Clim Dyn (2015) 45:47–65
DOI 10.1007/s00382-014-2300-8

Optimal forcing of ENSO either side of the 1970's climate shift and its implications for predictability

Christopher M. Aiken · Agus Santoso · Shayne McGregor · Matthew H. England

Received: 25 February 2014 / Accepted: 12 August 2014 / Published online: 28 August 2014

Abstract Inverse methods are used to investigate changes in the precursors to El Niño Southern Oscillation (ENSO) events since the so-called 1970's climate shift, associated with a change in the phase of the Interdecadal Pacific Oscillation (IPO). Linear Inverse Models (LIMs) constructed from tropical sea surface temperature, thermocline depth and zonal wind stress anomalies from each of the periods 1959–1978 and 1979–1998, are able to reproduce the major observed characteristics of ENSO, including its amplitude, frequency and time evolution. Each LIM possesses low-frequency and biennial ENSO modes, the former being both the least damped and the mode responsible for strongest pseudoresonance, as quantified via calculation of the resolvent norm. Because these modes are damped, ENSO variability is sustained in the stochastically forced LIMs by transiently growing perturbations, and predictability is determined by the character of the transiently growing subspace of perturbations. The optimal linear precursor over any given lead time is equivalent to the optimal perturbation of the LIM, that represents the most rapidly growing linear perturbation over that timescale. In both periods linear ENSO growth occurs through one of two trajectories associated with the 7 and 15 month optimal perturbations. The structure of these two optimal perturbations change significantly between the two periods, and their ability to predict ENSO degrades dramatically when applied to the alternate period. This suggests that ENSO precursors changed following the 1970s climate shift over both 7 and 15 month time-scales. In particular, while prior to the climate shift the heat content of the equatorial Pacific alone is a skillful ENSO predictor on 7 month lead times, afterwards Indian and south Atlantic sea surface temperature anomalies are inferred to have become important. Optimal ENSO growth over 15 months also contains a significant extra-Pacific contribution, and it is possible to skillfully hindcast some (but not all) ENSO events in both periods over 15 month lead times. Of the four considered linear precursors, only the 7 month optimal perturbation corresponding to the period 1959–1978 is able to skillfully hindcast ENSO amplitude from 1998 to the present, correctly predicting the development of El Niño conditions since February 2014. As such the optimal precursor structure appears to be related to the phase of the IPO, and we conjecture that extra-Pacific teleconnections may gain importance during a positive phase of the IPO.

Keywords ENSO precursors · Linear inverse model · Climate shift · Generalised stability theory

C. M. Aiken (✉) · A. Santoso · S. McGregor · M. H. England
ARC Centre of Excellence for Climate System Science, University of New South Wales, Sydney, NSW 2052, Australia
e-mail: c.aiken@unsw.edu.au

1 Introduction

The El Niño-Southern Oscillation (ENSO) is the dominant mode of interannual climate variability on Earth (McPhaden et al. 2006). While four decades have elapsed since Bjerknes (1969) first identified the positive feedback between the trade winds, sea surface temperature and thermocline depth in the equatorial Pacific that supports ENSO variability, an explanation for the differences between individual ENSO events remains an active focus of research. Uncertainty still exists regarding the mechanisms that initiate and terminate ENSO events, the roles played by the mean state and external forcing, and the very dynamical nature of ENSO, be-it an unstable non-linear chaotic mode, an unstable linear mode modulated by noise, or a damped mode that is

sustained by external forcing. Under the latter scenario in particular, the appearance of ENSO events depends sensitively upon the details of the external forcing. While early studies sought for local ENSO precursors in the tropical Pacific, finding a role of zonal wind (Moore and Kleeman 1999; Clarke and Van Gorder 2003) and surface heat content (Penland and Sardeshmukh 1995; Xue et al. 2000; Meinen and McPhaden 2000), more recent work has investigated the possibility that ENSO-initiating signals may come from further afield (Vimont et al. 2003; Kug and Kang 2006; Dommenget et al. 2006; Keenlyside and Latif 2007; Alexander et al. 2008; Jansen et al. 2009; Rodríguez-Fonseca et al. 2009; Izumo et al. 2010; Frauen and Dommenget 2012; Keenlyside et al. 2013; Zhang et al. 2014). In addition, there is some evidence to suggest that the ENSO precursors may have changed since the so-called 1970s climate shift (Boschat et al. 2013), corresponding to an apparent shift from a negative to positive phase of the Interdecadal Pacific Oscillation (IPO; Power et al. 1999; England et al. 2014). In the following we investigate mechanisms for ENSO generation, and changes in these, over the climate shift under the assumption that ENSO variability is consistent with a stochastically forced damped linear system. Optimal ENSO growth is determined by applying generalised stability analysis to the linear dynamical operators governing ENSO dynamics, determined empirically via inverse methods.

While ENSO dynamics have commonly been simulated by reduced-physics analytical or numerical models, Linear Inverse Models (LIMs) offer the advantage that no a priori assumptions are made regarding the physics of ENSO, other than that it is a stochastically forced linear system that obeys a fluctuation–dissipation relationship. Implicit in this assumption is that non-linear processes are of sufficiently short autocorrelation timescales so as to be modelled as part of the Gaussian white noise process representing external forcing, or may be subsumed into a linear term. Given sufficient observations it is possible to deduce the most likely underlying dynamical operator and hence investigate the sensitivity properties of the system empirically.

Linear inverse models (LIMs) have shown significant success in reproducing seasonal tropical sea surface temperature (SST) variability associated with ENSO (Penland and Magorian 1993; Penland and Matrosova 1994; Penland and Sardeshmukh 1995; Xue et al. 2000; Newman et al. 2009, 2011a) and other climate processes (Winkler et al. 2001; Newman et al. 2003; Alexander et al. 2008). For example, the Penland and Magorian (1993) LIM of tropical Indo-Pacific SST demonstrated forecast skill out to 9 months that rivals that of coupled GCMs, and the LIM developed in Newman et al. (2011a) extended the lead time over which forecast skill is significant out to 24 months. Operational LIM-based weekly forecasts of the tropical ocean-atmosphere, based on the LIM of Newman et al. (2009), are provided on-line by the NOAA/ESRL Physical Sciences Division and CIRES/Climate Diagnostics Center, Boulder, CO from their website at http://www.esrl.noaa.gov/psd/forecasts/sstlim/. The fact that LIMs can produce skillful forecasts supports the hypothesis that a substantial degree of ENSO behavior can be explained by the interaction of stochastically forced damped linear modes. Because the underlying dynamical operator is nonnormal, interference between the damped non-orthogonal modes can allow SST anomalies (SSTA) to grow transiently (Moore and Kleeman 1999). The particular composition of modes that produces optimal transient growth of tropical SST anomalies—the so-called optimal perturbation—has been shown to correlate well with ENSO metrics at various lead times, and to be consistent with observed predictability limits (Penland and Sardeshmukh 1995; Newman et al. 2011a). These authors demonstrate that optimal ENSO growth can be linked to interference between two or three modes that are not necessarily the least damped. In fact, Newman et al. (2011b) demonstrate that changes in ENSO character, such as between eastern and central Pacific events, do not require fundamental changes in the dynamics of ENSO and may instead result from the natural variability in the projection of the forcing onto these modes.

These previous LIM based studies of ENSO have used the observational record since the 1950s in the generation of the LIMs, implicitly assuming that the dynamics of ENSO have not changed over this period. However, the observational record of tropical SST is suggestive of a significant change in the character of ENSO since the mid to late 1970s concurrently with the change in the mean state of the climate system (An and Jin 2000). Indeed, application of a two-sided Kolmogorov-Smirnov test to the 20 year time-series' of Nino3 index either side of the climate shift allows rejection at the 99 % confidence level of the hypothesis that the distribution of ENSO amplitude has not changed. The two decades following the climate shift saw an increase in ENSO variance (eg McGregor et al. 2013), a change in the leading ENSO mode (An et al. 2006), a decrease in their frequency, and an apparent shift in the propagation direction of SST anomalies along the equator, from westwards pre-1970s to eastwards post-1970s (Fedorov and Philander 2000; Wang and An 2002; Trenberth et al. 2002; McPhaden and Zhang 2009; Aiken et al. 2013; Santoso et al. 2013). Through analysis of reduced physics analytical and numerical models of the coupled tropics, it has been demonstrated that such changes in ENSO characteristics appear to stem from a change in the mean state of the Tropical Pacific climate and hence in the stability of the system's coupled modes (e.g. An and Jin 2000; Fedorov and Philander 2001). In particular, the deepening of the eastern Pacific thermocline over recent decades, resulting from the mean weakening of the trades, favors the longer period delayed oscillator

mode and decreases the growth rate of the shorter period SST mode (Neelin et al. 1998; An and Jin 2000; Fedorov and Philander 2001). However, while the observed mean state changes would support a change in the underlying dynamics of coupled modes in the tropics, it is still possible that at least some of the changes in ENSO character may be simply a consequence of variability in the stochastically forced system and the relatively short observational record (Kirtman and Schopf 1998; Thompson and Battisti 2001; Fedorov and Philander 2000; McPhaden et al. 2011; Newman et al. 2011b). Given the high likelihood of a continued evolution in the mean state of the tropical Pacific over coming decades (Vecchi et al. 2006; Santoso et al. 2013), significant interest remains in understanding the sensitivity of ENSO to climate shifts. In contrast to previous studies, here we analyse the linear dynamics of ENSO derived empirically from the 20 year periods prior to and following the late 1970s climate shift, finding that the ENSO modes and their optimal precursors changed. This suggests that the trajectory of ENSO development, and hence ENSO predictability, may also change in the future.

In the following we provide a brief overview of linear inverse modelling and describe the generation of the LIMs corresponding to the periods pre- and post-1978. We discuss the spectral properties of these LIMs and then investigate the mechanisms responsible for the generation of ENSO in each case, with a special focus on the external forcing associated with remote oceanic basins.

2 Linear inverse models

2.1 LIM overview

The application of linear inverse models to ENSO has been thoroughly developed by Penland and co-workers (Penland and Matrosova 1994; Penland and Sardeshmukh 1995). The principal assumption underlying the LIM is that ENSO variability may be reasonably modelled as a linear system driven by additive Gaussian white noise of the form

$$d\mathbf{x}/dt = \mathbf{B}\mathbf{x} + \xi(t), \tag{1}$$

where $\mathbf{x}$ is the state vector, $\mathbf{B}$ is an autonomous (time-independent) linear operator and $\boldsymbol{\xi}$ is a Gaussian white noise process. Knowledge of the covariance properties of the observations of $\mathbf{x}$ allows the best fit estimate of $\mathbf{B}$ in the least squares sense to be determined with

$$\mathbf{B} = \tau_0^{-1} \ln(\mathbf{C}(\tau_0)\mathbf{C}(0)^{-1}), \tag{2}$$

where $\mathbf{C}(\tau_0)$ and $\mathbf{C}(0)$ are the covariance matrices at lags of τ_0 and zero respectively. Under these circumstances the fluctuation–dissipation relationship reduces to

$$\mathbf{B}\mathbf{C}(0) + \mathbf{C}(0)\mathbf{B} + \mathbf{Q} = 0, \tag{3}$$

where $\mathbf{Q}$ is the product of the white noise covariance matrix and the sampling period. An eigenvalue decomposition of $\mathbf{Q}$ yields the coherent spatial structures that span the temporally white forcing, here called the forcing orthogonal functions, of FOFs.

