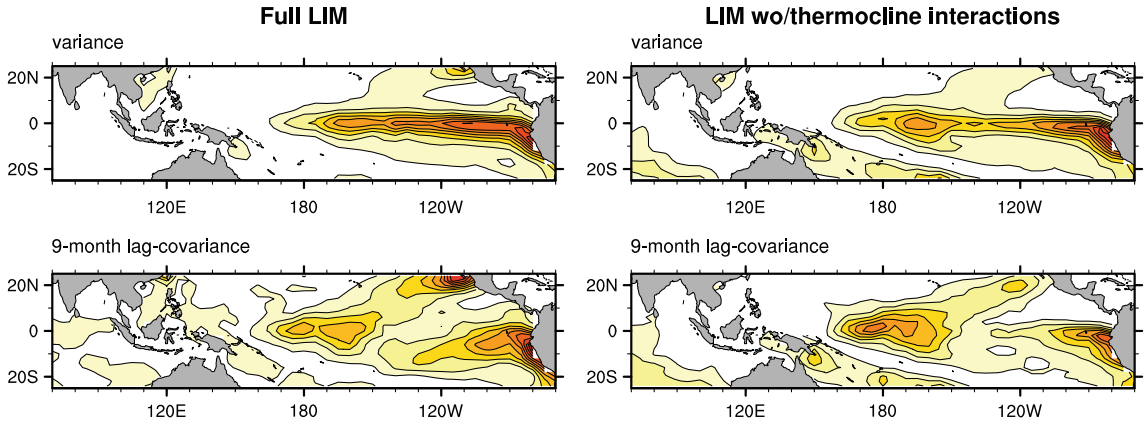


**Fig. 3** $\mathbf{T}_O$ variance and lag-covariance determined from $\mathbf{L}$ (*left panels*) and from $\mathbf{L}_{\text{noz}}$, where the thermocline interactions are removed by setting $\mathbf{L}_{ZT} = \mathbf{L}_{TZ} = \mathbf{L}_{\tau Z} = \mathbf{L}_{Z\tau} = 0$ in $\mathbf{L}$ (*right panels*). *Top* variance; contour interval is 0.16 $K^2$. *Bottom* 9-month lag-covariance; contour interval is 0.04 $K^2$

Fig. 13). Removing the thermocline interactions leads to a small increase of central equatorial Pacific $\mathbf{T}_O$ variability and simultaneously a pronounced decrease of eastern Pacific variability (cf. top panels of Fig. 3), while central Pacific variability becomes more persistent over a 9-month period (cf. bottom panels of Fig. 3). In essence, ENSO events have characteristics more similar to "Modoki" or Central Pacific type events (e.g., Ashok et al. 2007; Kao and Yu 2009) when thermocline interactions are removed from $\mathbf{L}$. Additionally, overall decadal variability is reduced (not shown).

The covariance budget of 3-months running mean SST anomalies, averaged over the annual cycle and derived from (3) and (4) as

$$\underbrace{\left[\mathbf{L}_{TT}\mathbf{C}_{TT} + \mathbf{C}_{TT}\mathbf{L}_{TT}^{T}\right]}_{\text{SST}} + \underbrace{\left[\mathbf{L}_{TZ}\mathbf{C}_{ZT} + \mathbf{C}_{TZ}\mathbf{L}_{TZ}^{T}\right]}_{\text{Thermocline interaction}} + \underbrace{\left[\mathbf{L}_{T\tau}\mathbf{C}_{\tau T} + \mathbf{C}_{T\tau}\mathbf{L}_{T\tau}^{T}\right]}_{\text{Wind stress}} + \underbrace{\mathbf{Q}_{TT}}_{\text{Stochastic}} = \mathbf{O}, \quad (5)$$

where $\mathbf{C}_{xy} = \langle \mathbf{x}\mathbf{y}^T \rangle$, is another gauge of anomalous thermocline and wind stress influences on SST variability. The diagonal elements in each bracketed term in (5) may be interpreted as their contribution to the local $\mathbf{T}_O$ variance tendency. Note that although these terms relate to the maintenance of local variance, they contain non-local effects: $\mathbf{L}_{TT}$ and $\mathbf{C}_{TT}$ are not diagonal matrices, and thermocline and zonal wind stress variability contribute to SST variability. In principle this "inverse" method of determining the interannual heat budget should agree with a more traditional "forward" approach in which all terms in the surface-layer heat budget are explicitly calculated (e.g., Wang and McPhaden 2000; Kang et al. 2001).

It is interesting to break down the bracketed SST term further into "local" (Fig. 4a) and "non-local" (Fig. 4b) contributions. The local contribution, twice the product of the diagonal elements of $\mathbf{L}_{TT}$ with the corresponding diagonal elements of $\mathbf{C}_{TT}$, results in a dissipative term of the form $-2\langle (T_O)^2 \rangle \big/ \tau_d$ (where $\tau_d$ is a local decay time scale) with stronger damping in the eastern and equatorial central Pacific ($\tau_d \sim$ 6–7 months) than in the western and subtropical central Pacific ($\tau_d \sim$ 10–15 months). This regional variation of damping includes effects both of local SST-related variations of radiative and surface heat fluxes (e.g., Kessler and McPhaden 1995; Wang and McPhaden 2000) and of vertical mixing related to stronger mean upwelling in the east. The non-local term is the difference between the total term (not shown) and the local term, representing how all other values of $\mathbf{T}_O$ impact local $\mathbf{T}_O$ variability, such as through advection and/or wave propagation. This term also generally acts to damp equatorial SST variance as might be expected from the mean advection of thermal anomalies (Jin et al. 2006) and meridional transport due to tropical instability waves (Wang and McPhaden 2000; Jochum et al. 2007), except in the central Pacific and around 5–10°N where it acts to generate SST variance, consistent with the downgradient transport of SST variance (cf. Fig. 1) by high frequency eddy fluxes (Wang and McPhaden 2000) and mean equatorial zonal and off-equatorial meridional currents (Kang et al. 2001).

The primary balance in Fig. 4 is between thermocline interactions (Fig. 4c) and the SST dynamical terms. The thermocline term, which is positive except in the central Pacific, includes the net SST variance generation by local thermocline and upwelling feedbacks, and also non-local zonal advective feedback (Picaut et al. 1996; Jin and An 1999) due to the dependence of anomalous zonal current on the anomalous equatorial thermocline gradient. $\mathbf{T}_O$ noise effects are very small (Fig. 4e) except in a narrow equatorial band where tropical instability waves are active (Chelton et al. 2000); in fact, atmospheric noise forcing has a greater but indirect effect upon SSTs primarily by maintaining zonal wind stress variability (not shown),

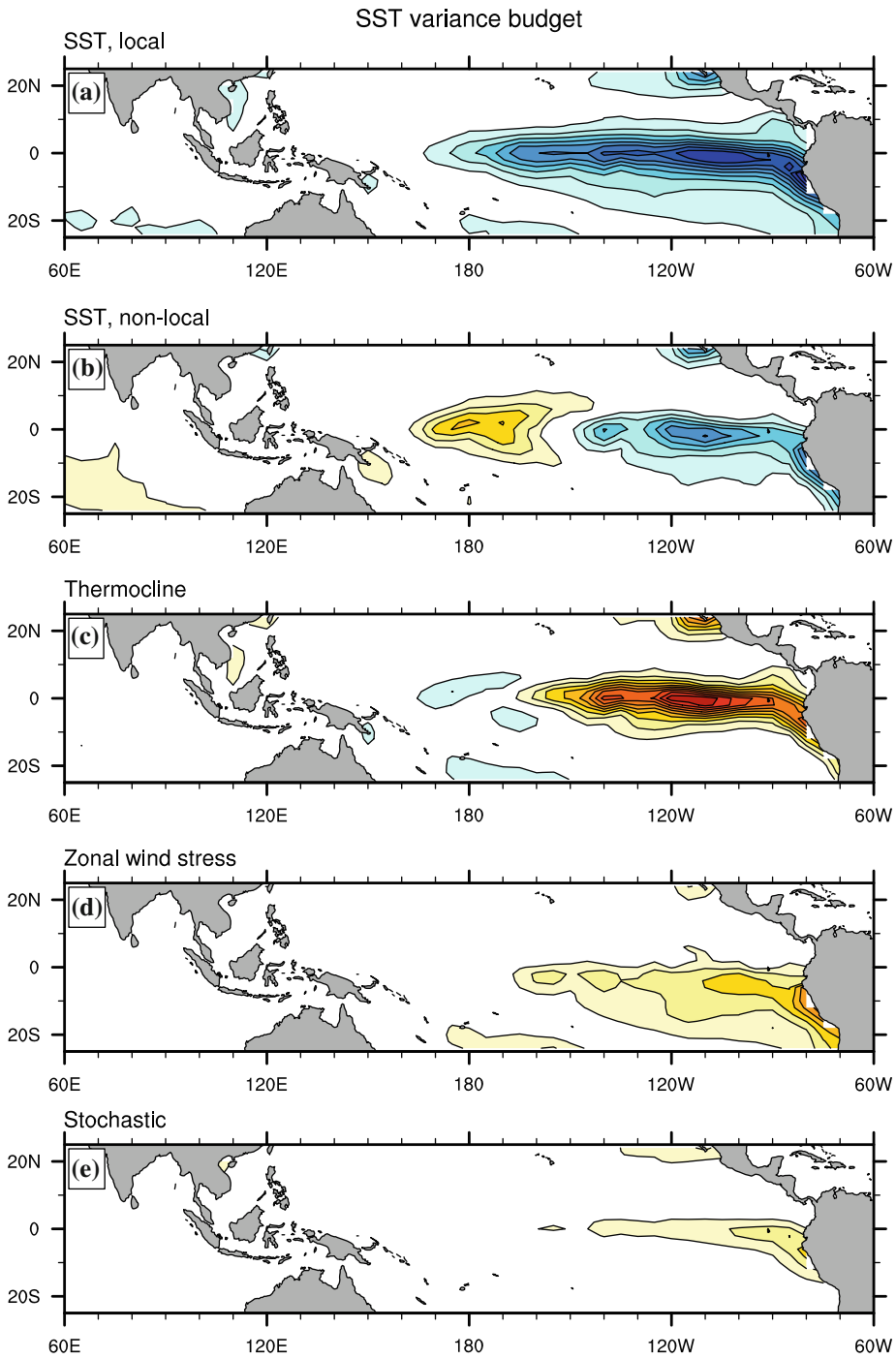


**Fig. 4** The local variance budget for SST (see text for description of terms). Note that the panels sum to zero. Contour and shading interval is 0.04 $K^2$ $month^{-1}$; positive values are *red/yellow* and negative values are *blue*, with the zero contour removed for clarity

which in turn forces thermocline variability. While there is a weaker direct zonal wind stress forcing term (Fig. 4d), it might also implicitly represent other anomalous surface wind effects on latent and sensible heat fluxes, and even radiative fluxes due to cloudiness. Note that advective feedbacks in the east-central Pacific appear to mostly be contained within the thermocline term and do not represent direct effects from the wind stress forcing term as in the "wind stress feedback" (Burgers and van Oldenborgh 2003).

## 4 Optimal evolution of ENSO

### 4.1 Maximum amplification of SST anomalies in SST and ocean LIMs

PS95 showed that the "optimal" initial condition for maximum amplification of tropical SST anomalies, obtained via a singular vector decomposition (SVD) of the system propagator $\mathbf{G}(\tau)$ under the L2 norm (i.e., domain-mean square amplitude) of $\mathbf{T}_O$ (e.g., Farrell 1988; PS95), is

also the most relevant initial condition for ENSO development. The SVD analysis yields a dominant pair $\mathbf{u}_1$,$\mathbf{v}_1$ of normalized singular vectors and maximum singular value $\lambda_1$, such that the initial condition $\mathbf{v}_1$ leads to the anomaly $\mathbf{G}$ $\mathbf{v}_1 = \lambda_1 \mathbf{u}_1$ at time $t = \tau$. The maximum possible anomaly growth factor $\lambda_1{}^2(\tau)$, sometimes called a "maximum amplification" (MA) curve (PS95), is displayed in Fig. 5a for each LIM. For consistent comparison, the L2 norm is defined in the space of the leading 13 EOFs of $\mathbf{T}_O$ in all three cases.

The shape and maximum (at $\tau \sim$ 8–9 months) of the MA curve for each SST-LIM is consistent with earlier studies including PS95. Although in principle an optimal initial condition for one value of $\tau$ does not have to be optimal for another value of $\tau$, in reality PS95 found relatively little difference in the leading optimal structure for a wide range of $\tau$. This result is also true for our SST-LIMs; for example the pattern correlation of $\mathbf{v}_1$(9 months) with $\mathbf{v}_1(\tau)$ is greater than 0.9 for $\tau$ ranging from about 4 to 14 months (Fig. 5b). That is, for all practical purposes the SST-LIM has only one leading optimal initial structure.

The shape of the MA curve for the ocean LIM is quite different, however, with a secondary maximum at $\tau \sim$ 17–18 months and a fairly flat curve between the two maxima. Moreover, these two maxima represent different optimal structures, with an 0.35 pattern correlation between $\mathbf{v}_1$(9 months) and $\mathbf{v}_1$(18 months) (Fig. 5b). However, although the initial states are different, the end results are similar, with both optimal initial conditions evolving into a large central and eastern Pacific SST anomaly (see below discussion). The increased maximum amplification for the ocean LIM implies increased predictability (Newman et al. 2003) for forecast leads greater than about nine months, consistent with the improved skill in Fig. 2.

### 4.2 Evolution of 9- and 18-month leading optimal structures

To explore the differences between these two leading optimal structures, we integrate (2) forwards from either initial condition $\mathbf{v}_1$(9), the optimal structure for growth over a $\tau = 9$ month interval (Fig. 6a), or from initial condition $\mathbf{v}_1$(18), the optimal structure for growth over a $\tau = 18$ month interval (Fig. 7a). For ease of comparison, both initial conditions are chosen with sign leading to a warm event, but in this linear model all signs can be reversed. The evolution from $\mathbf{v}_1$(9) [$\mathbf{v}_1$(18)] over the next 36 months is displayed in Figs. 6b–d [7b–d]. Forcing by $\mathbf{Z}_{20}$ ($\boldsymbol{\tau}_x$) and nonlocal terms in the $\mathbf{T}_O$ ($\mathbf{Z}_{20}$) tendency [from (4)] are indicated in Figs. 6e, f and 7e, f by shading (contours); these terms are evaluated at the equator but can contain effects driven from other latitudes.

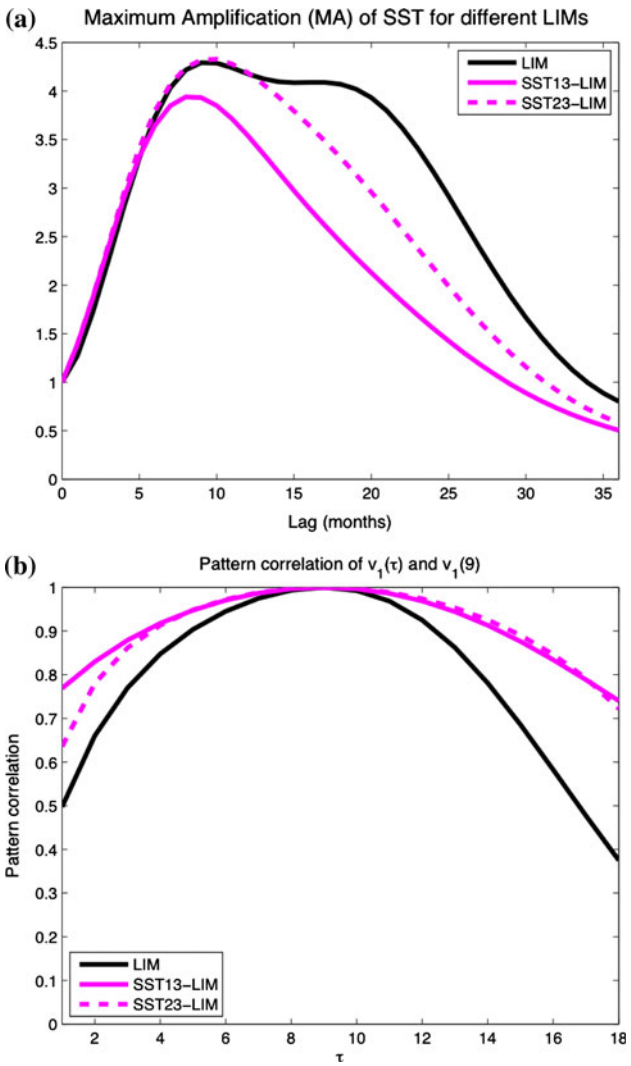


**Fig. 5** **a** Comparison of maximum amplification (*MA*) curves, defined as $[\lambda_1(\tau)]^2$ determined by the SVD of $\mathbf{G}(\tau)$ under the L2 norm of $\mathbf{T}_O$, for the LIM and the SST-LIMs. **b** Pattern correlation of $\mathbf{v}_1(9)$ with $\mathbf{v}_1(\tau)$ for the ocean LIM and the SST-LIMs

The initial anomaly for the shorter 9-month optimization time (Fig. 6a) has (1) SST that is broadly similar to the corresponding SST-LIM optimal initial condition (not shown, but see PS95) except in the central equatorial Pacific where it has half the amplitude, (2) a positive basin-wide thermocline anomaly centered in the east central Pacific similar to that often viewed as an important ENSO precursor (e.g., Wyrkti 1985; Jin 1997; Meinen and McPhaden 2000; Clarke and Van Gorder 2003; McPhaden et al. 2006), and (3) a westerly wind stress anomaly at roughly the same longitude with off-equatorial maxima. Optimal SST anomaly amplification over the next 9 months is consistent with Bjerknes (1969) feedback: the thermocline term $\mathbf{L}_{TZ}\mathbf{Z}_{20}$ (Fig. 6f, shading) forces rapid amplification of anomalously warm SST in the east Pacific, which in turn (along with the initial SST anomaly) drives

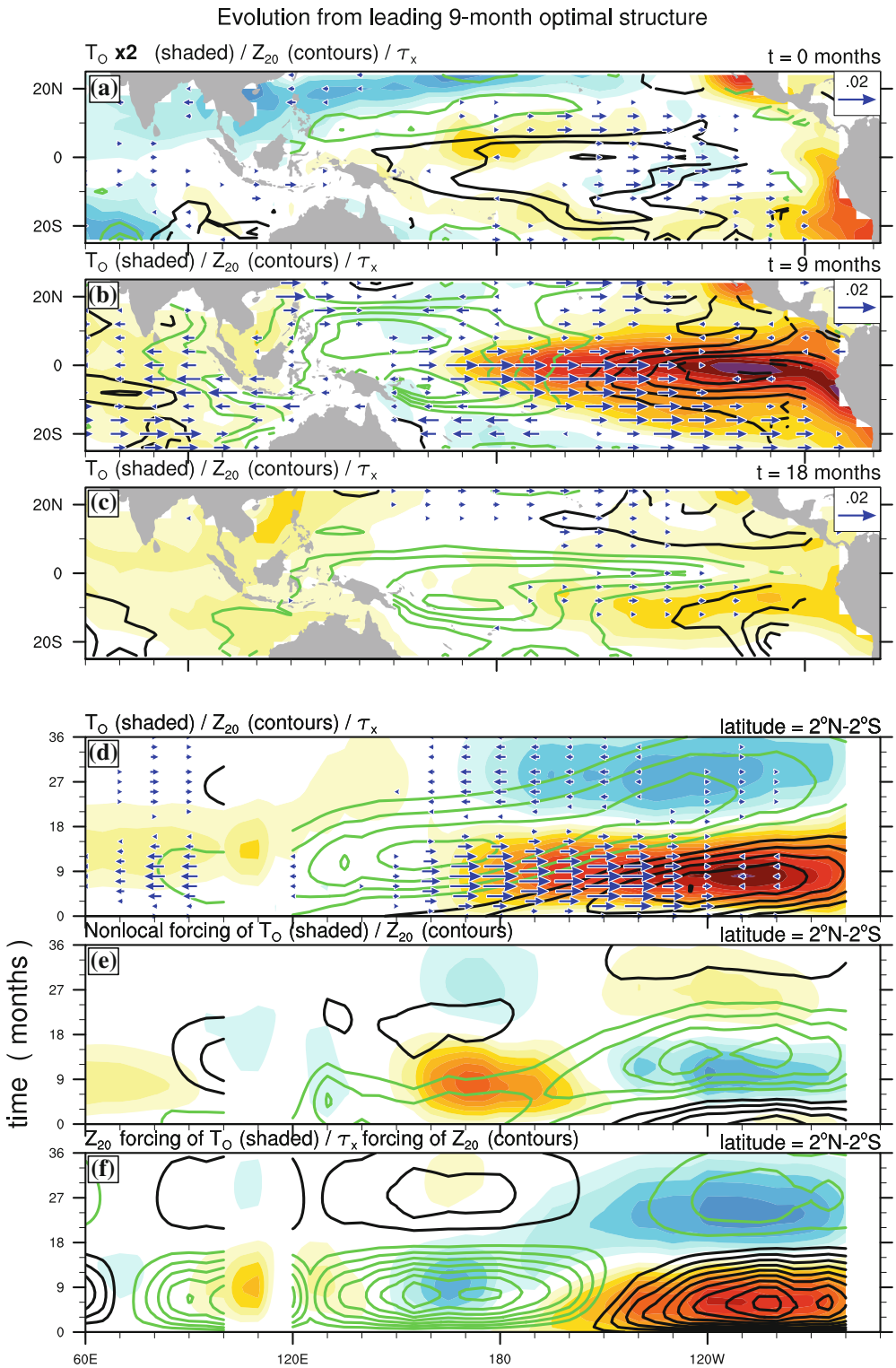


**Fig. 6** Evolution of the optimal initial condition for amplification of SST anomalies in the LIM over a 9-month interval. **a** Initial state, and the evolved states, **b** 9 and **c** 18 months later. **d** Time-longitude cross-section of the LIM evolution averaged between 2°S–2°N. $\mathbf{T}_O$ is indicated by shading (contour interval 0.175 K), $\mathbf{Z}_{20}$ by contours (contour interval 6 m), and $\tau_x$ by blue vectors (scaled by the reference vector 0.02 N m$^{-2}$, with values below 0.002 N m$^{-2}$ removed for clarity). **e** and **f** Time-longitude cross-section of selected forcing terms in the tendency equations for $\mathbf{T}_O$ (*shading* interval 0.05 K/month) and $\mathbf{Z}_{20}$ (*contours* interval 0.5 m/month): **e** Nonlocal terms; **f** $\mathbf{Z}_{20}$ forcing of $\mathbf{T}_O$ tendency and $\tau_x$ forcing of $\mathbf{Z}_{20}$ tendency. Note that sign in all panels can be flipped, but that black contours and red shading are of one sign, and green contours and blue shading are the other sign. Amplitudes are also arbitrary, but are scaled by a single factor here to have representative values. Note that $\mathbf{T}_O$ is doubled in **a** to be more visible