Penland and Sardeshmukh (1995) present a justification for why we may expect this to be a reasonable assumption for ENSO, and offer a series of tests that may be applied to assess the validity of the assumption. At a minimum we may expect that the LIM is able to produce skillful forecasts of the system state $\mathbf{x}$ and reproduce the lagged covariance matrix $\mathbf{C}(\tau)$ via

$$\mathbf{x}(\tau) = \exp(\mathbf{B}\tau)\mathbf{x}(0), \tag{4}$$

and

$$\mathbf{C}(\tau) = \exp(\mathbf{B}\tau)\mathbf{C}(0). \tag{5}$$

Given the limited observational record and the ubiquitous presence of observational errors, it is convenient to operate in the reduced state space spanned by just the leading EOFs. While Penland and co-workers found that the first 10 to 15 EOFs of SST were adequate to construct LIMs that recover much of the ENSO variability, Newman et al. (2011a) demonstrated that a significant improvement in model skill on lead times beyond 9 months could be achieved by extending the state vector to also include information of the surface wind stress and thermocline anomalies. Here, as in Aiken et al. (2013), we follow Newman et al. (2011a) and consider a state vector of the form

$$\mathbf{x} = [T_0 \quad Z_{20} \quad \tau_x], \tag{6}$$

containing the projections of the anomalies onto the first 13 EOFs of sea surface temperature (T_0), the first 7 EOFs of 20 C isotherm depth (Z_{20}), and the first 3 EOFs of zonal wind stress (τ_x). The EOF truncation retains over 72, 42 and 33 % of the variance in T_0, Z_{20} and τ_x, respectively.

LIMs were generated for the periods 1959–1978 and 1979–1998, corresponding to the 20 years either side of the 1970s climate shift, and for convenience are hereafter referred to as LIM1 and LIM2, respectively. The relatively short interval used is dictated by the limitations of the observational record on one hand (1958 taken to mark the commencement of robust SST monitoring), and on the other the phase of the dominant mode of decadal climate variability (ie the IPO) that, through changes in the background climate state, may possibly modulate ENSO dynamics. The time-series used to generate LIM1 (LIM2) coincide with a negative (positive) phase of the IPO. Notwithstanding the shortness of the record used in their generation, the LIMs possess significant cross-validated forecast skill that is superior to that of the LIM formed from

the combined period (Aiken et al. 2013). We choose to use a common basis for both LIMs, constructed from the EOFs for the period 1959–1998, after application of a 3 month running mean. The use of a common basis facilitates intercomparison at the cost of a slight decrease in the skill of each LIM. The results that follow do not otherwise substantially change if the EOF bases are determined separately for each period.

2.2 LIM generation

Observations for the period 1959–1998 were taken from the International Comprehensive Ocean–Atmosphere Data Set Project (ICOADS, Woodruff et al. 2011) for T_0 and from ERA-40 (Uppala et al. 2005) for τ_x. Given the sparsity of historical observations of upper ocean heat content, Z_{20} was calculated from the Simple Ocean Data Assimilation version 2.1.6 global ocean reanalysis (Carton and Giese 2008; Giese et al. 2010). SODA employs a sequential approach in which the POP model was first forced by ERA-40 winds and available freshwater flux data to provide a first guess, followed by the assimilation of ICOADS data via a Kalman filter. In this way the SODA-derived thermocline depth may be expected to be consistent with the surface wind and temperature fields. All data were subject to a 3 month running mean prior to removing the seasonal cycle. These filtered anomalies were then consolidated into bins of 2° latitude by 5° longitude within the region between the 25°N and 25°S parallels, and normalised by the global mean RMS anomaly. The resulting binned and normalised anomalies were then projected onto the state vector (6), and used to generate the linear inverse model **B** via (2). The covariance time-scale τ used for calculating the LIMs was 3 months in both cases. As discussed in Penland and Magorian (1993) the estimate of the FOFs degrades with increasing τ. Nonetheless, the inferred dynamical operator **B** proved to be relatively insensitive to values of τ ranging from 3 to 9 months.

A significant degree of uncertainty exists in the trends present in historical observational datasets due to sparse sampling and changes in observational methods, and data extrapolation techniques (e.g., Deser et al. 2010), over time. The apparent change in ENSO character around 1977 also coincides with the sudden increase in the quantity and quality of SST observations from the tropics heralded by the satellite era (Tokinaga et al. 2012). Because in this study the mean and seasonal cycle are removed independently from each period, the analysis that follows is not sensitive to any such artifacts in the observational record. Only observational errors that affect the covariance will degrade the LIMs.

Notwithstanding the above, for reference the change in the mean states of the three variables between the 1959–1978 and 1979–1998 periods are shown in Fig. 1. The post-1977 period is characterised by surface warming over the central to eastern Pacific, a shoaled thermocline in the western Pacific, and weakened trade winds in the central Pacific. These trends are consistent with the recent analysis of Tokinaga et al. (2012). As discussed by Fedorov and Philander (2001), the flattening of the equatorial Pacific thermocline tends to favour a longer period ENSO. The lagged SST autocovariance along the equator (Fig. 2) reveals the tendency for ENSO events to decorrelate more quickly prior to the climate shift, consistent with the weaker and shorter ENSO events pre-1970s.

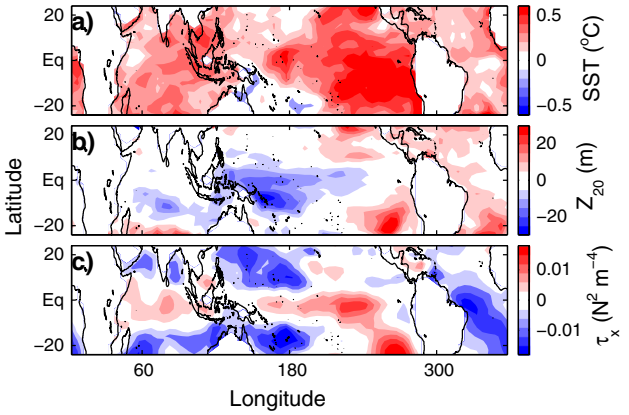

Fig. 1 Change in observed **a** SST, **b** 20 °C isotherm depth, and **c** zonal wind stress, from the periods 1959–1978 to the period 1979–1998

2.3 Validity of the LIMs

Previous studies have demonstrated that the linear inverse model assumptions are often valid for tropical ocean-atmosphere variability (Penland and Sardeshmukh 1995; Alexander et al. 2008; Newman et al. 2009). Penland and Sardeshmukh (1995) outline a series of tests with which to quantify the validity of a LIM, and demonstrate that their SST-only LIM of the equatorial Indo-Pacific satisfies these tests to an acceptable degree. In particular, the LIM of Newman et al. (2011a), upon which ours is based, has also been shown to be a valid model for ENSO. Here we concentrate on the ability of LIMs to reproduce the observations, while noting that the other criteria discussed in Penland and Sardeshmukh (1995) are also satisfactorily met.

The cross-validated forecast skill was determined via a jackknife procedure, involving the calculation of the correlation between the observed SST for each consecutive 2 year period and the forecast for that same period by the LIM generated using the remaining 18 years of data. The so-determined forecast skill of each LIM (Fig. 3) beats persistence and climatology forecasts throughout the tropics on

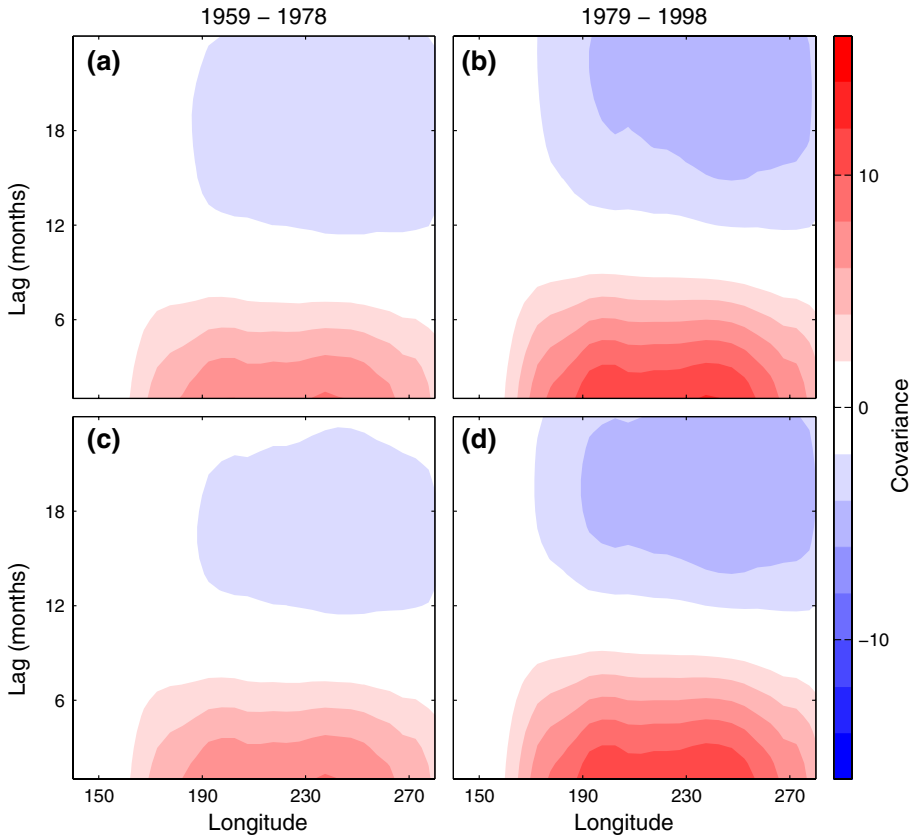

Fig. 2 Autocovariance of observed equatorial SST versus lag for the periods **a** 1959–1978 and **b** 1979–1998 ; **c**, **d**, as in (**a**, **b**) for the linear inverse models

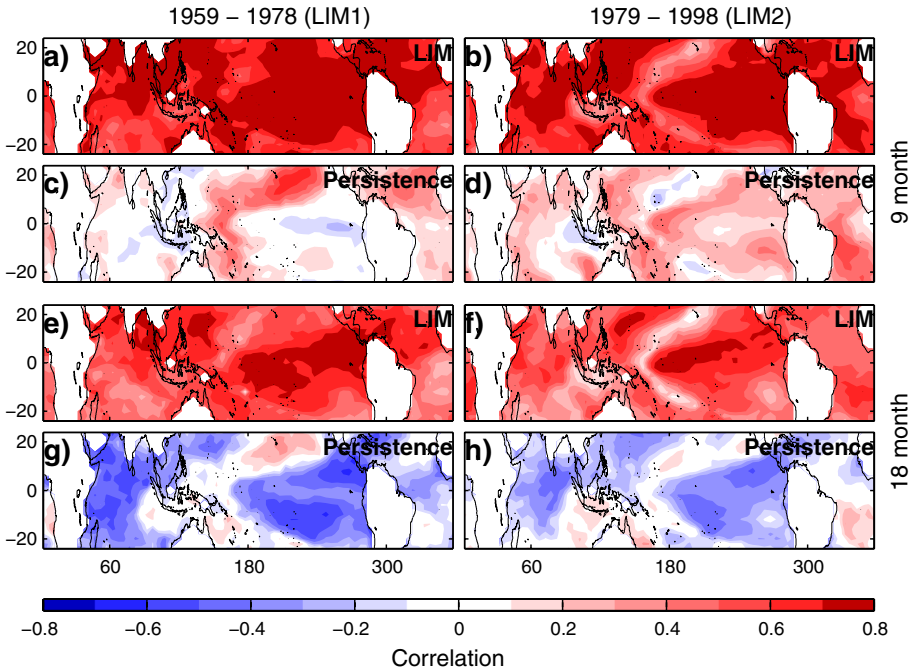

Fig. 3 Spatial structure of the cross-validated skill of each LIM compared to persistence on forecast lead times of 9 and 18 months

lead times past 18 months. As shown in Aiken et al. (2013) (their Fig. 1), the RMS error in forecasts of the principal component of the first EOF of SST (i.e. the first element of the state vector) is also substantially better than both persistence and climatology for lead times up to 2 years for both LIMs. In fact, it is shown below that each LIM's optimal perturbation can, by itself, explain over three quarters of the ENSO variance on a 7 month lead time.

Both LIMs also reproduce the equatorial Pacific lag covariance to a excellent degree (Fig. 2). The equatorial SST autocovariance predicted by (5) closely matches that of the observations at lags out to 2 years, including the tendency for equatorial Pacific SST to decorrelate more slowly in the eastern (western) Pacific pre (post) 1977. The fact that the LIMs derived from each period are able to reproduce the observations suggests that they capture the principal dynamics of ENSO and hence can represent useful analogues with which to investigate the real system.