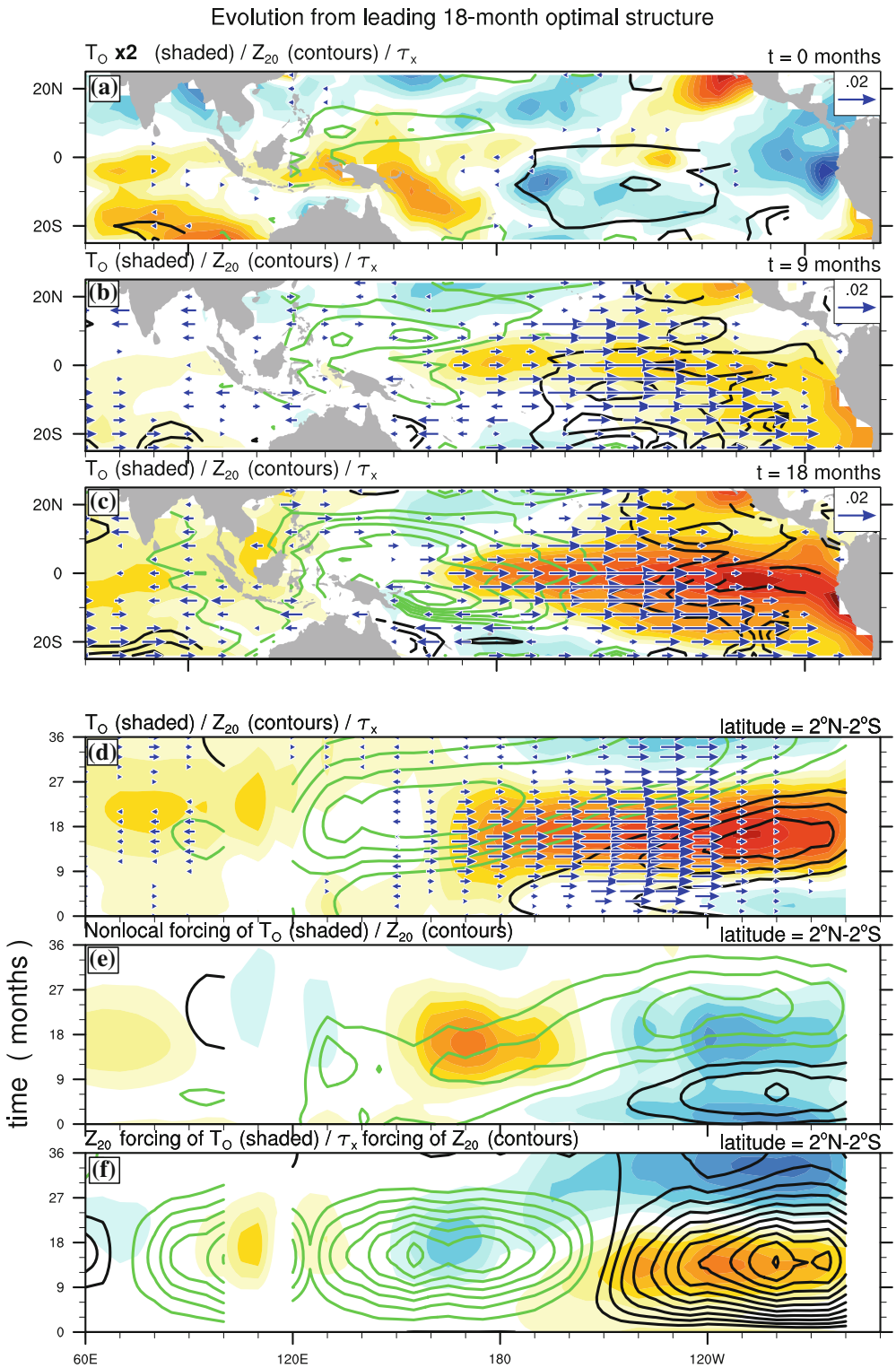


**Fig. 7** Same as Fig. 6 but for the evolution of the optimal initial condition for amplification of SST anomalies in the LIM over an 18-month interval

intensification and westward expansion of anomalous equatorial westerly wind stress (not shown) that drives (Fig. 6f, contours) the $\mathbf{Z}_{20}$ anomaly to intensify and move eastward (Fig. 6d), and so on. Maximum $\mathbf{L}_{TZ}\mathbf{Z}_{20}$ forcing of SST occurs near the $\mathbf{Z}_{20}$ maximum at about 110°W; a secondary $\mathbf{L}_{TZ}\mathbf{Z}_{20}$ maximum near 140°W might represent advective feedback since it is located near large zonal gradients of both mean SST and anomalous $\mathbf{Z}_{20}$.

Optimal growth is broadly consistent with the "discharge" (Jin 1997) of anomalous equatorial heat content (as represented by equatorial $\mathbf{Z}_{20}$ integrated in the region 2°S–2°N, 120°E–85°W), which after increasing the first 2 months decreases to zero at the peak SST time of $t = 9$ months (not shown, but cf. Fig. 6d). The initial increase of equatorial heat content coincides with initial off-equatorial wind stress that weakens after about 2 months, when the initial SST anomaly has evolved into its ENSO pattern. Zonal wind stress forcing of $\mathbf{Z}_{20}$ in the east Pacific and $\mathbf{Z}_{20}$ forcing of SST both peak before the SST anomaly, at $t = 7$ months, at which time zonal wind stress forcing of equatorial heat content also becomes negative. SST anomaly amplification continues for two more months, until the (negative) total SST term grows to balance the thermocline term (e.g., Jin 1997).

Equatorial heat content decreases during anomalous SST amplification since as the positive $\mathbf{Z}_{20}$ anomaly grows and moves eastward, zonal wind stress forcing $\mathbf{L}_{Z\tau}\boldsymbol{\tau}_x$ (Fig. 6f) generates a negative $\mathbf{Z}_{20}$ anomaly in the west Pacific (Fig. 6d), with minima both north and south of the equator (Fig. 6b) consistent with Ekman pumping from the curl of easterly wind stress anomalies centered at about 15°N and 15°S. The non-local portion of $\mathbf{L}_{ZZ}\mathbf{Z}_{20}$ forcing also enhances the west Pacific negative $\mathbf{Z}_{20}$ anomaly, but is generally much weaker (Fig. 6e) except near the secondary $\mathbf{Z}_{20}$ minimum located around 130°E and 6°N (Fig. 6b). SST anomaly decay accelerates when the negative western Pacific thermocline anomaly finally expands eastwards along the equator (Fig. 6d), reaching South America by about $t = 18$ months (Fig. 6c) and causing equatorial SSTs to become cool even as subtropical SSTs remain warm. This implies that ocean wave propagation processes may be secondary to direct wind forcing in discharging the heat content during optimal growth, and is consistent with the suggestion that wind forcing in the west Pacific was of primary importance in initiating the termination of the 1997/98 El Nino (McPhaden 1999; Boulanger and Menkes 2001).

Note also that while the growing negative thermocline anomaly drives cooler SSTs via $\mathbf{L}_{TZ}\mathbf{Z}_{20}$, it is offset in the central equatorial Pacific by the non-local SST term (Fig. 6e, shading). This is consistent with equatorial SST tendencies during the 1997/98 El Niño just east of the dateline, where negative vertical advection was offset by positive advection by the mean zonal current (Boulanger and Menkes 2001).

Repeating this integration by first setting the initial $\mathbf{Z}_{20}$ and $\tau_x$ anomalies to zero results in similar SST evolution but with peak amplitude reduced by almost half (not shown), demonstrating that the initial thermocline and SST anomalies contribute roughly equally to SST amplification. If the initial conditions in Fig. 6a evolve using $\mathbf{L}_{noz}$ instead (not shown), the peak anomaly is not only weaker but also more localized, centered much closer to the dateline, similar to the results shown in Fig. 3; also, as it grows the SST anomaly propagates westward from its initial location along the South American coast.

In contrast to $\mathbf{v}_1(9)$, the initial SST portion of $\mathbf{v}_1(18)$ (Fig. 7a) is relatively much stronger than the initial thermocline anomaly. Notably, the southwest Pacific and Indian Ocean SST anomalies persist for several months (e.g., Fig. 7d) while off-equatorial westerly wind stress anomalies rapidly intensify in both hemispheres of the eastern Pacific (not shown) and deepen the equatorial thermocline there (Fig. 7d, f). [Adding additional $\tau_x$ PCs to $\mathbf{x}$ to capture more west Pacific variance does not change this picture.] As the anomaly evolves, it never becomes similar to $\mathbf{v}_1(9)$, with perhaps the closest correspondence occurring at about $t = 6$ months primarily for central and eastern Pacific SST (cf. $t = 0$ in Fig. 6d with $t = 6$ months in Fig. 7d). Subsequent anomaly evolution ($t = 6$–18 months) is only broadly similar to that shown for the 9-month optimal and moreover extends over a longer period consistent with more persistent wind stress forcing of $\mathbf{Z}_{20}$ (Fig. 7f).

Most of the evolution from $\mathbf{v}_1(18)$ is due to the initial SST, in contrast to $\mathbf{v}_1(9)$; a second integration with the initial $\mathbf{Z}_{20}$ anomaly set to zero reduces peak SST by less than 15% (not shown). However, the evolution of $\mathbf{Z}_{20}$ remains essential, since the thermocline term ultimately drives most of the eastern Pacific SST amplification (Fig. 7f). Alternatively, when $\mathbf{v}_1(18)$ is evolved with the $\mathbf{L}_{noz}$ operator, there is very little SST growth.