3 Properties of the LIMs

In the absence of stochastic forcing, the solutions to (1) are found through eigenanalysis of the LIMs, yielding the spectrum (eigenvalues) and modes (eigenvectors) corresponding to each period. The spectrum is illustrated by the dark dots in Fig. 4e, f. In order to gauge their sensitivity to the relative shortness of the records used to generate each LIM, the eigenspectrum was recalculated for each of the jackknife estimates of each LIM. While the various estimates of the spectra differ (Fig. 4e, f, gray dots), in each case two eigenmodes exist with a spatial structure and evolution that are highly reminiscent of ENSO (Fig. 4e, f, large circle and square), but with differing periods of approximately 40 and 24 months. The evolving structure of these two modes (Figs. 5, 6), referred to hereafter as the "slow ENSO mode" and "fast ENSO mode", respectively, in reference to their natural periods of oscillation, are consistent with the classical Bjerknes type mechanism (Bjerknes 1969). In each of the ENSO modes the basin-wide equatorial thermocline deepens at the initial state of developing El Niño, followed by a shoaling (deepening) of the thermocline in the west (east) enhanced by anomalous westerlies which are further reinforced by the eastern Pacific surface warming. Note, however, that the extent of the equatorial Pacific thermocline anomaly prior to the mature phase is larger in the slow ENSO modes (Fig. 5a, c) than in the fast ENSO modes (Fig. 6a, c). The RMS size of thermocline anomalies in the slow modes is approximately 12 and 33 % more than is the case for the fast modes of LIM1 and LIM2, respectively. This suggests that, consistent with its longer period, the slow ENSO mode is representative of the recharge-discharge oscillator paradigm, in which a recharge (discharge) of the equatorial Pacific warm water volume acts as a necessary precursor to El Niño (La Niña) development (Jin 1997; Meinen and McPhaden 2000). The fast modes, on the other hand, are characterised by a more westerly maximum in SSTA, and a greater relative importance of anomalous SST and zonal wind stress anomalies relative to thermocline, and may be identified with the SST mode (Neelin et al. 1998; An and Jin 2000; Fedorov and Philander 2001). The slow ENSO mode is the least damped eigenmode in each LIM, and as a result would be expected to dominate the system's response to forcing. Although strongly damped, the fast ENSO modes nonetheless play a fundamental role in ENSO growth, as discussed below.

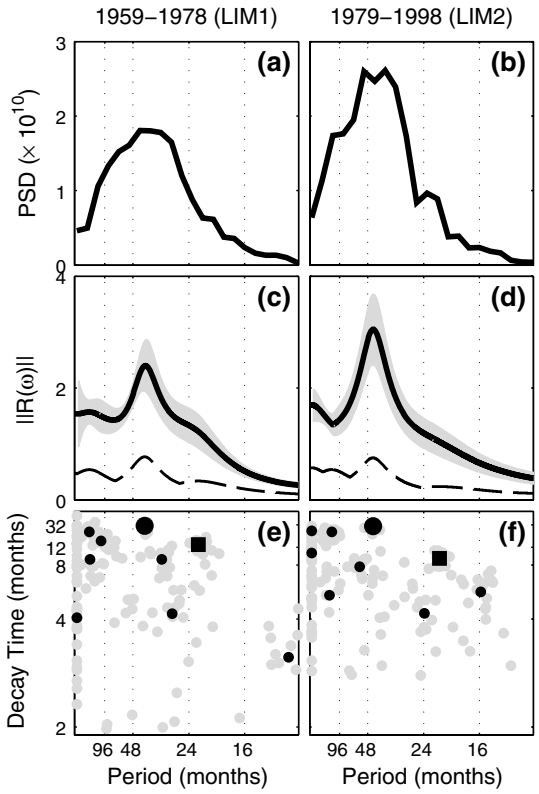

Fig. 4 **a**, **b** Power spectral density of the principal component of the first EOF of observed SSTA for the period indicated. **c**, **d** The value of the resolvent norm at zero real frequency (that is, purely harmonic forcing of frequency ω, *solid line*), the spread in jackknife estimates (*shaded*), and the corresponding resonant response curve (*dashed line*) of each LIM. **e**, **f** The spectrum of each LIM (*dark small dots*). The eigenvalue corresponding to the slow and fast ENSO modes are indicated by the *larger circle* and *square*, while faint *dots* show the ensemble of spectra from the jackknife estimates. Eigenvalues with negative imaginary frequency are not shown, but are simply the complex conjugates of those shown

From LIM1 to LIM2 the period of the slow ENSO mode increases from 40 ± 1 to 44 ± 2 months, consistent with the observed change post-1977 (Fedorov and Philander 2000). There is no significant change in the decay time scale of the slow ENSO mode, however, despite the observed increase in ENSO variance following the climate shift. The fact that ENSO amplitude can increase without a significant change in the ENSO mode decay rate owes to the essentially non-modal nature of ENSO growth. As previously described in Penland and Sardeshmukh (1995), the ENSO variance sustained through external forcing does not depend simply upon the modal decay rates alone, but also upon the

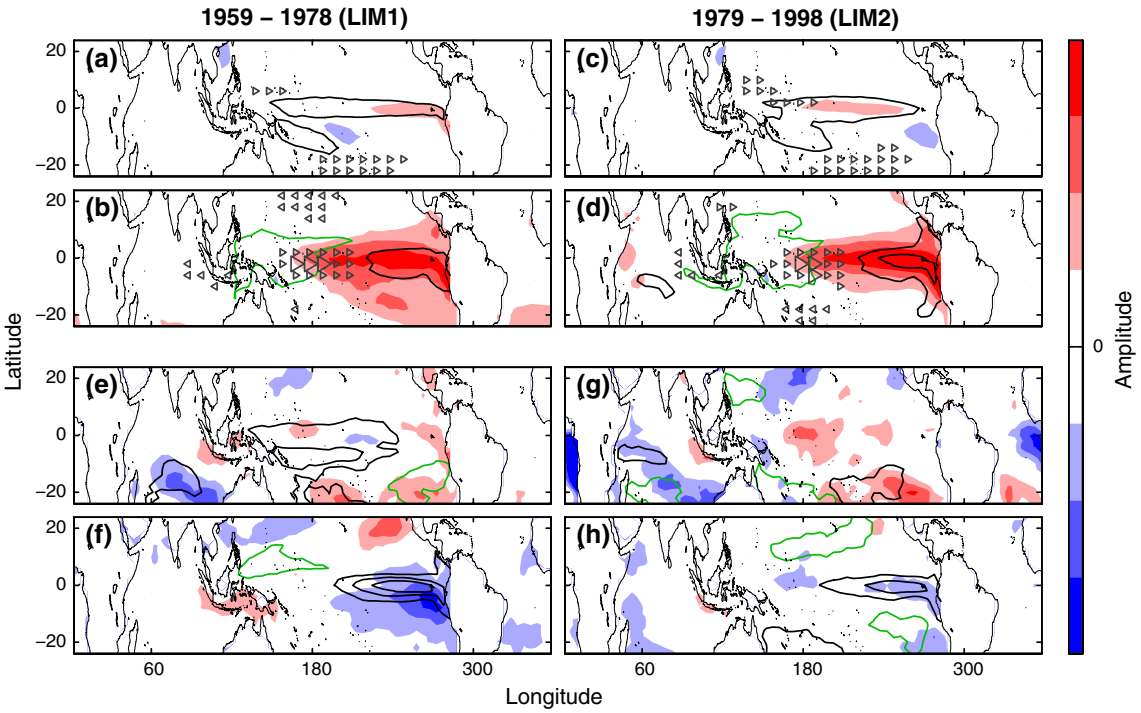

Fig. 5 Quadrature phases of the slow ENSO modes for **a**, **b** LIM1, and **c**, **d** LIM2. The quadrature phases of the corresponding adjoint modes are shown in (**e**, **f**) and (**g**, **h**), respectively. T_0, Z_{20} and τ_x are indicated by *shading*, *contours* and *arrowheads*, respectively. The interval spacing in each case is a constant fraction of the normalised amplitude. Each *shading*/*contour*/*arrow* level represents an identical fraction of the variance of $T_0/Z_{20}/\tau_x$. *Black* (*green*) *contours* indicate positive (negative) thermocline anomalies. The evolution of each pair of quadrature phases proceeds as (*upper*), (*lower*), -(*upper*), -(*lower*), (*upper*), etc

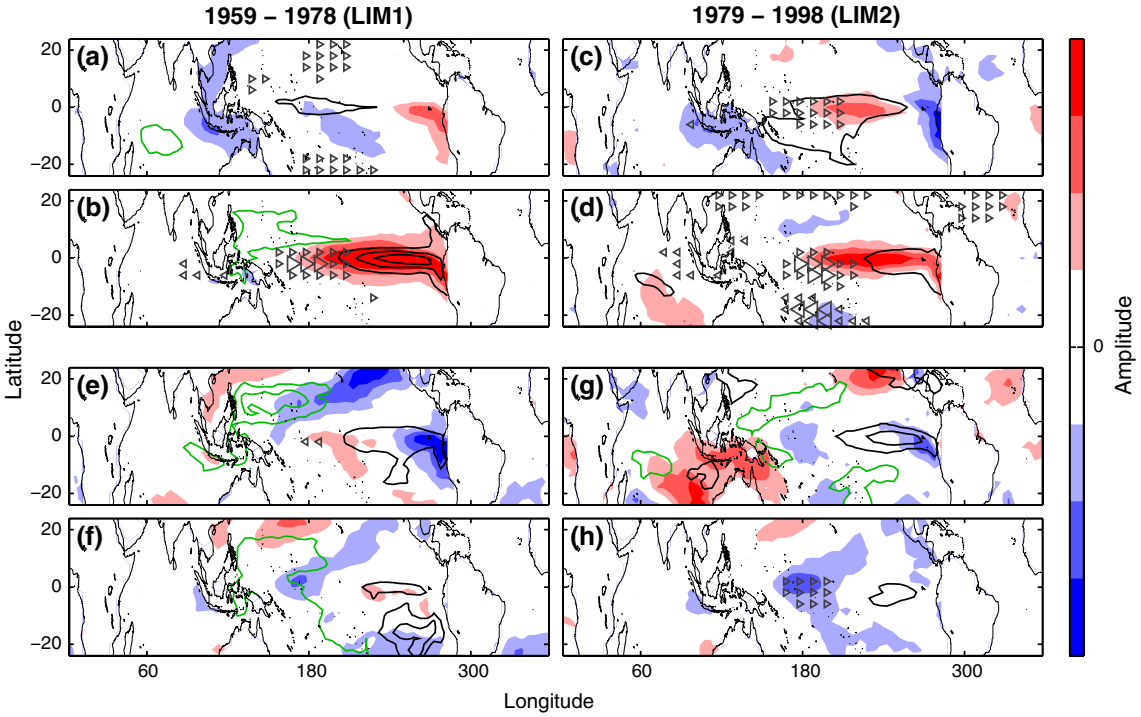

Fig. 6 As in Fig. 5, but for the fast ENSO modes

degree of linear dependence between the modes. The magnitude of the response of a dissipative linear system to a forcing at complex frequency ω (that is, with harmonic and exponential components) is bounded above by the norm of the resolvent $R(\omega) = (\mathbf{I}\omega - \mathbf{B})^{-1}$, where $\mathbf{I}$ is the identity matrix and the matrix norm is evaluated by the maximum singular value. The resolvent norm satisfies the following inequality

$$\frac{1}{\text{dist}(\omega, \Lambda)} \leq \| R(\omega) \| \leq \frac{\kappa}{\text{dist}(\omega, \Lambda)}, \tag{7}$$

in which the denominators indicate the Euclidean distance in the complex plane of ω from the nearest member of the eigenspectrum Λ, and κ is the condition number of the matrix having the eigenvectors as columns. Inequality (7) illustrates that the system response to harmonic forcing can exceed resonance (given by the first term and corresponding to the excitation of the eigenmode with eigenfrequency closest to the forcing frequency) whenever $\kappa > 1$. This condition ensues when the eigenvectors are not mutually orthogonal; that is, when the dynamical system is nonnormal. It is the linear interference between the modes that allows an enhanced system response to harmonic forcing, a phenomenon referred to as a pseudoresonance (Trefethen et al. 1993).

In each LIM the resolvent norm $\| R(\omega) \|$ evaluated at real forcing frequencies ω (Fig 4c, d) possesses local maxima that coincide with the frequency of the slow ENSO mode and with that of the maximum in the power spectral density of the principal component corresponding to the first EOF of observed SST. While the LIM2 slow ENSO mode is slightly more damped than that of LIM1, the associated resolvent norm maximum is greater, consistent with the observed increase in ENSO amplitude. As such the resolvent norm reconciles the model eigenspectrum with the spectrum of the system's forced response. The resolvent norm is larger than the response that would be expected from unimodal resonance (dashed curves Fig 4c, d) at all frequencies, and especially so at those corresponding to the slow ENSO modes. This indicates that the observed large system response at ENSO frequencies is consistent with a pseudoresonance and is inherently nonmodal in nature.

Contours of $\epsilon^{-1} = \| R(\omega) \|$ in complex frequency space mark the boundaries of the ϵ-pseudospectra (Trefethen 1991). These indicate the sensitivity of the spectrum to changes in the operator $\mathbf{B}$ itself (as opposed to perturbations of the state vector $\mathbf{x}$). $\epsilon(\omega)$ can be interpreted as the necessary size of the change to the system dynamics such that ω can become part of the spectrum. That is, frequencies at which the resolvent norm is large require only a relatively small change to the system dynamics in order to become part of the spectrum. Thus the fact that the resolvent norm was found to be large in the vicinity of the slow ENSO mode frequency suggests that changes in the underlying dynamical system, as could be anticipated from an altered background state, may influence most strongly the frequency and decay rate of this mode. That is, significant variability in the complex frequency of ENSO could be expected from only small changes in the background state. A corollary is that observational errors may also have a relatively large, spurious effect on the diagnosed frequency of ENSO. This may be especially relevant to estimating ENSO frequency from weakly constrained state estimates, and explains the broad spectrum of ENSO frequency estimates derived from a stochastically forced model with constant dynamics in Aiken et al. (2013).

The pseudoresonance at the slow ENSO mode frequency involves linear interference between a number of nonorthogonal modes. The degree of nonorthogonality may be quantified by consideration of the adjoint operator $\mathbf{B}^{\dagger}$ (where $\dagger$ indicates the Hermitian transpose). Decomposition of a perturbation into its component modes requires calculation of each mode's *biorthogonal*—the structure that is perpendicular to all but that mode. The modes of the adjoint operator $\mathbf{B}^{\dagger}$ (where $\dagger$ indicates the Hermitian transpose) possess the required biorthogonality condition, and hence are relevant for understanding the forcing of the modes. For each mode a corresponding adjoint mode exists that is orthogonal to all other modes and hence can be interpreted as being the optimal forcing of that mode. If the LIMs were normal (with all eigenmodes mutually orthogonal) then the ENSO modes would be identical to their corresponding adjoint modes. Conversely, disparity between mode and adjoint mode indicates linear dependence of the mode, and nonnormality of the operator. This is the case for each of the ENSO modes (Figs. 5, 6). Such linear dependence opens the possibility for perturbations involving the ENSO modes to grow transiently, even though they must subsequently decay asymptotically. Although the structures of the ENSO modes from each period are broadly similar, more significant differences exist among the corresponding adjoint modes. Nonetheless, the adjoint ENSO modes do share a number of common elements, including a relatively weak wind stress component and an anti-correlation between SST and Z_{20} anomalies in the central and eastern Pacific. This relationship between SST and Z_{20} is opposite to that associated with ENSO (in which SST and thermocline covary positively). Nonetheless, both of these common elements of the adjoint ENSO modes reappear in the optimal ENSO precursors discussed in sequel.

4 Optimal linear ENSO growth

Although all eigenmodes of both LIMs have negative real eigenvalues and hence decay monotonically, the fact that the modes are not mutually orthogonal allows for the transient growth of disturbances. As discussed above for purely harmonic forcing, when transient growth is possible

the system variance under stochastic forcing will exceed that expected from consideration of the modal decay rates alone (Farrell and Ioannou 1996). While the asymptotic behavior of a perturbation to the system will always follow the decay rate of the least damped mode, the growth rate of the transient phase is highly dependent upon the initial disturbance. The perturbation $\mathbf{x}_t$ that maximises a generalised growth factor $\mu = \| \mathbf{G}\mathbf{x}(t) \| / \| \mathbf{H}\mathbf{x}(0) \|$ over time t is given by the first singular vector of $\mathbf{G}\mathbf{R}\mathbf{H}^{-1}$, where $\mathbf{G}\mathbf{G}^T$ and $\mathbf{H}\mathbf{H}^T$ define quadratic norms and the propagator $\mathbf{R}(t) = \exp(\mathbf{B}t)$. In the following $\mathbf{G}$ and $\mathbf{H}$ are chosen such that $\mathbf{x}_t$ maximises the projection onto the first EOF of SST relative to the initial perturbation size as evaluated in the L2 or Euclidean norm, equivalent to the RMS of the state vector. In this case, the singular value can be interpreted as the magnitude of ENSO at time t when the initial perturbation has unit length in the Euclidean norm. Hence all basis elements are penalised equally in the optimisation problem. While the first EOF of SST is the first element of the state vector, and hence a convenient metric for ENSO amplitude, the results that follow are, in fact, qualitatively robust when other $\mathbf{G}$, such as a Niño3 or L2 norm, are used instead.

4.1 Optimal growth over 7 months

The calculated maximum linear growth factors reveal that each LIM supports perturbations that can grow over timescales beyond 2 years (Fig. 7). For our choice of norm, positive growth factors indicate that the final projection onto SST EOF1 exceeds the initial perturbation size as measured with the L2 norm. In each case, maximum growth occurs over a timescale of approximately 7 months. The slightly faster transient growth in LIM2 mirrors the enhanced pseudoresonance associated with the slow ENSO mode of LIM2, and is consistent with the larger amplitude ENSOs post-1977. The structure of the perturbations responsible for maximum growth in each period (referred to hereafter as the 7 month optimal perturbations, or $\mathbf{OP7}_i$ for $\mathbf{LIM}i$) possess a number of significant differences (Fig. 8), and inspection of the spread of the jackknife estimates of each OP7 in state space demonstrates that these differences are significant (Fig. 9). Some common elements do exist of optimal ENSO growth in each LIM, such as a deepening of the western Pacific thermocline, reminiscent of the "east–west" mode of Meinen and McPhaden (2000), and surface warming in the eastern and western equatorial Pacific. Major differences between the two OP7s include the inferred sensitivity to cooling in the Atlantic, south Indian and the Maritime Continent sectors in $OP7_2$, and the greater importance of eastern Pacific SST and western Pacific thermocline disturbances in $OP7_1$. It may be noted that SST anomalies in the Atlantic and Indian Oceans occur exclusively in LIM2. The positive SSTA pattern traversing the northern tropical Pacific

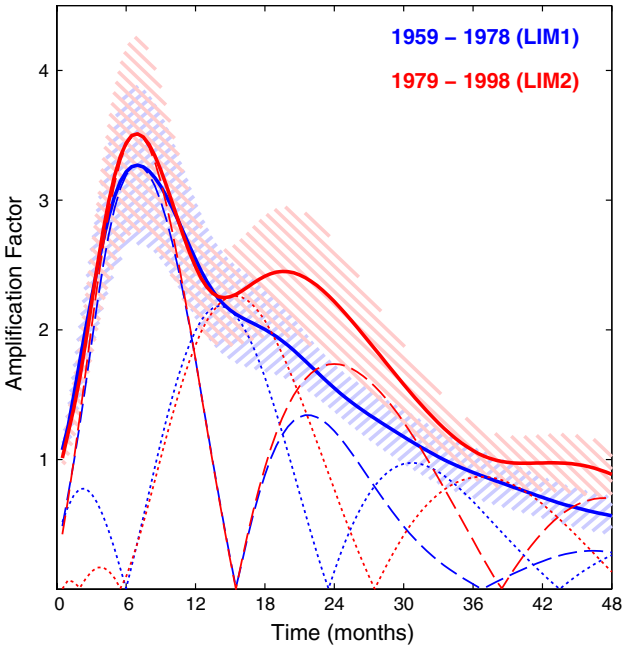

Fig. 7 Maximum linear growth curves for 1959–1978 (LIM1, *blue*) and 1979–1998 (LIM2, *red*). The evolution of the amplitudes of the 7 (15) month optimal perturbations are marked by the *dashed* (*dotted*) curves. The spread in the jackknife estimates of the maximum amplification factor is *shaded*

from Indonesia to California in $OP7_1$ was also found in the 7 month optimal perturbation of Penland and Sardeshmukh (1995), and strongly resembles the North Pacific Meridional Mode (Vimont et al. 2003). It is also notable that the contribution of zonal wind stress forcing to both OP7s is neglible. This implies that the component of the wind variability resolved within the LIM plays little role in ENSO generation. We note, however, that this does not necessarily contradict the role of westerly wind bursts in triggering certain El Niño events, given that much of the associated variability is removed from the LIM by the EOF truncation.

Notwithstanding the differences between the OP7s, their subsequent evolution follows a similar trajectory in both LIMs (Fig. 8). In both cases within the first month anomalous westerlies develop above the western equatorial Pacific positive SST anomalies, consistent with the beginning the Bjerknes feedback. The following months witness the concurrent intensification of eastern Pacific warming, western Pacific cooling, and westerly wind stress anomalies in the equatorial and southern Pacific. After 7 months each perturbation has evolved into a mature El Niño event, including a slight southward shift of the central Pacific westerlies. The development of the OP7s reproduce the post-1977 change in the direction of SST phase propagation (Santoso et al. 2013): in LIM1 the eastern Pacific SST anomaly expands westwards, while in LIM2 central Pacific warming is synchronous with that in the east.