When considering optimal structures, it is important whether this potential growth is actually observed to occur as predicted by the LIM. In addition, do the optimal structures shown in Figs. 6 and 7 improve upon those from the SST-LIM; that is, are they more relevant to the observed system? To answer this, following PS95 and others we compared the projection of observed anomalies on the initial conditions shown in Figs. 6a and 7a to the projection of observed SST anomalies upon the predicted evolved structure 9 (Fig. 6b) and 18 (Fig. 7c) months later, respectively, as shown in Fig. 8a ($\tau = 9$ months) and Fig. 8c ($\tau = 18$ months). This analysis was repeated using the corresponding SST13-LIM optimal structures (not shown); results are shown in Fig. 8b ($\tau = 9$ months) and d ($\tau = 18$ months). The increased linear correlation and reduced scatter about the least squares lines in the panels on the left relative to those on the right indicate that

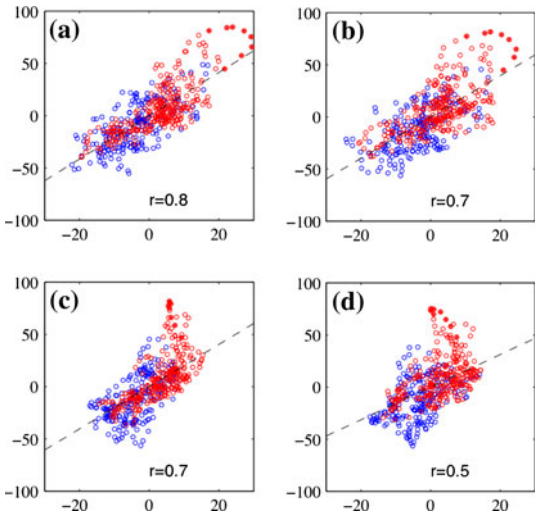


**Fig. 8** Projection of observations upon the optimal initial condition for amplification of SST anomalies in the LIM over a 9-month interval, versus the optimal evolved SST state 9 months later, in **a** the ocean LIM and **b** the SST13-LIM. Projection of observations upon the optimal initial condition for amplification of SST anomalies in the LIM over a 18-month interval, versus the optimal evolved SST state 18 months later, in **c** the ocean LIM and **d** the SST13-LIM. The dashed lines in all four panels represent the expected evolution, with slopes $\lambda_1(\tau)$. Blue circles represent the 1959–1976 period, and red circles the 1977–2000 period. *Red dots* represent the 1997–1998 El Niño event

optimal SST growth is better captured by the ocean LIM. Of course, some scatter about the least squares line is still expected due to the noise $\mathbf{F}_s$. The slopes of the dashed lines should correspond to the amplification factors $\lambda_1(\tau)$, and in fact they are quite close. This also serves as additional evidence that $\mathbf{L}$ is independent of $\tau_0$, since these calculations have all been made for intervals $\tau \gg \tau_0 = 3$ months.

Stratifying into pre- (blue) and post- (red) 1976/1977 periods shows no significant long-term change in the relationship (i.e., change in slope) for the ocean LIM. Observed decadal changes in variability could then be represented by random variation of noise forcing projecting differently upon the optimal initial conditions, as opposed to a decadal change in underlying dynamics. Finally, while the 1997/1998 El Niño event is somewhat better simulated with the ocean LIM than the SST-LIM, it remains a potential outlier and may represent an event for which we cannot exclude a role for predictable nonlinearity.

### 4.3 Eigenmodes contributing to optimal structures

For the SST-LIM, the leading optimal structure can be well approximated by just three nonnormal eigenmodes of $\mathbf{L}$ (PS95; Penland and Matrosova 2006). We find this too for $\mathbf{v}_1(9)$ and its evolution into $\mathbf{u}_1(9)$, where again all three are propagating eigenmodes (that is, complex conjugate pairs with complex eigenvalues each representing period and

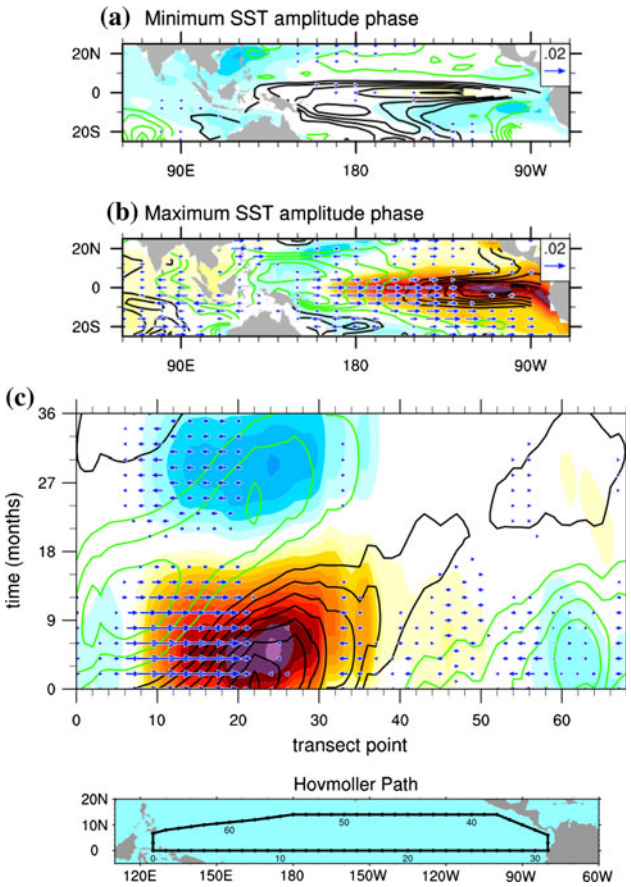


**Fig. 9** The third least damped eigenmode of $\mathbf{L}$, with period of 4.0 years and e-folding time of 2.1 years. **a** Minimum and **b** Maximum SST phases, defined from the root-mean-squared domain SST anomaly. **c** Time-space cross-section of the eigenmode evolution, starting from the phase shown in Fig. 11a, along the transect indicated by **d**. $\mathbf{T}_O$ is indicated by *shading* (contour interval 0.175 K), $\mathbf{Z}_{20}$ by *contours* (contour interval 6 m), and $\tau_x$ by *blue vectors* (scaled by the reference vector 0.02 N m$^{-2}$). Note that sign in all panels is arbitrary, but that *black contours* and *red shading* are of one sign, and *green contours* and *blue shading* are the other sign. Amplitudes are also arbitrary, but are scaled by a single factor here to have representative values

e-folding time) with periods of about 4 years ('4-year eigenmode'; Fig. 9), 1.9 years ('2-year eigenmode'; Fig. 10), and 21 years ('bidecadal eigenmode'; not shown). [The complete set of 23 eigenmodes $\boldsymbol{\varphi}_j$ and corresponding complex eigenvalues $\sigma_j$ are determined from the eigenanalysis $\mathbf{L}\boldsymbol{\varphi}_j = \sigma_j\boldsymbol{\varphi}_j$.] While the SST components and periods are similar to the corresponding SST-LIM eigenmodes (not shown, but see Penland and Matrosova 2006), the e-folding times of both the 4-year and bidecadal eigenmodes are about double those of their SST-LIM counterparts; the 2-year eigenmode e-folding time is unchanged.

The 4- and 2-year eigenmodes are not orthogonal (with a minimum 40° angle between them) as is evident in their similar structures for SST minimum (Figs. 9a, 10a) and

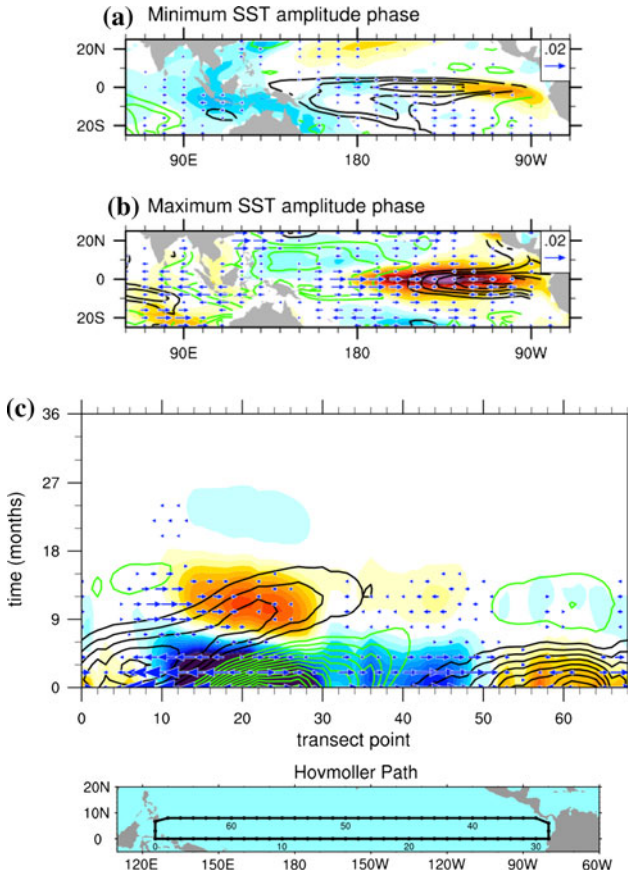


**Fig. 10** The eighth least damped eigenmode of L with period of 1.9 years and e-folding time of 0.65 years. **a** Minimum and **b** maximum SST phases, defined from the root-mean-squared domain SST anomaly. **c** Time-space cross-section of the eigenmode evolution, starting from the phase shown in Fig. 11b, along the transect indicated by **d**. Plotting conventions as in Fig. 9 except all values in Fig. 10c are scaled by a factor of 3; note that the transect points here are at the same longitudes as the transect points in Fig. 9

SST maximum (Figs. 9b, 10b) phases. Consequently, an initial state quite similar to $\mathbf{v}_1(9)$ can be simply constructed by taking the 4-year eigenmode in a phase about 9 months prior to its SST maximum (Fig. 11a), and then mostly "covering it up" by adding the 2-year eigenmode with amplitude and phase corresponding to Fig. 11c, resulting in the anomaly shown in Fig. 11e (cf. Fig. 6a). [The statistics of the noise determined by (3) show these two eigenmodes on average are excited together and in opposite phases.] Subsequently, the 2-year eigenmode undergoes relatively more rapid propagation and decay (Fig. 10c), "revealing" the more slowly evolving and decaying 4-year eigenmode (Fig. 9c) and even beginning to reinforce it by $t = 9$ months (Fig. 11b,d,f). This simple transition from destructive to constructive interference between these two eigenmodes is the basis of optimal growth, with the bidecadal and other eigenmodes playing important but lesser roles.

Some features of these empirical eigenmodes appear similar to theoretical eigenmodes determined from various linearized versions of the Zebiak and Cane (ZC; 1987) intermediate complexity coupled atmosphere-ocean model (e.g., Jin and Neelin 1993; Dijkstra and Neelin 1999; Fedorov and Philander 2001; An and Jin 2001; MacMynowski and Tziperman 2008; Bejarano and Jin 2008). In particular, the 4-year eigenmode resembles a damped "recharge-discharge oscillator" (Jin 1997; Burgers et al. 2005), evolving from minimum SST phase when anomalous equatorial heat content is a maximum towards maximum SST phase (Fig. 9b) when anomalous equatorial heat content has decreased to zero, during which time the eastward moving thermocline anomaly forces SST amplification (through the $\mathbf{L}_{TZ}\mathbf{Z}_{20}$ term, not shown). Subsequent off-equatorial wind stress at and south of the equator, driven by the SST anomaly, then partly "recharges" the equatorial heat content anomaly (through the $\mathbf{L}_{Z\tau}\boldsymbol{\tau}_x$ term, not shown) with opposite sign. This eigenmode also appears to have some features characteristic of a "delayed-oscillator" (Schopf and Suarez 1988; Battisti and Hirst 1989), with a thermocline extremum in the far west equatorial Pacific (at about 5°N, 135°E) driven largely by $\mathbf{L}_{ZZ}\mathbf{Z}_{20}$ that could be consistent with equatorial Rossby waves (e.g., Battisti and Hirst 1989), and a second thermocline anomaly at about 14°N that moves westward (Fig. 9c), curving slightly southward as it crosses the Pacific, consistent with Rossby wave propagation (e.g., Schopf et al. 1981). But these are both fairly weak features, and most of the apparent evolution of the thermocline anomaly along the entire loop shown in Fig. 9c and d can be reproduced without the nonlocal portion of the $\mathbf{L}_{ZZ}\mathbf{Z}_{20}$ (not shown), except in the central subtropical Pacific (transect points 45–55). Additionally, the wind stress term $\mathbf{L}_{Z\tau}\boldsymbol{\tau}_x$ drives amplification of the east–west $\mathbf{Z}_{20}$ gradient on the equator, consistent with many earlier studies dating back to Wyrtki. As a result, the thermocline anomaly evolution does not appear as free Kelvin or Rossby wave propagation.