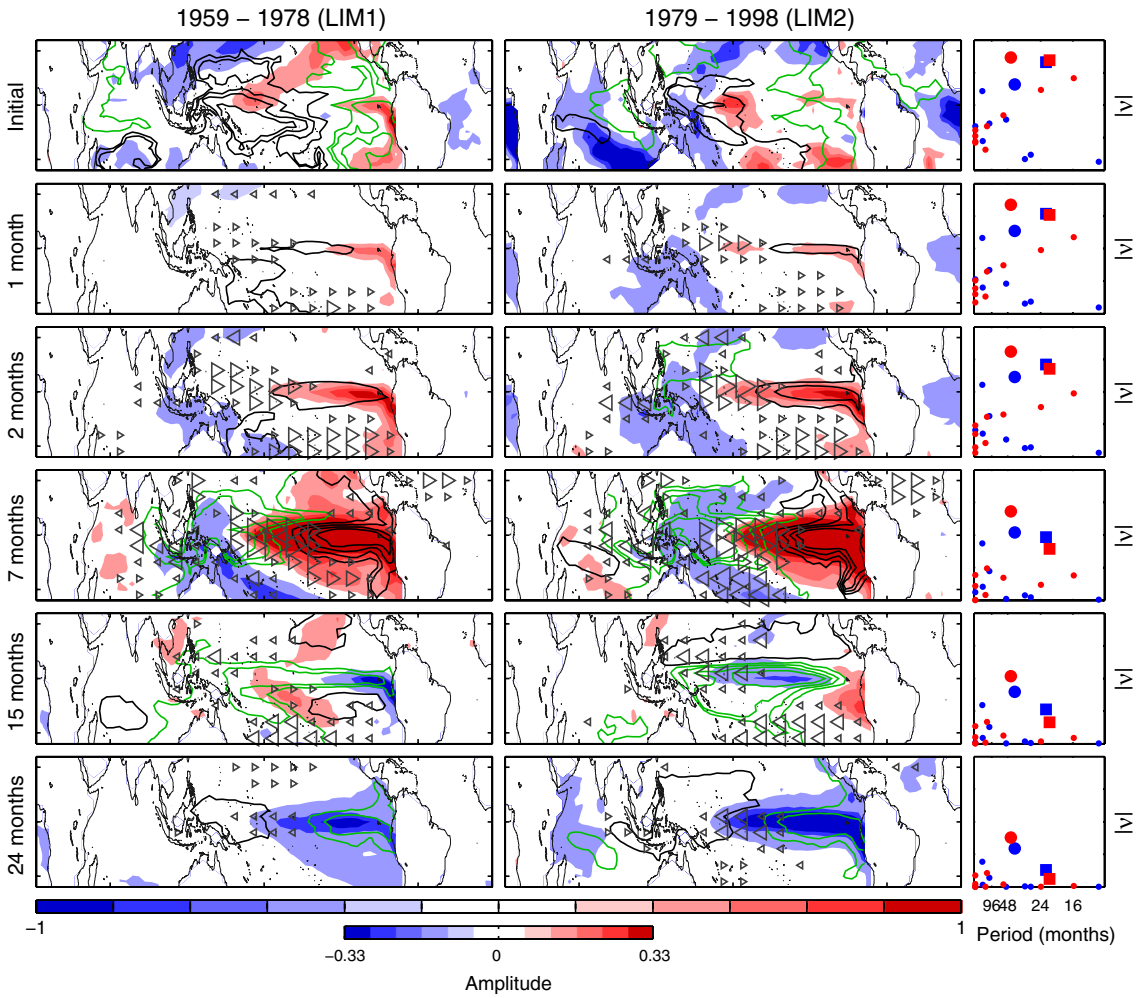

Fig. 8 Initial structure and evolution of the 7 month optimal perturbations in each LIM. T_0, Z_{20} and τ_x are indicated by *shading*, *contours* and *arrowheads*, respectively. *Black* (*green*) *contours* indicate positive (negative) thermocline anomalies. The interval spacing in each variable is an arbitrary constant fraction of the normalised amplitude. That is, each *shading*/*contour*/*arrow* level represents an identical fraction of the variance of $T_0/Z_{20}/\tau_x$. The initial structure of the optimal in each case is displayed with interval spacing reduced by a factor of three, as indicated by the *lower colour-scale* at the *bottom* of the figure. The *right hand panels* illustrate the evolution of the projection $|\nu|$ of the perturbation onto each of the modes of LIM1 (*blue*) and LIM2 (*red*). The *larger circle* and *square* indicate the values of $|\nu|$ for the slow and fast ENSO modes, respectively

As discussed above, each adjoint mode represents the structure that projects exclusively onto the corresponding mode, and in general does not coincide exactly with the mode itself. Thus, when the modes do not form an orthogonal basis, the projection ν_i of the optimals onto the i-th mode must be determined via the biorthogonality relationship with the adjoint modes,

$$\nu_i = \frac{< \mathbf{a}_i, \mathbf{p} >}{|\mathbf{a}_i||\mathbf{p}|}, \tag{8}$$

where $< ., . >$ indicates the inner product, $\mathbf{a}_i$ is the adjoint mode corresponding to the ith mode, and $\mathbf{p}$ is the optimal perturbation. The evolution in the values of $|\nu_i|$ during optimal growth are depicted in the right hand panels of Fig. 8, illustrating how the modes interact to produce transient ENSO growth despite their monotonic decay. The initial strong projection of each OP7 onto the slow ENSO modes (large dots in right hand panels of Fig. 8) is "concealed" by destructive interference with other nonorthogonal modes, and in particular with the fast ENSO modes (squares in the right hand panels of Fig. 8). The more rapid decay of these "concealing" modes causes the destructive interference to subside over time, allowing the slow ENSO mode to emerge over the following 7 months.

4.2 Optimal growth over 15 months

The maximum amplification curve for the LIM of Newman et al. (2011a) contained a secondary local maximum at approximately 18 months, suggestive of the existence of a

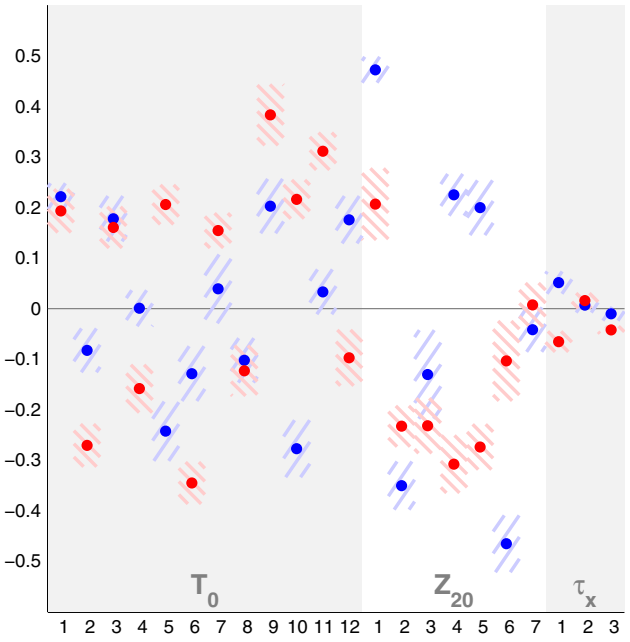

Fig. 9 The structure of the 7 month optimals, highlighting their significant differences, is illustrated in state space. The value of each state vector element is given by the *dots*, and *colour shading* indicates the spread from jackknife estimates. *Blue* (*red*) corresponds to LIM1 (LIM2). The optimals are normalised to have unit length

second linear growth mechanism. While secondary maxima are also found here in both amplification curves, albeit at approximately 20 months (Fig. 7), these do not correspond to a separate transient growth process. Rather they owe to the natural oscillation of the slow ENSO modes producing a second, weaker, ENSO event of opposite sign to the first. This may be appreciated in the amplitudes of the OP7s as they evolve, which coincide closely with the local maxima in maximum linear growth rate (Fig. 7, dotted line). The relatively large inner products between the OP7s and the 20 month optimals, of 0.78 and 0.81 for LIM1 and LIM2, respectively, confirm that it is essentially the same initial structure that is responsible for the two local maxima in Fig. 7.

Nonetheless, the fact that each OP7 produces no growth at 15 months indicates that a distinct perturbation must be responsible for the non-zero value of the linear growth curve at that time period. The initial state of the 15 month optimal perturbations (hereafter $OP15_1$ and $OP15_2$) do indeed differ substantially from those of the OP7s (c.f. Figs. 8, 10) and from each other. The differences between the two OP15s are also robust amongst the jackknife estimates (Fig. 11). In summary, we find that each LIM possesses two distinct perturbation structures that can grow into ENSO events, and that these operate over 7 and 15 month time-scales, respectively. That OP7 and OP15 also control sustained ENSO variance is demonstrated in the following section.

Both $OP15_1$ and $OP15_2$ involve coincident surface cooling together with thermocline deepening in the eastern equatorial Pacific. This out-of-phase feature, also reflected in the adjoint ENSO modes (Figs. 5, 6), may represent the role of equatorial waves in returning the tropical Pacific towards a neutral state as a precondition for the development of the next ENSO. The associated dynamics may be similar to the mechanism proposed by Izumo et al. (2010), wherein anomalous warming of the eastern Indian (Indian Ocean Dipole, or IOD, in its negative state) during the boreal autumn produces downwelling Kelvin waves in the equatorial Pacific that acts to terminate a La Niña and, through zonal advection, initiate an El Niño that peaks the following winter. A negative IOD pattern is apparent in $OP15_2$ (Fig. 10), consistent with the Izumo et al. (2010) mechanism, and the negative Atlantic SST anomalies seen in $OP15_1$ could potentially initiate ENSO via a similar modulation of the Walker circulation (Rodríguez-Fonseca et al. 2009).

It may be noted that after 7 months, however, each OP15 has evolved into a structure strongly reminiscent of its corresponding OP7, and the subsequent development of mature ENSO events follows a very similar trajectory to that of the OP7s. As such, in each LIM, OP15 may be interpreted as the perturbation that evolves into, or excites, OP7. Moreover, significant growth of the OP15s only occurs over the final 7 months, after reaching a state represented by OP7s that immediately precedes commencement of the Bjerknes feedback (Fig. 7, dotted curves). That is, OP7 and OP15 represent distinct mechanisms to trigger the Bjerknes feedback. As will be discussed below, the existence of perturbations that can grow over a period of 15 months offers the possibility of ENSO predictability on 15 month lead times.

4.3 Response to stochastic forcing

In the following we demonstrate the role of the optimals in controlling (and hence predicting) ENSO variability under the linear paradigm in a series of extended stochastically forced simulations.

The fluctuation–dissipation relationship (3) yields the covariance matrix of the spatially coherent but temporally white forcing $\mathbf{Q}$. The eigenvectors of $\mathbf{Q}$, referred to here as forcing orthogonal functions or FOFs, represent the spatial patterns inferred to dominate the stochastic forcing, and the corresponding eigenvalues indicate the fraction of variance explained. If the assumptions underlying the LIM held exactly then $\mathbf{Q}$ should be positive definite—that is, its eigenvalues should all be positive. As the assumption of linearity is clearly not met exactly, it may be anticipated that some eigenvalues of $\mathbf{Q}$ may be negative, with corresponding FOFs that are spurious. Here the $\mathbf{Q}$ calculated from the

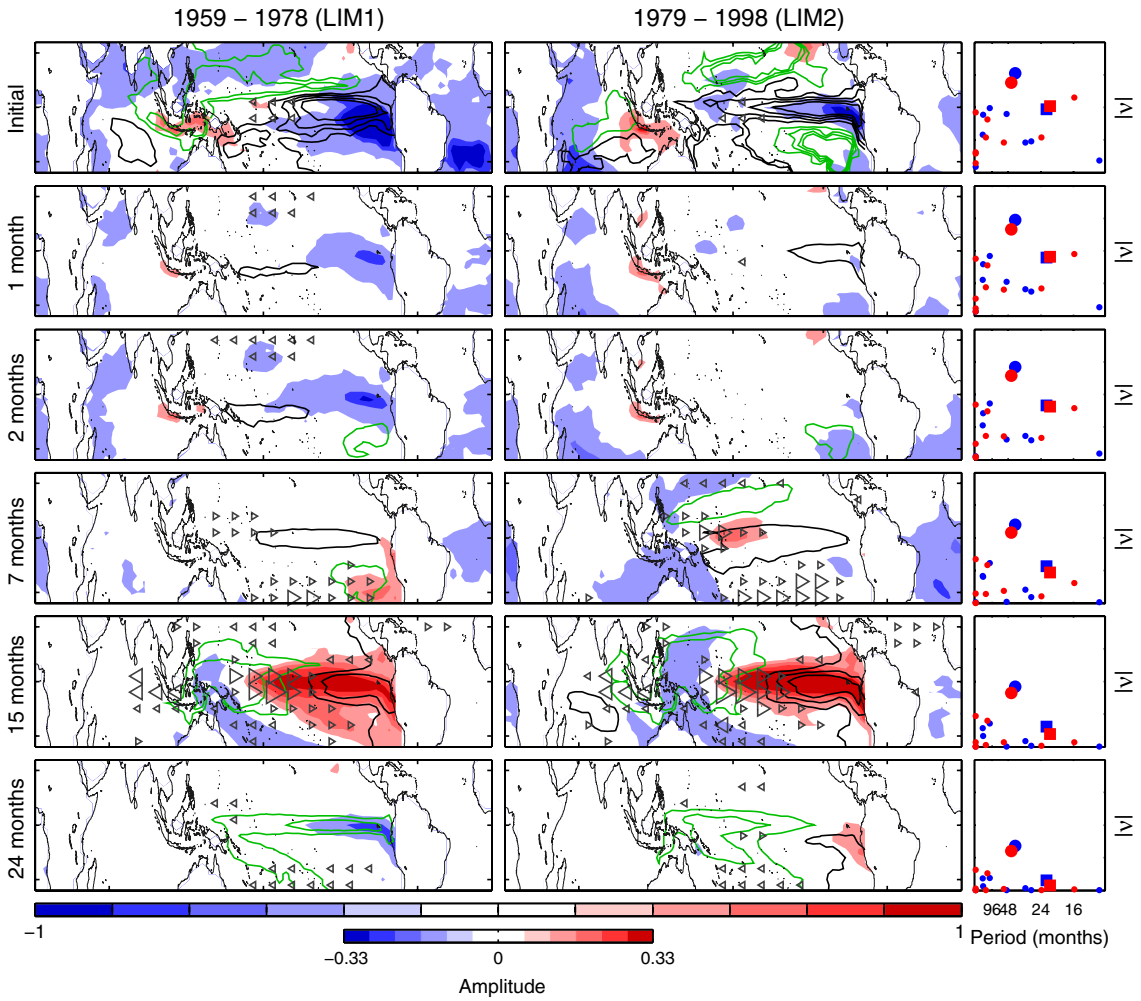

Fig. 10 As in Fig. 8 but for the 15 month optimal perturbations

jackknife estimates of **B** and corresponding to LIM1 and LIM2 were all found to have 2 or 3 negative eigenvalues of relatively small magnitude. In all cases wind stress fluctuations account for the vast majority of the white noise forcing variance, and over 40 % of the variance occurs in structures strongly weighted to the northern tropical Pacific. While the FOFs suggest that wind stress fluctuations make the dominant contribution to the external stochastic forcing, it was seen above that the disturbances responsible for transient ENSO growth contain a negligible contribution from zonal wind stress anomalies. This suggests that the spatial structure of the noise forcing in the real system is highly suboptimal, in the sense that most of the stochastic variability projects weakly onto the optimal perturbations.

To illustrate the importance of the spatial distribution of the white noise forcing, simulations of the stochastic differential Eq. (1) were performed using the scheme outlined in Penland and Matrosova (1994), wherein firstly the intermediate state vector $\mathbf{y}(t)$ is generated from

$$\mathbf{y}(t+\delta) = \mathbf{y}(t) + \delta\mathbf{B}\mathbf{y} + \mathbf{P}\Lambda\mathbf{r}(t)'\delta^{1/2}, \tag{9}$$

where the timestep $\delta = 6$ h, $\mathbf{P}$ is a matrix whose columns contain the orthonormal spatial forcing structures, Λ is a diagonal matrix defining the amplitude of each structure, and $\mathbf{r}(t)$ is a vector of independent Gaussian deviates of unit variance. The final state vector $\mathbf{x}(t)$ is then given by

$$\mathbf{x}(t+\delta/2) = (\mathbf{y}(t) + \mathbf{y}(t+\delta))/2. \tag{10}$$

Simulations were performed for each LIM using as columns of $\mathbf{P}$ either the FOFs of the corresponding period or normalised random vectors. So that the total forcing variance in each simulation was the same while still preserving the noise covariance, the elements of Λ were given by the square roots of the corresponding eigenvalues of $\mathbf{Q}$, normalised such that $\| \Lambda \| = 1$ in each simulation. Given that the system is linear the results are qualitatively insensitive to this normalisation. In the second case, the unitary forcing variance was distributed evenly between the random

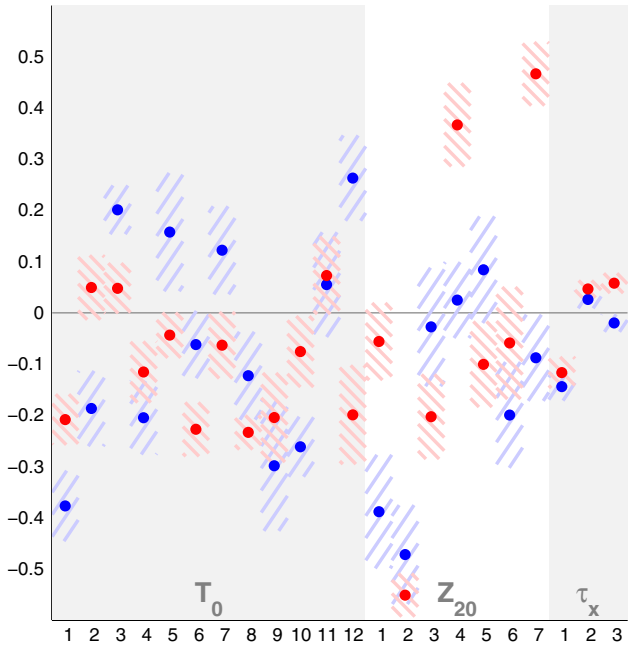

Fig. 11 As in Fig. 9 but for the 15 month optimal perturbations

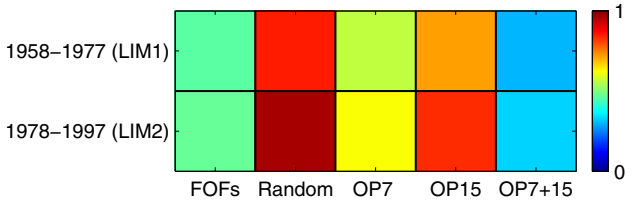

Fig. 12 The *colour* of each *box* indicates the standard deviation in the principal component of the first EOF of SSTA in stochastically forced simulations using each of the LIMs under different stochastic forcing scenarios. In each case the forcing variance was identical, but the spatial distribution of the forcing variance was determined by, from *left* to *right*: the forcing orthogonal functions of the corresponding period (FOFs); spatially random forcing (random); random forcing from which the projection onto the 7 month optimal perturbation has been removed (OP7); random forcing from which the projection onto the 15 month optimal perturbation has been removed (OP15); random forcing from which the projection onto both the 7 and 15 month optimal perturbations have been removed (OP7 + 15). The standard deviations have been normalised by a constant factor

vectors, and as a result each basis element received equal forcing variance. The FOFs with negative eigenvalue were removed. As noted above, these spurious FOFs explain very little of the forcing variance. In each simulation the model was spun up from rest for 10 years and then run for a further 20 thousand years. Monthly mean values of the state vector were recorded, and the model response assessed by calculating the standard deviation in the first state vector element, corresponding to the first EOF of SST. The results are illustrated in Fig. 12.

The system response to spatially random forcing (Fig. 12, second column) greatly exceeds that when the FOFs are used (Fig. 12, first column). This may be anticipated, as wind stress variations are the dominant component of the FOFs but appear only weakly in the OP, and therefore the spatially random forcing projects more strongly onto the optimally growing perturbations. Notwithstanding the differences in the spatial character of the forcing, in each case the frequency of maximum power spectral density remains close to that of the slow ENSO mode of the corresponding LIM. Even though this mode is slightly less damped in LIM1 (dynamics pre-1977), ENSO amplitude under spatially random forcing is significantly greater in the simulation suing LIM2 (dynamics post-1977), consistent with the stronger non-modal growth supported in LIM2.

The relationship between the simulated ENSO amplitude and the projection of the forcing onto the optimal perturbations is further illustrated by considering three additional scenarios, in which the random forcing case was manipulated to remove the projection of the forcing onto (1) OP7, (2) OP15, and (3) both. As this process removes degrees of freedom, the variance was redistributed evenly between the remaining structures such that the forcing variance was identical to the previous cases. Removal of both optimal structures reduces the simulated ENSO variability by over 80 % relative to the random forcing case (Fig. 12, fifth column). The reduction in the sustained ENSO variability is significantly less, however, if only one of the OPs is removed (Fig. 12, third and fourth columns). That is, when only one of OP7 and OP15 is filtered from the noise forcing, significant ENSO growth remains possible via the other. This illustrates that in each LIM it is indeed the projection of the forcing onto the two identified optimal structures (OP7 and OP15), and only these two, that completely controls ENSO development.

4.4 The optimal perturbations as ENSO precursors

The ability of OP7 and OP15 to predict ENSO events is quantified here by the fraction of ENSO variance that can be explained by the projection of each onto the the anomalous state of the climate system on a given lead time. We refer to this as hindcast skill, mindful that in some cases the OP are hindcasting the same observations used in their generation. As such we consider below whether differences in the skill of each OP in different periods may owe to artificial skill (e.g. Michaelsen 1987; Mielke et al. 1996).

As previously, ENSO amplitude is quantified by the first element of the state vector, which represents the projection onto the first EOF of SST. The ability of the optimals to reproduce observed ENSO evolution is illustrated in Fig. 13. The OP7s exhibit significant hindcast skill for their

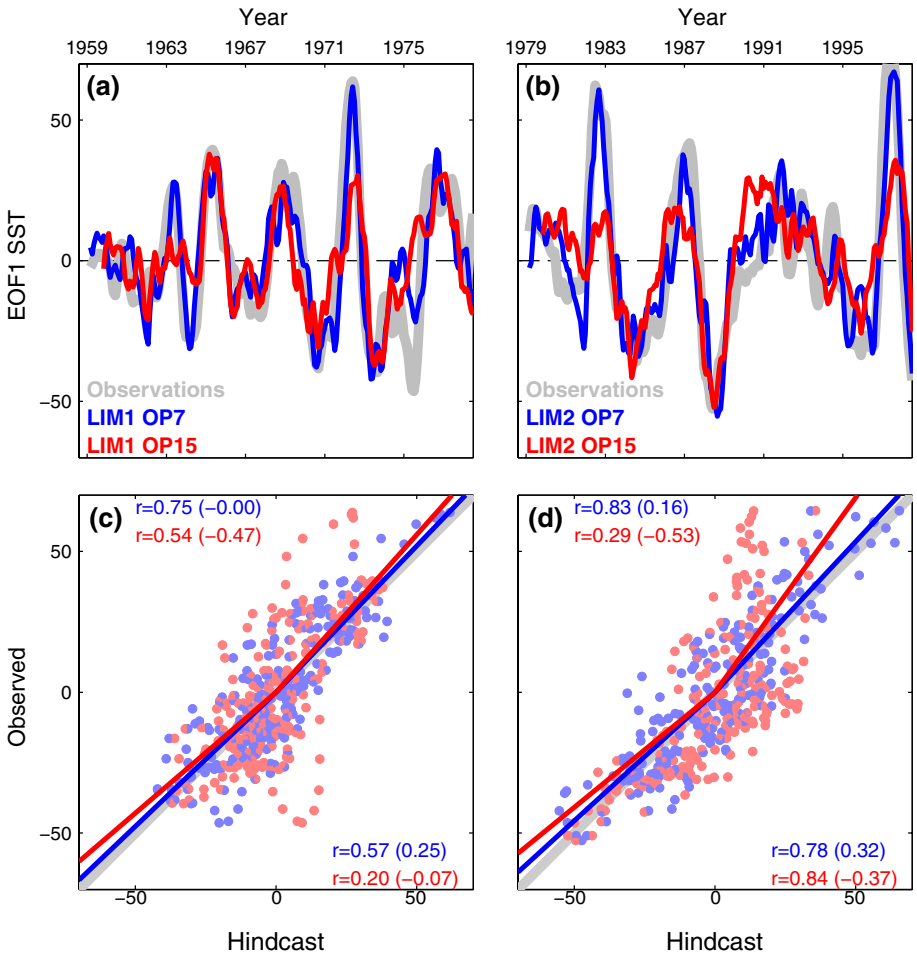

Fig. 13 Hindcasts of the principal component of the first EOF of SSTA using the optimal perturbations for each of the periods **a** 1959–1978 and **b** 1979–1998. **c**–**d** Scatter plots of observed versus hindcast ENSO amplitude for each of the two periods. The correlation coefficient for positive and negative events are shown in the *top* and *bottom*, respectively, of each panel. *Lines* of least squares best fit are drawn for positive and negative events. The value in *brackets* corresponds to the same quantity for a persistence forecast. Throughout results for the 7 month optimal (OP7) are shown in *blue*, and for the 15 month optimal (OP15) in *red*. Observed values are drawn in *grey*

corresponding time period (Fig. 14a), explaining 74 % (79 %) of ENSO variance at a lead time of 7 months for the 20 years pre- (post-) 1977 (anomaly correlation coefficients of 0.86 and 0.89, respectively). That is, in each period roughly three quarters of the variance in ENSO can be explained by a linear relationship with the projection of the observations onto the structures shown in the first row of Fig. 8. It may be appreciated in Fig. 13 that both positive and negative ENSO events are well correlated with the OP7 hindcasts. Of large ENSO events the OP7s only fail to preempt the 1975 La Niña. The OP7 display far less skill, however, when used to hindcast the alternate time period: the projection onto $OP7_1$ ($OP7_2$) explains just 25 % (23 %) of ENSO variance for the period 1979–1998 (1959–1978) (anomaly correlation coefficients of 0.50 and 0.58, respectively), only marginally better than a persistence hindcast (Fig. 14a). The fact that the optimal perturbation from each period is highly suboptimal in the other is consistent with a change in the precursors to ENSO either side of the 1970s climate shift.