A different balance of forcing terms related to different $\tau_x$ and SST structure drives the faster evolution of the 2-year eigenmode. Along the equator, $\tau_x$ for the 4-year eigenmode extends from the dateline through the SST maximum at about 120°W, whereas $\tau_x$ reverses sign for the 2-year eigenmode at about 140°W, the location of the eigenmode's more westward SST maximum. SST and zonal wind stress anomalies also have greater meridional extent, particularly south of the equator, for the slower eigenmode. On the other hand, the two eigenmodes have only minor thermocline differences, primarily of amplitude, and consequently have similar thermocline forcing of SST evolution, including a discharge of anomalous equatorial heat content as SST evolves from minimum to maximum phase. However,

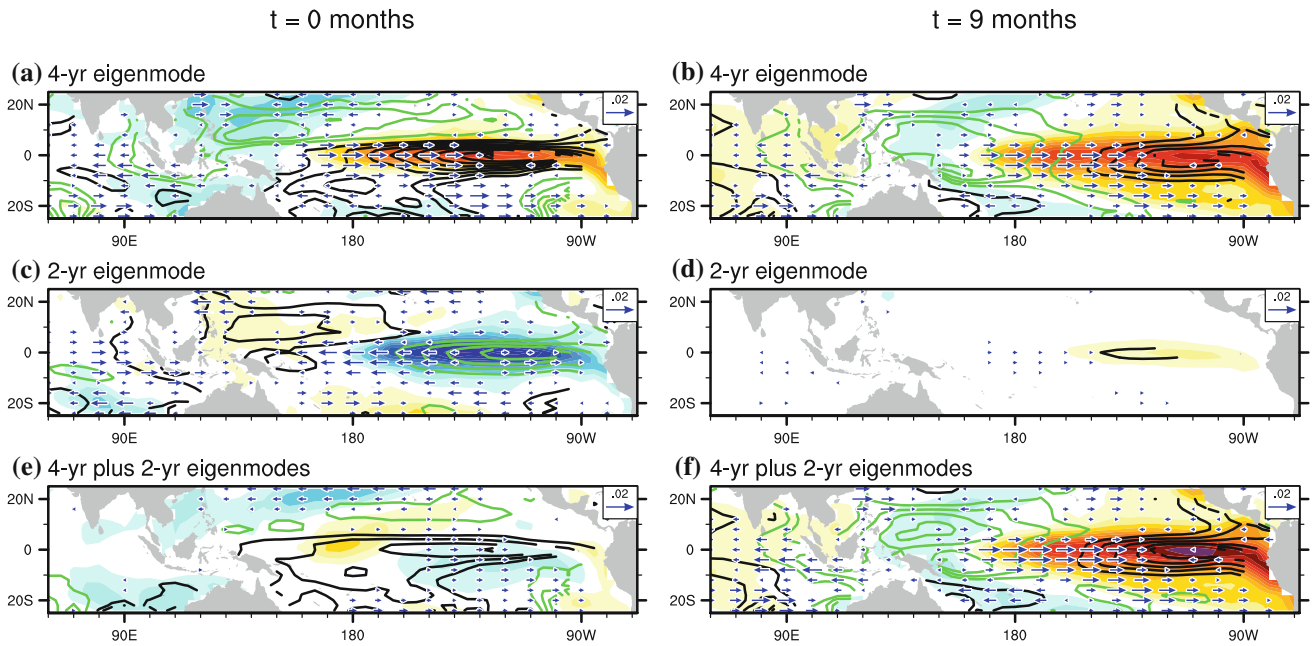


**Fig. 11** Evolution of the 4- and 2-year eigenmode components of $\mathbf{v}_1(9)$, with the initial state in the left panels and the evolved state in the right panels. **a** 4-year eigenmode component at $t = 0$. **b** 4-year eigenmode component at $t = 9$. **c** 2-year eigenmode component at $t = 0$. **d** 2-year eigenmode component at $t = 9$. **e** 4-year plus 2-year eigenmode components at $t = 0$. **f** 4-year plus 2-year eigenmode components at $t = 9$. Plotting conventions as in Fig. 6

while the thermocline forcing term is stronger than the non-local SST term for the 4-year eigenmode, it is weaker for the 2-year eigenmode (not shown). The imbalance between the two terms is also relatively larger, allowing for a more rapid evolution of the 2-year eigenmode in which the SST anomaly moves westward along the equator from the South American coast to near the dateline (Fig. 10c), where the SST anomaly becomes almost exactly out of phase with the thermocline anomaly below and so is weakened (consistent with the overall SST variance sink in Fig. 4c).

The bidecadal eigenmode bears some resemblance to "decadal" ENSO variability (e.g., Zhang et al. 1997) and its contribution to $\mathbf{v}_1(9)$ is essential for capturing the decadal tail of the $\mathbf{T}_O$/PC1 spectrum (Fig. 14). Its much longer time scale is consistent with its more poleward subtropical $\mathbf{Z}_{20}$ anomalies in both hemispheres, which after being induced by anomalies moving from the equator in the east central Pacific propagate more slowly toward the western boundary as might be expected of subtropical Rossby waves (Capotondi and Alexander 2001) at these higher latitudes (e.g., Chelton and Schlax 1996).

For longer optimization periods, the optimal structure projects more strongly upon other eigenmodes besides the three eigenmodes discussed above, although the 4-year eigenmode remains the largest component. In this case, predictable ENSO evolution over periods of a year and more may not be so clearly identified with a few eigenmodes.

## 5 Concluding remarks

Including thermocline depth and wind stress in the PS95 SST-only LIM has provided a foundation for empirical diagnosis of observed coupled atmosphere-ocean dynamics. We find that the SST pattern leading to optimal ENSO excitation, originally found by PS95, also corresponds to an east-central Pacific thermocline anomaly similar to an ENSO precursor found by many previous studies. Each contributes roughly equally to subsequent ENSO evolution and amplification, consistent with a large ENSO event possibly requiring both surface and subsurface processes as a trigger (e.g., McPhaden 1999; Vimont et al. 2003; Anderson 2007; Chang et al. 2007). Also, since anomalous thermocline evolution is primarily forced by wind stress, free ocean wave dynamics are likely of secondary importance to ENSO evolution, although rapid adjustment of the ocean on a time scale shorter than the 3 month data-averaging period (such as through Kelvin waves) can be a part of the wind stress forcing term.

Our results support the view of ENSO as an episodic and not periodic (e.g., Kessler 2002) effectively linear process. Essentially, a rapidly evolving westward-propagating

eigenmode driven predominantly but not exclusively by shallow SST-atmosphere dynamics first destructively and then constructively interferes with a more slowly evolving eastward-propagating eigenmode driven by mixed SST/ocean dynamics. Since the slower mode has a damping time scale about half its period, it continues to predictably evolve for another quarter cycle; that is, an ENSO episode involves anomaly growth, decay, and overshoot, the last stage occurring only along the equator with a weak sign reversal of anomalous SST and zonally-integrated warm water volume. This picture is enriched but not fundamentally altered by the presence of the other eigenmodes. Some previous observational studies have also emphasized the combined contribution of similar tropical Pacific structures with 4- and 2-year periods (e.g., Jiang et al. 1995; White et al. 2003), although no decay time scales were determined. Also, eigenanalysis using the full ZC intermediate coupled model (Bejarano and Jin 2008) results in two leading eigenmodes, again with 4- and 2-year periods, that capture many but not all of the characteristics noted above, including a relative difference in the balance between processes forcing SST anomalies. Theoretical eigenanalyses focusing on the weakly unstable portion of a large parameter space have led to the view that ENSO variations (or "flavors") are a consequence of changes in the background state altering stability characteristics of the two leading eigenmodes (e.g., Fedorov and Philander 2001; Bejarano and Jin 2008), and/or of nonlinear interaction between them (e.g., Jin et al. 1994). Our empirical results support the alternative view that both (stable) eigenmodes are always relevant since rapid observed ENSO amplification is a consequence of their modal interference. Hence, events differ because noise variations differently excite each eigenmode, although on average these two eigenmodes are preferentially excited with an initial (mostly opposing) relationship to each other. As a consequence, the relative importance of key physical processes changes from event to event, yet background state changes in stability properties may have only minor impact. Moreover, the leading optimal structure is the "best-case scenario"; suboptimal ENSO events are also possible, including events where anomalies develop less rapidly but in different tropical regions (such as "Modoki" or central Pacific ENSOs), depending on the particular noise realization.

While over short time intervals, the ocean LIM yields minimal SST forecast improvement, over time intervals greater than several months it improves upon the SST-LIM in all measures, including long-lead tropical SST predictions, simulation of power spectra and lag co-variance statistics, and the observed relevance of leading optimal structures (Fig. 8). Much of this improvement may be due to a second optimal structure, allowing for significant SST anomaly growth over time intervals longer than a year, not present in the SST-LIM. Growth still occurs when initial destructive interference with the 4-year mode evolves into constructive interference, although this is now the result of many rather than a few eigenmodes. While the initial anomaly lies mostly within the SST field, subsequent ENSO development cannot be captured by SST alone since it is driven by persistent off-equatorial wind stress in the eastern Pacific that deepens (shallows) the equatorial thermocline to drive warm (cold) SST anomaly amplification (e.g., Wang et al. 2003).

The improved validation of the ocean LIM gives further support to the effectively linear, stochastically forced view of tropical ocean variability. It is particularly encouraging that, despite inherent limitations on ocean data assimilation, $\mathbf{Z}_{20}$ has improved the LIM's representation of ENSO dynamics. While some remaining deficiencies of this LIM may represent truly predictable (i.e., slowly-varying) nonlinear internal ocean dynamics representable only by a high-dimensional fully nonlinear CGCM (Chen et al. 2004), we might hope for further improvement of the LIM by explicit representation of additional processes within its state vector, such as vertical stratification, currents, the extratropics, and/or atmospheric variability. For example, due to the dependence of current strength on thermocline slope, one clear deficiency of this study is that not explicitly including currents complicates distinguishing advective feedback from other thermocline feedbacks.

One key issue that remains is determining the correct seasonal dependence of both $\mathbf{L}$ and $\mathbf{Q}$, given the seasonal dependence of $\mathbf{C}(0)$ and $\mathbf{C}(\tau)$ as well as the tropical base state. We have followed the approach of some previous analyses (e.g., PS95, Penland 1996; Chang et al. 2007; Newman et al. 2009) which found that ENSO seasonality, including its phase locking, can be captured by a LIM where $\mathbf{L}$ is fixed but $\mathbf{Q}$ is seasonally-varying. As noted in the "Appendix", we have also investigated potential seasonality of the dynamical operator but obtained results that were actually slightly worse. Whether this is because even 42 years of data is too short to correctly capture the seasonality of $\mathbf{L}$, or because noise effects make the seasonality of the effectively linear operator $\mathbf{L}$ weaker than a linearization of the physical equations would lead us to expect, or because seasonality is better determined not piecewise by season (e.g., Johnson et al. 2000; Xue et al. 2000) but instead from the entire year in a cyclo-stationary LIM (e.g., Ortiz-Bevía 1997), remains to be determined.

"Metrics" of important phenomena within coupled GCMs generally do not directly measure dynamical processes of interest but rather their resulting variability. For example, Nino3 power spectra typically used to compare ENSO SST variability (e.g., Guilyardi et al. 2009) do not determine how, only whether, ENSO dynamics differ between models. By constructing even a rudimentary air-

SST-subsurface ocean LIM, we go beyond simpler metrics to estimate the relative impact of different coupled feedback terms, and their importance both to the SST variance budget and to the evolution of ENSO events. This can serve as an important baseline for coupled dynamics as simulated by coupled GCMs. In particular, since the 4- and 2-year eigenmodes together provide a fairly compact representation of observed ENSO dynamics, it may prove useful to compare corresponding eigenmodes determined from the output of coupled GCM runs. Moreover, recall that in the extreme case where thermocline interactions were entirely decoupled from the surface in the LIM, SST spectra became more sharply peaked with reduced decadal variability, and SST variability shifted towards the central Pacific. Similarly, Newman et al. (2009) found that removing air-sea coupling in their tropical LIM led to a greatly weakened ENSO whose period was too short and whose maximum amplitude was too far west. These are all common CGCM failings, suggesting that climate LIMs might also be useful for diagnosing feedbacks leading to errors in comprehensive coupled climate models.

**Acknowledgments** The authors thank Antonietta Capotondi, Cécile Penland, Amy Solomon, David Battisti and two anonymous reviewers for helpful comments. This work was partially supported by a grant from NOAA CLIVAR-Pacific.

## Appendix: Construction and validation of the LIM

In any multidimensional statistically stationary system with components $x_i$, one may define a time lag covariance matrix $\mathbf{C}(\tau)$ with elements $Cij(\tau) = \langle x_i(t+\tau)x(t)\rangle$, where angle brackets denote a long term average. In linear inverse modeling, one assumes that the system satisfies $\mathbf{C}(\tau) = \mathbf{G}(\tau)\mathbf{C}(0)$, where importantly $\mathbf{G}(\tau) = exp(\mathbf{L}\tau)$ and $\mathbf{L}$ is a constant matrix, which follows from (2). One then uses this relationship to estimate $\mathbf{L}$ from observational estimates of $\mathbf{C}(0)$ and $\mathbf{C}(\tau_0)$ at some lag $\tau_o$. For the statistics of this system to be stationary, $\mathbf{L}$ must be dissipative, i.e its eigenvalues must have negative real parts. In a forecasting context, $\mathbf{G}(\tau)\mathbf{x}(t)$ represents the "best" forecast (in a least squares sense) of $\mathbf{x}(t+\tau)$ given $\mathbf{x}(t)$. Note that unlike multiple linear regression, determination of $\mathbf{G}$ at one lag $\tau_o$ identically gives $\mathbf{G}$ at all other lags. Also, statistics of the noise forcing, not just the error covariance, are determined by LIM since the positive-definite noise covariance matrix $\mathbf{Q} = \langle \xi\xi^T \rangle dt$ is determined from a Fluctuation-Dissipation relationship (3), given the observed $\mathbf{C}(0)$ and $\mathbf{L}$.

A training lag of $\tau_o = 3$ months was used to determine $\mathbf{L}$. The EOF truncations (Sect. 2) and training lag were chosen to maximize the LIM's cross-validated forecast skill for leads up to 18 months, while avoiding the Nyquist problem for $\mathbf{L}$ (PS95) that inhibits analysis of interactions amongst the model variables. In no other respect do these choices qualitatively affect the points made in this paper. Estimates of $\mathbf{L}$ and of forecast skill were cross-validated by sub-sampling the data record by sequentially removing one five-year period, computing $\mathbf{L}$ for the remaining years, and then generating forecasts for the independent years. This procedure was repeated for the entire period. Forecast skill is then determined by comparing the local anomaly correlation between the cross-validated model predictions and gridded untruncated verifications.

Explicitly including both $\tau_x$ and $\mathbf{Z}_{20}$ in the LIM state vector increases skill of 9 (Figs. 12a, b) and 18 (Fig. 12c, d) month $\mathbf{T}_o$ forecasts. Skill improvement is mostly due to the inclusion of $\mathbf{Z}_{20}$ rather than $\tau_x$, confirming that the $\mathbf{Z}_{20}$ data provides useful information. In addition, the LIM makes $\mathbf{Z}_{20}$ forecasts whose skill is significantly better than persistence (not shown). The difference in SST skill between the ocean and SST-LIMs generally increases with forecast lead because the ocean LIM captures the slower $\mathbf{Z}_{20}$

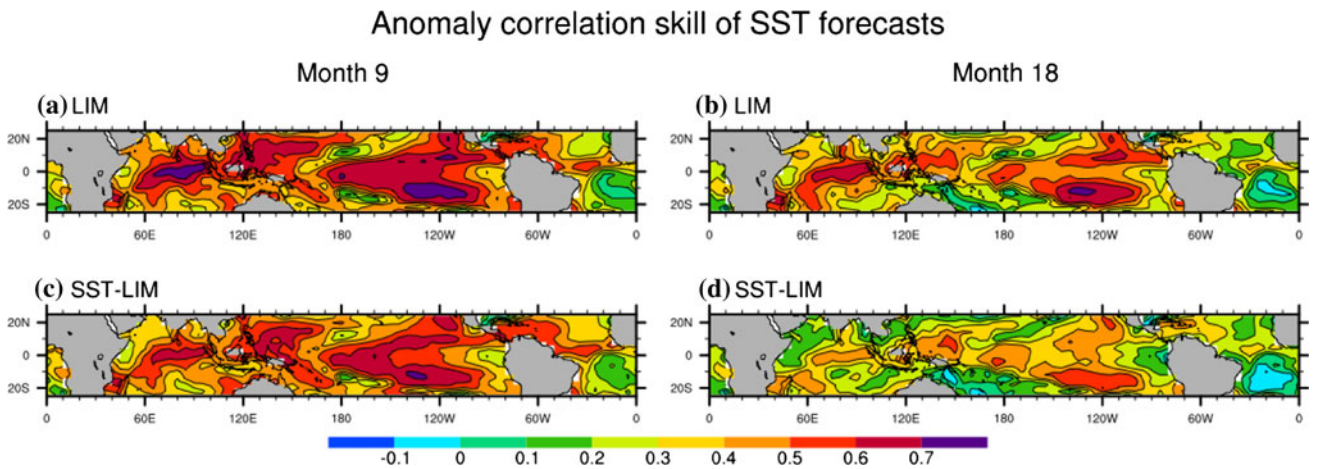


**Fig. 12** Cross-validated forecast skill (1959–2000) of $\mathbf{T}_O$ for forecast leads of 9 and 18 months, for the LIM (*top*) and the SST13-LIM (*bottom*)

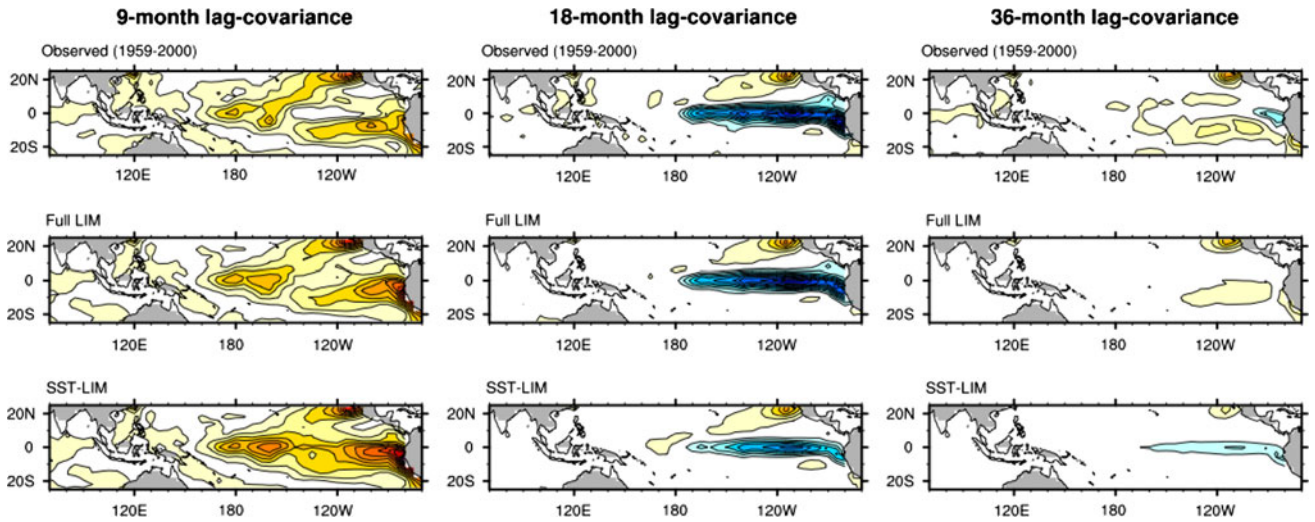


**Fig. 13** Observed (*top panels*), LIM (*middle panels*) and SST13-LIM (*bottom panels*) $\mathbf{T}_O$ lag-covariance. (*left*) 9-month lag-covariance; (*center*) 18-month lag-covariance; (*right*) 36-month lag-covariance. Note that the observed lag-covariances are based on the full (that is, not truncated in the EOF basis) gridded anomaly fields. Contour interval is 0.04 $K^2$