Changes over time in the skill of ENSO precursors have been previously reported, and are consistent with the modulation of ENSO dynamics by an altered mean state. An alternate explanation, however, is that the differences in hindcast skill seen here between the two intervals is symptomatic of artificial skill (e.g. Mielke et al. 1996)—i.e. that each OP7 simply performs better when tested against the observations used to generate it—rather than reflecting any substantial change in the ENSO precursors. We address this by determining the variability in the skill of the precursors over an extended period, and specifically the probability that the difference in skill between the two OP7s could result from a system governed by constant ENSO dynamics. To this end, the hindcast skill of the OP7s was determined in each 20 year interval of a 20 thousand year stochastically forced simulation. The simulation was identical to those described in Sect. 4.3, with spatial structure given by the FOFs, but governed by the LIM determined from the combined period (termed LIM0 in Aiken et al. 2013). As above, and as demonstrated by Newman et al. (2011a, b) in a very similar LIM, the resulting simulation produces realistic ENSO statistics. Although substantial variability exists in the skill of each OP7, the maximum instantaneous difference between the skill of $OP7_1$ and $OP7_2$ is always less than that calculated above for hindcasts of the 20 year periods

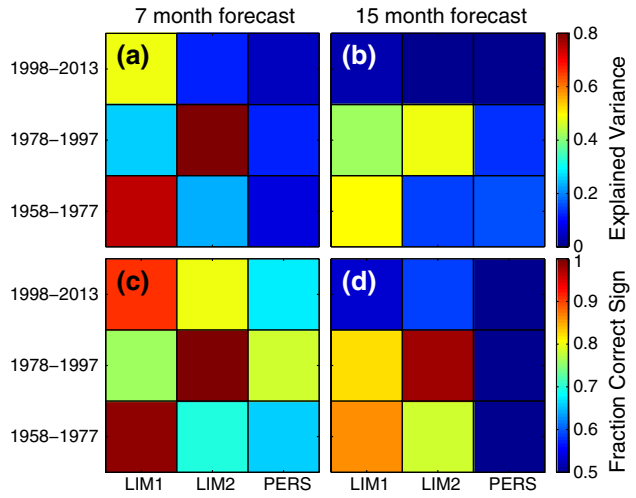

Fig. 14 **a** Hindcast skill of the optimal perturbations as determined by the fraction of variance that can be explained with a linear relationship between the projection of the observations onto the 7 months optimal perturbations and the subsequent value of an ENSO metric furnished by the projection of the observed state onto the first EOF of SST. **b** As in (**a**) for the 15 month optimal perturbations. **c** Fraction of the time that each optimal perturbation correctly hindcasts the sign of the same ENSO metric whenever its absolute value is greater than the standard deviation. A random hindcast would be correct half the time. **d** As in (**c**) for the 15 month optimal perturbations. In all cases the equivalent quantity for persistence hindcasts is shown for reference. The *colour* in each *box* indicates the skill score for the corresponding experiment

either side of the climate shift. Hence, the large differences in hindcast skill in subsequent 20 year periods cannot be explained by sampling variations from a system governed by constant dynamics. Instead we conclude that real differences exist in the dynamics and optimal precursors of ENSO straddling the climate shift.

The existence of perturbations that can grow over timescales of 15 months suggests the possibility to forecast ENSO beyond the 9 month predictability barrier (e.g. Latif et al. 1998), and we find that the OP15 can indeed often represent a skillful ENSO precursor on 15 month lead times. The 15 month hindcasts using the OP15 capture 49 and 48 % of the variance in ENSO, for 1959–1978 and 1979–1998, respectively, (anomaly correlation coefficients of 0.70 and 0.69, respectively) which is substantially better than persistence hindcasts (15 and 12 %, respectively) (Fig. 14b). Despite this skill, it is clear that OP15 fails to reproduce many of the major ENSO events in their respective study periods (Fig. 13a, b).

There is also a significant amount of asymmetry in the hindcast skill of OP15 in both periods. For instance, El Niño events are better hindcast than La Niña events in the earlier LIM1 period, while the opposite is true for the later LIM2 period. Of particular note during this later period is the significant skill of $OP15_2$ at predicting La Niña on a 15 month lead time, explaining over 70 % of the variance of negative ENSO events during that period.

These pre and post 1977 differences suggest that while predictability of ENSO at 15 month leads is possible, there is a considerable ENSO phase asymmetry in the skill and this asymmetry appears to change on multi-decadal timescales. It is interesting to note, however, that the OP15 do at least perform substantially better at predicting the sign of ENSO, correctly hindcasting the sign of ENSO events over 86 and 97 % of the time, in the periods 1959–1978 and 1979–1998, respectively (Fig. 14d).

The above demonstrates the existence of two skillful linear precursors of ENSO variability for each of the 20 year periods bracketing the climate shift. There is a clear interest in knowing which of these structures, if any, possess forecast skill since 1998. For this purpose an estimate of the ocean state from January 1998 to February 2014 was obtained from the NCEP Global Ocean Data Assimilation System (GODAS) (Behringer and Xue 2004), a data constrained solution of the GFDL MOM.v3 GCM. Given the negligible contribution of wind stress in the various optimals no wind stress product was considered. Of the four precursors discussed above, we find that only OP7 from LIM1 is able to make skillful ENSO hindcasts for the period 1998–2013, capturing approximately 55 % of the variance (Fig. 14a) and successfully reproducing the majority of the recent ENSO events (Fig. 15). Hindcasts based on $OP7_2$ and both OP15 are poor. The shorter lead times for predictability, implied by the lack of any forecast skill from either of the OP15 forecasts, is consistent with the observational study of McPhaden (2012).

5 Discussion and conclusions

We have investigated the possibility that the 1970s climate shift marked a change in the character of ENSO and in the mechanism that initiates it. Our approach centered on performing generalised stability analysis of linear inverse models (LIMs) of the coupled tropics corresponding to the periods 1959–1978 and 1979–1998, 20 years before and after the climate shift. ENSO has been previously shown to be well described as a stochastically forced damped linear system, and as a result LIMs can provide convenient empirically derived estimates of ENSO dynamics. Indeed the LIMs from each period displayed significant skill at reproducing the observed SST and covariance, implying that they capture much of the physics underlying ENSO. Eigenanalysis of the LIMs reveals that both possess two ENSO-like modes that share similar spatial structure but differing complex frequencies. These two ENSO modes resemble the 4-year and 2-year eigenmodes analysed in the LIMs of Penland and Sardeshmukh (1995), Penland

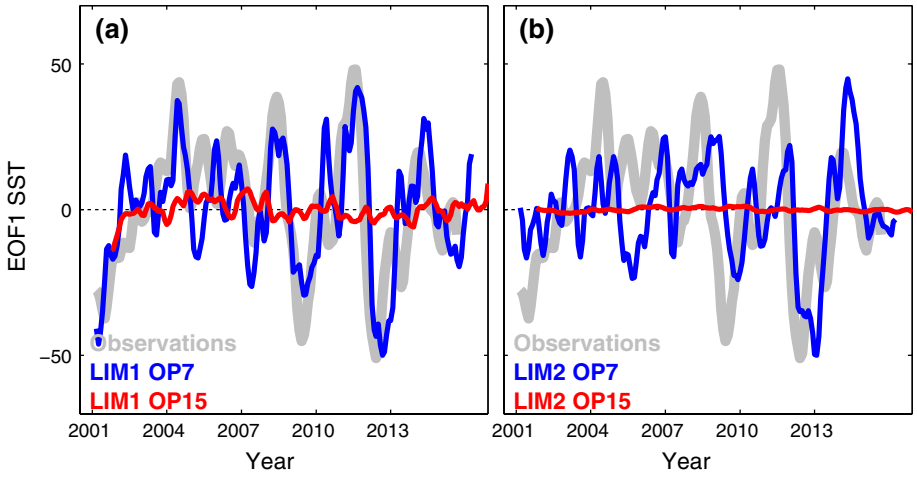

Fig. 15 As in Fig. 13a, b, but for the period since January 1998. These observations were not used in the calculation of the OP

and Matrosova (2006) and Newman et al. (2011a). In both LIMs it is the graver of the two ENSO modes, the "slow mode" with period close to 40 months and resembling the "thermocline" mode, that is the least damped mode and hence dominates the response of the forced system. The period prior to the climate shift is associated with a slow ENSO mode that is slightly less damped and of higher frequency than that corresponding to the following period. The "fast mode" in both LIMs has period close to 20 months and may be identified with the SST mode. Because these and all other modes are not mutually orthogonal (ie both LIMs are nonnormal), transient growth of perturbations is possible, and even though all perturbations must eventually decay asymptotically, the ENSO variability sustained under continued forcing greatly exceeds what would be expected from the decay rate of the ENSO modes. The linear interference between the two ENSO modes in particular underpins the transient growth and augmented variability of ENSO. Thus the existence of two distinct mechanisms for generating ENSO (ie delayed oscillator and SST advection physics) may provide a physical explanation for the nonnormality of the ENSO operator and the possibility for nonmodal ENSO growth.

In addition to enhancing the variability sustained in the system, the nonnormality of the LIMs increase their sensitivity to changes in the mean state. Calculation of the resolvent norm of each LIM for real forcing frequencies identified the slow ENSO modes as being especially sensitive to such changes. As a result, it may be anticipated that the character of ENSO is particularly sensitive to changes in the mean state of the tropical atmosphere/ocean.

For each LIM we identify two distinct optimal perturbations (OP) that grow into mature ENSO events over the course of 7 and 15 months, respectively, and between them embody the two pathways by which significant ENSO variability can be excited in the linear model. The projection of the climate state onto each of the four OP can explain a significant proportion of the observed ENSO variability for their corresponding period, confirming their relevance as ENSO precursors in the real system and suggesting that ENSO may be occasionally predictable on 15 month lead times. The difference in hindcast skill when the OP are applied to the alternate period, however, together with the significant differences in the structures of the OP, supports the hypothesis that the system's sensitivity to the spatial structure of the forcing differed following the climate shift.

This result is not surprising, given the dependence of ENSO upon the background climate state, and the body of evidence supporting altered ENSO statistics and dynamics. In fact, the possibility that the change in ENSO precursors is related to low frequency variations in the climate state, such as the IPO (Power et al. 1999; England et al. 2014), is further supported by our finding that the 7 month OP from the earlier period (1959–1978) provides skillful hindcasts of recent ENSO events (since 1998). The projection of the observed climate state onto this OP explains over half the variance in ENSO since 1998 on a 7 month lead time, and correctly forecasts the tendency towards El Niño conditions since February 2014. As such, a single linear precursor has been identified that can skillfully hindcast ENSO events during the periods of negative phase of the IPO prior to 1978 and since 1999, but not during the period with positive phase corresponding to the climate shift. Thus we postulate that the mechanisms initiating ENSO events may be modulated by the phase of the IPO. Recalling that $OP7_1$ is the only one of the four OP that does not involve a contribution from the Indian or Atlantic Oceans, it may be proposed that the ability of extra-Pacific teleconnections to initiate ENSO is most pronounced during a positive phase of the IPO. That the 1979–1998 period also saw the occurrence of the two largest recorded ENSO events may also be suggestive of a relationship between ENSO amplitude, extra-Pacific initiation and the climate state.

While it may be that the involvement of extra-Pacific teleconnections led directly to stronger ENSO, it is also possible that the stronger ENSO events enhanced remote climate variability which in turn fedback to ENSO processes within the Pacific (e.g. Santoso et al. 2012). It may also be the case that extra-Pacific SSTA has played an active role in the generation of recent ENSO events, but that the corresponding SST patterns are distinct from those that acted in the previous periods. It is possible that the increased importance of remote SST forcing following the late 1970s may be tied to changes in the mean climate. In any event, our analysis makes clear that the precursors to ENSO are not time invariant, and particularly so the contributions from remote basins.

While the global reach of ENSO has been appreciated for many decades, there has been a growing understanding of the active role played by teleconnections from outside the tropical Pacific upon the development of ENSO. Our findings are in general consistent with recent results in this regard. As inferred here, the IOD has previously been shown to anticipate a subsequent ENSO event of opposite sign, and hence extend the limit of ENSO predictability (Izumo et al. 2010; Terray 2011; Boschat et al. 2013). These studies also find that, as was found here, this relationship is most apparent from the time of the climate shift until the end of the 1990s. Unlike these previous studies, however, we find that the 15 month skill of IOD based forecasts following the climate shift is improved when eastern Pacific thermocline and SST is also included. In fact, the superposition of positive thermocline and negative SST anomalies in the eastern Pacific was a common element to the 15 month optimal precursors in both periods, and also in the adjoint ENSO modes. We also find an indication that, prior to the climate shift, Atlantic SST anomalies contributed to ENSO growth on 15 month time scales, possibly by an equivalent mechanism to Izumo et al. (2010). The possibility for south Atlantic SST anomalies to preempt ENSO events has also been previously demonstrated (Keenlyside and Latif 2007; Rodríguez-Fonseca et al. 2009; Terray 2011; Boschat et al. 2013; Keenlyside et al. 2013). On the other hand, zonal wind stress anomalies were found here to provide a negligible contribution to the development of ENSO, consistent with previous studies on the role of wind fluctuations as an ENSO precursor (Eisenman et al. 2005). It should be noted, however, that high frequency variability such as westerly wind bursts are not resolved within the limited state space considered here. Zonal current anomalies in the equatorial Pacific have also been implied to play a role in ENSO initiation (An et al. 2006), yet the addition of the principal components of the first three EOFs of anomalous zonal current to the state vector made no qualitative difference to the findings presented herein. This result may simply reflect the fact that this choice of basis mostly resolves the component of zonal current that is strongly linearly related to the thermocline and wind components, and so may miss the component involved in ENSO generation. This highlights a limitation in our method for deciding on the composition of the state vector, which is chosen to resolve the system response, but not necessarily the forcing. Ongoing work is addressing methods to optimize the choice of state space, such as through balanced truncation (Farrell and Ioannou 2001).

We have found that a significant number of ENSO events during both periods were predictable on 15 month lead times. The existence of skillful precursors for these two periods suggests that ENSO may be occasionally predictable out to 15 months, and that there may concurrently exist state-dependent "short" and "long" routes to the development of mature ENSO events. As demonstrated in the stochastically forced simulations, which of these trajectories is followed by the climate system depends upon the detail of the external forcing. While the prospect may exist for ENSO predictability over 15 month lead times, nonetheless, here we find considerable asymmetry between the forecast skill as a function of ENSO phase: during the 1959–1978 period El Niño events were hindcast with greater skill by the OP15 than La Niña events, while the reverse is true for the 1979–1998 period.

Acknowledgments This work was supported by the Australian Research Council (ARC) including the ARC Centre of Excellence in Climate System Science. We are grateful to three anonymous reviewers for their constructive comments.

References

Aiken CM, Santoso A, McGregor S, England MH (2013) The 1970s shift in ENSO dynamics: a linear inverse model perspective. Geophys Res Lett 40:1612–1617. doi:10.1002/grl.50264

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 S-I, Jin F-F (2000) An eigenanalysis of the interdecadal changes in the structure and frequency of ENSO Mode. Geophys Res Lett 27:1573–2576

An S-I, Ye Z, Hsieh WW (2006) Changes in the leading ENSO modes associated with the late 1970s climate shift: role of surface zonal current. Geophys Res Lett 33:L14609. doi:10.1029/2006GL026604

Behringer DW, Xue Y (2004) Evaluation of the global ocean data assimilation system at NCEP: the Pacific Ocean. In: Eighth symposium on integrated observing and assimilation systems for atmosphere, oceans, and land surface, AMS 84th annual meeting, Washington State Convention and Trade Center, Seattle, Washington, pp 11–15

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

Boschat G, Terray P, Masson S (2013) Extratropical forcing of ENSO. Geophys Res Lett 40:1605–1611. doi:10.1002/grl.50229

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

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

Deser C, Phillips AS, Alexander MA (2010) Twentieth century tropical sea surface temperature trends revisited. Geophys Res Lett 37:L10701. doi:10.1029/2010GL043321

Dommenget D, Semenov V, Latif M (2006) Impacts of the tropical Indian and Atlantic oceans on ENSO. Geophys Res Lett 33(L11):701. doi:10.1029/2006GL025871

Eisenman I, Yu L, Tziperman E (2005) Westerly wind bursts: ENSO's tail rather than the dog? J Clim 18:5224–5238

England MH, McGregor S, Spence P, Meehl GA, Timmermann A, Cai W, Sen Gupta A, McPhaden MJ, Purich A, Santoso A (2014) Recent intensification of wind-driven circulation in the Pacific and the ongoing warming hiatus. Nat Clim Change. doi:10.1038/nclimate2106

Farrell BF, Ioannou PJ (1996) Generalized stability theory. Part I: autonomous operators. J Atmos Sci 53:2025–2040

Farrell BF, Ioannou PJ (2001) Accurate low-dimensional approximation of the linear dynamics of fluid flow. J Atmos Sci 58:2771–2789

Fedorov AV, Philander SGH (2000) Is El Niño changing? Science 288:1997–2002

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

Frauen C, Dommenget D (2012) Influences of the tropical Indian and Atlantic Oceans on the predictability of ENSO. Geophys Res Lett 39:L02706. doi:10.1029/2011GL050520

Giese BS, Chepurin GA, Carton JA, Boyer TP, Seidel HF (2010) Impact of bathythermograph temperature bias models on an ocean reanalysis. J Clim 24:84–93. doi:10.1175/2010JCLI3534.1

Izumo T, Vialard J, Lengaigne M, de Boyer Montegut C, Behera SK, Luo JJ, Cravatte S, Masson S, Yamagata T (2010) Influence of the state of the Indian Ocean dipole on the following years El Ñino. Nat Geosci 3:168–172. doi:10.1038/ngeo760

Jansen MF, Dommenget D, Keenlyside N (2009) Tropical Atmosphere-Ocean interactions in a conceptual framework. J Climate 22(3):550–567

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

Keenlyside N, Latif M (2007) Understanding equatorial Atlantic interannual variability. J Clim 20:131–142

Keenlyside N, Ding H, Latif M (2013) Potential of equatorial Atlantic variability to enhance El Niño prediction. Geophys Res Lett 40:2278–2283

Kirtman BP, Schopf PS (1998) Decadal variability in ENSO predictability and prediction. J Clim 11:2804–2822

Kug J-S, Kang I-S (2006) Interactive feedback between ENSO and the Indian Ocean. J Clim 19:1784–1801. doi:10.1175/JCLI3660.1

Latif M, Anderson D, Barnett T, Cane M, Kleeman R, Leetmaa A, O'Brien J, Rosati A, Schneider E (1998) A review of the predictability and prediction of ENSO. J Geophys Res 103(C7):14375–14393. doi:10.1029/97JC03413

McGregor S, Timmermann A, England MH, Elison Timm O, Wittenberg AT (2013) Inferred changes in El Nio-Southern oscillation variance over the past six centuries. Clim Past Discuss 9:2929–2966

McPhaden MJ (2012) A 21st century shift in the relationship between ENSO SST and warm water volume anomalies. Geophys Res Lett 39:L09706. doi:10.1029/2012GL051826

McPhaden MJ, Zhang X (2009) Asymmetry in zonal phase propagation of ENSO sea surface temperature anomalies. Geophys Res Lett 36:L13703. doi:10.1029/2009GL038774

McPhaden MJ, Zebiak SE, Glantz MH (2006) ENSO as an integrating concept in Earth science. Science 314:1740–1745

McPhaden MJ, Lee T, McClurg D (2011) El Niño and its relationship to changing background conditions in the tropical Pacific Ocean. Geophys Res Lett 38:L15709. doi:10.1029/2011GL048275 (2011)

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

Michaelsen J (1987) Cross-validation in statistical climate forecast models. J Clim Appl Meteorol 26:1589–1600

Mielke PW Jr, Berry KJ, Landsea CW (1996) Artificial skill and validation in meteorological forecasting. Weather Forecast 11:153–169

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

Neelin JD, Battisti DS, Hirst A, Jin F-F, Wakata Y, Yamagata T, Zebiak SE (1998) ENSO theory. J Geophys Res 103:14261–14290

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

Newman M, Alexander MA, Scott JD (2011a) An empirical model of tropical ocean dynamics. Clim Dyn. doi:10.1007/s00382-011-1034-0

Newman M, Shin S-I, Alexander MA (2011b) Natural variation in ENSO flavors. Geophys Res Lett 38:L14705. doi:10.1029/2011GL047658

Penland C, Magorian T (1993) Prediction of Niño-3 sea surface temperatures using linear inverse modelling. J Clim 6:1067–1076

Penland C, Matrosova L (1994) A balance condition for stochastic numerical models with application to the El Nino-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

Power S, Casey T, Folland C, Colman A, Mehta V (1999) Interdecadal modulation of the impact of ENSO on Australia. Clim Dyn 15:319–324

Rodríguez-Fonseca B, Polo I, Garcia-Serrano J, Losada T, Mohino E, Mechoso CR, Kucharski F (2009) Are Atlantic Niños enhancing Pacific ENSO events in recent decades? Geophys Res Lett 36:L20705. doi:10.1029/2009GL040048

Santoso A, England MH, Cai W (2012) Impact of Indo-Pacific feedback interactions on ENSO dynamics diagnosed using ensemble climate simulations. J Clim 25:7743–7763

Santoso A, McGregor S, Jin F-F, Cai W, England MH, An S-I, McPhaden MJ, Guilyardi E (2013) Late-twentieth-century emergence of the El Niño propagation asymmetry and future projections. Nature 504:126–130

Terray P (2011) Southern hemisphere extra-tropical forcing: a new paradigm for El Niño-Southern Oscillation. Clim Dyn 36:2171–2199. doi:10.1007/s00382-010-0825-z

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

Tokinaga H, Xie1 S-P, Timmermann A, McGregor S, Ogata T, Kubota H, Okumura YM (2012) Regional patterns of tropical Indo-Pacific climate change: evidence of the Walker circulation weakening. J Clim 25:1689–1710

Trefethen LN (1991) Pseudospectra of matrices. In: Griffiths DF, Watson GA (eds) Numerical analysis. Longman, pp 234–266

Trefethen LN, Trefethen AE, Reddy SC, Driscoll TA (1993) Hydrodynamic stability without eigenvalues. Science 261:578–584

Trenberth KE, Stepaniak DP, Caron JM (2002) Interannual variations in the atmospheric heat budget. J Geophys Res 107(D8):4066. doi:10.1029/2000JD000297

Uppala SM, Kallberg PW, Simmons AJ, Andrae U, Da Costa Bechtold V, Fiorino M, Gibson JK, Haseler J, Hernandez A, Kelly GA, Li X, Onogi K, Saarinen S, Sokka N, Allan RP, Andersson E, Arpe K, Balmaseda MA, Beljaars ACM, Van de Berg L, Bidlot J, Bormann N, Caires S, Chevallier F, Dethof A, Dragosavac M, Fisher M, Fuentes M, Hagemann S, Hólm E, Hoskins BJ, Isaksen L, Janssen PAEM, Jenne R, McNally AP, Mahfouf J-F, Morcrette J-J, Rayner NA, Saunders RW, Simon P, Sterl A, Trenberth KE, Untch A, Vasiljevic D, Viterbo P, Woollen J (2005) Quart J R Meteorol Soc 131:2961–3012

Vecchi GA, Soden BJ, Wittenberg AT, Held IM, Leetmaa A, Harrison MJ (2006) Weakening of tropical Pacific atmospheric circulation due to anthropogenic forcing. Nature 441:73–76

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

Wang B, An S-I (2002) A mechanism for decadal changes of ENSO behavior: roles of background wind changes. Clim Dyn 18:475–486

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

Woodruff SD, Worley SJ, Lubker SJ, Ji Z, Eric Freeman J, Berry DI, Brohan P, Kent EC, Reynolds RW, Smith SR, Wilkinson C (2011) ICOADS Release 2.5: extensions and enhancements to the surface marine meteorological archive. Int J Climatol 31:951–967. doi:10.1002/joc.2103

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

Zhang H, Clement A, Di Nezio P (2014) The South Pacific Meridional mode: a mechanism for ENSO-like variability. J Clim 27:769–783