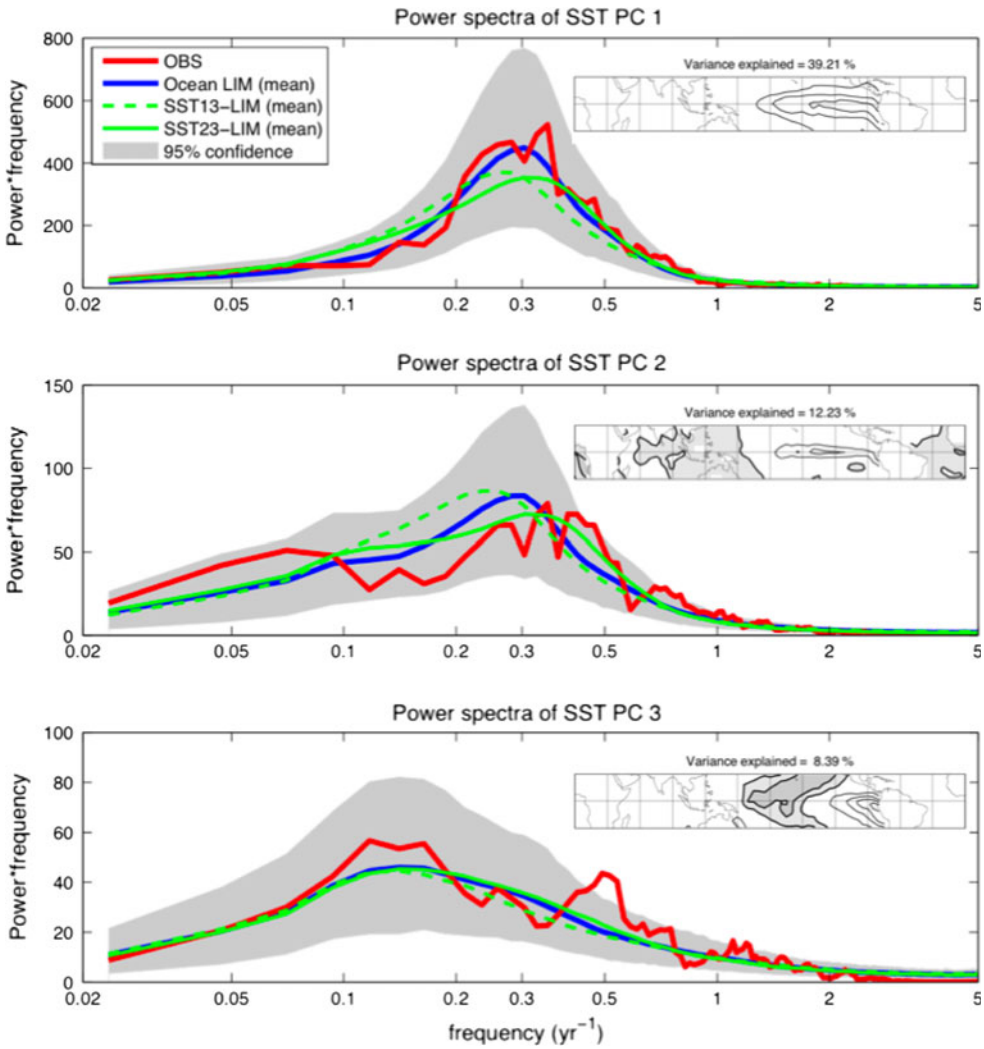


**Fig. 14** Power spectra for the three leading SST ($\mathbf{T}_O$) PCs (*red lines*), compared to that predicted by the LIM (*blue lines*). *Gray shading* represents the 95% confidence interval determined from a 1,000-member ensemble of 42 year LIM model runs (see text for further details). In these log(frequency) versus power times angular frequency ($\omega$) plots, the area under any portion of the curve is equal to the variance within that frequency band. Note that displaying power times frequency slightly shifts the power spectral density peak centered at a period of 4.5 years to a variance peak centered at a period of 3.5 years. Insets in each panel show the corresponding EOF and the variance explained by that pattern

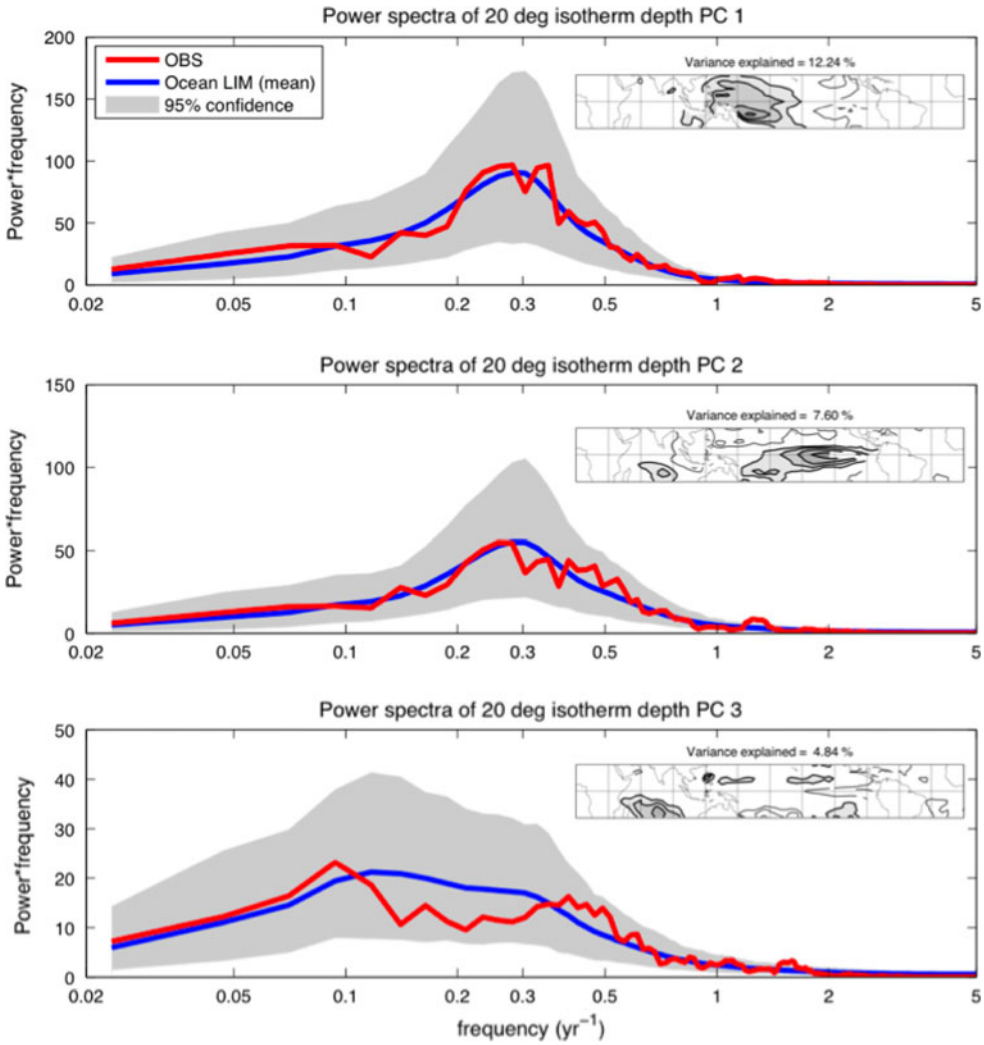


**Fig. 15** Same as Fig. 14 but for the three leading 20°C isotherm depth ($\mathbf{Z}_{20}$) PCs. Note that there is no SST-LIM equivalent

evolution. For example, a second set of "fixed-$\mathbf{Z}_{20}$" ocean LIM forecasts in which the initial $\mathbf{Z}_{20}$ anomaly is persisted throughout the forecast period (i.e., $\mathbf{Z}_{20}(t) = \mathbf{Z}_{20}(0)$; not shown) has greatly reduced 18-month forecast skill.

We test the validity of the linear approximation with a "tau-test" (PS95). For example, since (2) implies that $\mathbf{C}(\tau) = \mathbf{G}(\tau)\ \mathbf{C}(0)$, the LIM should be able to reproduce observed lag-covariance statistics at much longer lags than the 3 month lag on which the LIM was trained (e.g., $\mathbf{C}(18) = [\mathbf{G}(3)]^6\ \mathbf{C}(0)$). Figure 13 compares the observed and predicted lag-autocovariances of $\mathbf{T}_O$ for lags of 9, 18, and 36 months, using both the ocean LIM and the SST13-LIM. The SST13-LIM does a reasonably good job capturing the main aspects of the lag-autocovariance pattern, but for lags less than about a year it tends to overestimate persistence especially along the equator and for longer lags it errs in the amplitude of the negative lag-autocovariance. The ocean LIM improves upon these deficiencies, as well as reproducing observed lag-covariance for $\mathbf{Z}_{20}$ over the same lags (not shown).

A complementary test of linearity is to compare observed and LIM-predicted power spectra, by integrating (2) for 42,000 years using the method described in Penland and Matrosova (1994) and Newman et al. (2009) and then collecting statistics. The white noise forcing is determined from the noise covariance matrix $\mathbf{Q}$ determined as a residual from (3). $\mathbf{Q}$ should be positive-definite but determined this way it is only guaranteed to be symmetric. Ensuring positive-definiteness in the manner of Penland and Matrosova (1994), by rescaling the noise due to one small negative eigenvalue of $\mathbf{Q}$ that accounted for less than 0.25% of the trace of $\mathbf{Q}$, resulted in an almost negligible impact on $\mathbf{C}(0)$. The resulting model "data" is separated into 1,000 42-year time series. The observed spectra and the ensemble mean of the model spectra for the three leading PCs of $\mathbf{T}_O$ and $\mathbf{Z}_{20}$ are shown in Figs. 14 and 15, respectively. The corresponding EOF pattern for each spectrum is shown in the inset panels. The gray shading shows the 95% confidence intervals of these spectra, estimated using the 1,000 model realizations.

The LIM reproduces the main features of the observed power spectra for the leading PCs of each variable (including $\tau_x$, not shown). Obviously, the mean LIM spectra are much smoother than observed, due to the relatively few degrees of freedom in the truncated EOF space. On the other hand, the irregularity of the observed spectra is at least partly due to sampling, as indicated by the confidence intervals, which show how much variation in the spectra could occur simply from different realizations of noise.

Since $\mathbf{C}(0)$ and $\mathbf{C}(\tau)$ have seasonal dependence, both Johnson et al. 2000 and Xue et al. 2000 suggested that $\mathbf{L}$ should also be considered to be seasonally-varying. They constructed "Markov models" in which they determined $\mathbf{G}(\tau_o)$ for each season (but not $\mathbf{L}$) so that forecasts are made by an appropriate product of each $\mathbf{G}$. On the other hand, PS95 and Penland (1996) argued that the observed seasonality of ENSO, including its phase locking, can be captured with a fixed $\mathbf{L}$ with seasonally varying $\mathbf{Q}$. Newman et al. (2009) found that a tropical atmosphere-SST LIM constructed using weekly data had a poorer representation of coupled ENSO dynamics when segregated by season, although the internal subseasonal atmospheric dynamics were slightly improved. They also found pronounced seasonality of the noise, as did Penland (1996) and Chang et al. (2007). Similar to Penland, we found that the seasonal dependence of the tests above can be generally captured by assuming fixed $\mathbf{L}$ but with seasonally varying $\mathbf{Q}$. As in Newman et al. (2009), we also found that seasonal LIMs did poorer in these tests compared to those using a fixed $\mathbf{L}$.

## References


Alexander MA, Matrosova L, Penland C, Scott JD, Chang P (2008) Forecasting Pacific SSTs: linear inverse model predictions of the PDO. J Clim 21:385–402

An SI, Jin FF (2001) Collective role of thermocline and zonal advective feedbacks in the ENSO mode. J Clim 14:3421–3432

Anderson BT (2007) Intraseasonal atmospheric variability in the extratropics and its relation to the onset of tropical Pacific sea surface temperature anomalies. J Clim 20:926–936

Ashok K, Behere SK, Rao SA, Weng H, Yamagata T (2007) El Nino Modoki and its possible teleconnection. J Geophys Res 112. doi: 10.1029/2006JC003798

Battisti DS, Hirst AC (1989) Interannual variability in a tropical atmosphere-ocean model: influence of the basic state, ocean geometry, and nonlinearity. J Atmos Sci 46:1687–1712

Bejarano L, Jin FF (2008) Coexistence of equatorial coupled modes of ENSO. J Clim 21:3051–3067

Bjerknes J (1969) Atmospheric teleconnections from the equatorial Pacific. Mon Weather Rev 97:163–172

Boulanger JP, Menkes C (2001) The Trident Pacific model. Part 2: role of long equatorial wave reflection on sea surface temperature anomalies during the 1993–1998 TOPEX/POSEIFON period. Clim Dyn 17:175–186

Burgers G, van Oldenborgh GJ (2003) On the impact of local feedbacks in the central Pacific on the ENSO cycle. J Clim 16:2396–2407

Burgers G, Jin FF, van Oldenborgh GJ (2005) The simplest ENSO recharge oscillator. Geophys Res Lett 32:L13706. doi: 10.1029/2005GL022951

Capotondi A, Alexander MA (2001) Rossby waves in the tropical North Pacific and their role in decadal thermocline variability. J Phys Oceanogr 31:3496–3515

Carton JA, Giese BS (2008) A reanalysis of ocean climate using Simple Ocean Data Assimilation (SODA). Mon Weather Rev 136:2999–3017

Chang P, Zhang L, Saravanan R, Vimont DJ, Chiang JCH, Ji L, Seidel H, Tippett MK (2007) Pacific meridional mode and El Niño-Southern Oscillation. Geophys Res Lett 34:L16608. doi: 10.1029/2007GL030302

Chelton DB, Schlax MG (1996) Global observations of oceanic Rossby waves. Science 272:234–238

Chelton DB, Wentz FJ, Gentemann CL, de Szoeke RA, Schlax MG (2000) Satellite microwave SST observations of transequatorial tropical instability waves. Geophys Res Lett 27:1239–1242

Chen D, Cane MA, Kaplan A, Zebiak SE, Huang D (2004) Predictability of El Niño over the past 148 years. Nature 428:733–736

Clarke AJ, Van Gorder S (2003) Improving El Niño prediction using a space-time integration of Indo-Pacific winds and equatorial Pacific upper ocean heat content. Geophys Res Lett 30:1399. doi:10.1029/2002GL016673

Dijkstra HA, Neelin JD (1999) Coupled processes and the tropical climatology. Part III: instabilities of the fully coupled climatology. J Clim 12:1630–1643

Farrell B (1988) Optimal excitation of neutral Rossby waves. J Atmos Sci 45:163–172

Fedorov AV, Philander SG (2001) A stability analysis of tropical ocean-atmosphere interactions: Bridging measurements and theory for El Niño. J Clim 14:3086–3101

Flügel M, Chang P (1996) Impact of dynamical and stochastic processes on the predictability of ENSO. Geophys Res Lett 23:2089–2092

Frankignoul C, Hasselman K (1977) Stochastic climate models. Part II: application to sea-surface temperature variability and thermocline variability. Tellus 29:289–305

Guilyardi E, Wittenberg A, Fedorov A, Collins M, Wang C, Capotondi A, van Oldenborgh GJ, Stockdale T (2009) Understanding El Niño in ocean-atmosphere general circulation models: progress and challenges. Bull Am Meteor Soc 90:325–340

Hasselmann K (1976) Stochastic climate models. Part I. Theory. Tellus 28:474–485

Jiang N, Neelin JD, Ghil M (1995) Quasi-quadrennial and quasi-biennial variability in the equatorial Pacific. Clim Dyn 12:101–112

Jin FF (1997) An equatorial ocean recharge paradigm for ENSO. Part I: conceptual model. J Atmos Sci 54:811–829

Jin FF, An SI (1999) Thermocline and zonal advective feedbacks within the equatorial ocean recharge oscillator model for ENSO. Geophys Res Lett 26:2989–2992

Jin FF, Neelin JD (1993) Modes of interannual tropical ocean-atmosphere interaction—a unified view. Part III: analytical results in fully coupled cases. J Atmos Sci 50:3523–3540

Jin FF, Neelin JD, Ghil M (1994) El Nino on the "Devils Staircase": annual subharmonic steps to chaos. Science 264:710–713

Jin FF, Kim ST, Bejarano L (2006) A coupled-stability index for ENSO. Geophys Res Lett 33:L23708. doi:10.1029/2006GL027221

Jochum M, Cronin MF, Kessler WS, Shea D (2007) Observed horizontal temperature advection by tropical instability waves. Geophys Res Lett 34:L09604. doi:10.1029/2007GL029416

Johnson SD, Battisti DS, Sarachik ES (2000) Empirically derived Markov models and prediction of tropical pacific sea surface temperature anomalies. J Clim 13:3–17

Kallummal R, Kirtman B (2008) Validity of a linear stochastic view of ENSO in an ACGCM. J Atmos Sci 65:3860–3879. doi: 10.1175/2008JAS2286.1

Kalnay E et al (1996) The NCEP/NCAR 40-Year reanalysis project. Bull Am Meteor Soc 77:437–471

Kang IS, An SI, Jin FF (2001) A systematic approximation of the SST anomaly equation for ENSO. J Met Soc Jpn 79:1–10

Kao HY, Yu JY (2009) Contrasting Eastern-Pacific and Central-Pacific types of ENSO. J Clim 22:615–632

Kessler WS (2002) Is ENSO a cycle or a series of events? Geophys Res Lett 29:2125. doi:10:1029/2002GL015924

Kessler WS, McPhaden MJ (1995) The 1991–93 El Niño in the central Pacific. Deep-Sea Res II 42:295–334

Latif M, Graham NE (1992) How much predictive skill is contained in the thermal structure of an OGCM? J Phys Oceanogr 22:951–962

MacMynowski DG, Tziperman E (2008) Factors affecting ENSO's period. J Atmos Sci 65:1570–1586

McPhaden MJ (1999) Genesis and evolution of the 1997–98 El Niño. Science 283:950–954

McPhaden MJ, Zhang X, Hendon HH, Wheeler MC (2006) Large scale dynamics and MJO forcing of ENSO variability. Geophy Res Lett 33:L16702. doi:10.1029/2006GL026786

Meinen CS, McPhaden MJ (2000) Observations of warm water volume changes in the equatorial pacific and their relationship to El Nino and La Nina. J Clim 13:3551–3559

Moore A, Kleeman R (1997) The singular vectors of a coupled-atmosphere model of ENSO. II: sensitivity studies and dynamical interpretation. Quart J R Meteor Soc 123:983–1006

Moore A, Kleeman R (1999) The non-normal nature of El Niño and intraseasonal variability. J Clim 12:2965–2982

Neelin JD (1991) The slow sea surface temperature mode and the fast-wave limit: analytic theory for tropical interannual oscillations and experiments in hybrid coupled models. J Atmos Sci 48:584–606

Neelin JD, Jin FF (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

Newman M (2007) Interannual to decadal predictability of tropical and North Pacific sea surface temperatures. J Clim 20:2333–2356

Newman M, Sardeshmukh PD (2008) Tropical and stratospheric influences on extratropical short-term climate variability. J Clim 21:4326–4347

Newman M, Sardeshmukh PD, Winkler CR, Whitaker JS (2003) A study of subseasonal predictability. Mon Weather Rev 131:1715–1732

Newman M, Sardeshmukh PD, Penland C (2009) How important is air-sea coupling in ENSO and MJO evolution? J Clim 22:2958–2977

Ortiz-Bevía MJ (1997) Estimation of the cyclostationary dependence in geophysical data fields. J Geophys Res 102:13473–13486

Papanicolaou G, Kohle W (1974) Asymptotic theory of mixing stochastic ordinary differential equations. Commun Pure Appl Math 27:641–668

Penland C (1989) Random forcing and forecasting using principal oscillation pattern analysis. Mon Weather Rev 117:2165–2185

Penland C (1996) A stochastic model of IndoPacific sea surface temperature anomalies. Phys D 98:534–558

Penland C, Matrosova L (1994) A balance condition for stochastic numerical models with application to the El Niño-Southern Oscillation. J Clim 7:1352–1372

Penland C, Matrosova L (2006) Studies of El Niño and interdecadal variability in Tropical sea surface temperatures using a nonnormal filter. J Clim 19:5796–5815

Penland C, Sardeshmukh PD (1995) The optimal growth of tropical sea surface temperature anomalies. J Clim 8:1999–2024

Penland C, Flügel M, Chang P (2000) Identification of dynamical regimes in an intermediate coupled ocean-atmosphere model. J Clim 13:2105–2115

Picaut J, Ioualalen M, Menkes C, Delcroix T, McPhaden MJ (1996) Mechanism of the zonal displacements of the Pacific Warm Pool: Implications for ENSO. Science 274:1486–1489

Rayner NA, Parker DE, Horton EB, Folland CK, Alexander LV, Rowell DP, Kent EC, Kaplan A (2003) Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. J Geophys Res 108:4407. doi: 10.1029/2002JD002670

Schopf PS, Suarez MJ (1988) Vacillations in a coupled ocean-atmosphere model. J Atmos Sci 45:549–566

Schopf PS, Anderson DLT, Smith R (1981) Beta-dispersion of low frequency Rossby waves. Dyn Atmos Oceans 5:187–214

Thompson CJ, Battisti DS (2001) A linear stochastic dynamical model of ENSO. Part II: analysis. J Clim 14:445–466

Vimont DJ, Wallace JM, Battisti DS (2003) The seasonal footprinting mechanism in the Pacific: implications for ENSO. J Clim 16:2668–2675

Wang W, McPhaden MJ (2000) The surface-layer heat balance in the equatorial Pacific ocean. Part II: interannual variability. J Phys Ocean 30:2989–3008

Wang X, Jin FF, Wang Y (2003) A tropical ocean recharge mechanism for climate variability. Part I: equatorial heat content changes induced by the off-equatorial wind. J Clim 16:3585–3598

White WB, Tourre YM, Barlow M, Dettinger M (2003) A delayed action oscillator shared by biennial, interannual, and decadal signals in the Pacific Basin. Geophys Res 108. doi: 10.1029/2002JC001490

Winkler CR, Newman M, Sardeshmukh PD (2001) A linear model of wintertime low-frequency variability. Part I: formulation and forecast skill. J Clim 14:4474–4494

Wyrkti K (1985) Water displacements in the Pacific and the genesis of El Niño cycles. J Geophys Res 90:7129–7132

Xue Y, Leetmaa A, Ji M (2000) ENSO prediction with Markov models: the impact of sea level. J Clim 13:849–871

Zebiak SE, Cane MA (1987) A model El Niño-southern oscillation. Mon Weather Rev 115:2268–2278

Zelle H, Appeldoorn G, Burgers G, van Oldenborgh GJ (2004) The relationship between sea surface temperature and thermocline depth in the eastern equatorial Pacific. J Phys Ocean 34:643–655

Zhang Y, Wallace JM, Battisti DS (1997) ENSO-like interdecadal variability. J Clim 10:1004–1020