Forecasting Pacific SSTs: Linear Inverse Model Predictions of the PDO

MICHAEL A. ALEXANDER
NOAA/Earth System Research Laboratory, Boulder, Colorado

LUDEMILA MATROSOVA
NOAA/Earth System Research Laboratory, and CIRES Climate Diagnostics Center, Boulder, Colorado

CÉCILE PENLAND
NOAA/Earth System Research Laboratory, Boulder, Colorado

JAMES D. SCOTT
NOAA/Earth System Research Laboratory, and CIRES Climate Diagnostics Center, Boulder, Colorado

PING CHANG
Department of Oceanography, Texas A&M University, College Station, Texas

(Manuscript received 11 January 2007, in final form 14 May 2007)

ABSTRACT

A linear inverse model (LIM) is used to predict Pacific (30°S–60°N) sea surface temperature anomalies (SSTAs), including the Pacific decadal oscillation (PDO). The LIM is derived from the observed simultaneous and lagged covariance statistics of 3-month running mean Pacific SST for the years 1951–2000. The model forecasts exhibit significant skill over much of the Pacific for two to three seasons in advance and up to a year in some locations, particularly for forecasts initialized in winter. The predicted and observed PDO are significantly correlated at leads of up to four seasons, for example, the correlation exceeds 0.6 for 12-month forecasts initialized in January–March (JFM). The LIM-based PDO forecasts are more skillful than persistence or a first-order autoregressive model, and have comparable skill to LIM forecasts of El Niño SSTAs. Filtering the data indicates that much of the PDO forecast skill is due to ENSO teleconnections and the global trend.

Within LIM, SST anomalies can grow due to constructive interference of the empirically determined modes, even though the individual modes are damped over time. For the Pacific domain, the basinwide SST variance can grow for \( \approx 14 \) months, consistent with the skill of the actual predictions. The optimal structure (OS), the initial SST pattern that LIM indicates should increase the most rapidly with time, is clearly relevant to the predictions, as the OS develops into a mature ENSO and PDO event 6–10 months later. The OS is also consistent with the seasonal footprinting mechanism (SFM) and the meridional mode (MM); the SFM and MM involve a set of atmosphere–ocean interactions that have been hypothesized to initiate ENSO events.

1. Introduction

Accurate predictions of upper-ocean conditions, including sea surface temperature anomalies (SSTAs), can provide many economic and societal benefits. For example, forecasts based on the state of the climate system, especially the state of ENSO and Pacific SSTs in general have the potential to provide useful information to fishermen, farmers, water resource managers, and industry (Dettinger et al. 1999; Stern and Esterling 1999; Hill et al. 2000; Hamlet et al. 2002; Huppert et al. 2002).

Studies examining the predictability of North Pacific
A wide array of both physical and statistical methods has been used to predict ENSO variability (see reviews by Latif et al. 1998; Barnston et al. 1999; Mason and Mimmak 2002). Forecasts from coupled ocean–atmosphere models of varying complexity include “intermediate” coupled models, in which anomalous air–sea interaction is simulated in simplified physical atmosphere–ocean models within the tropical Pacific domain (e.g., Cane et al. 1986; Zebiak and Cane 1987), “hybrid” coupled models that consist of a physical ocean model coupled to a statistical model of the atmosphere (Barnett et al. 1993; Balmaceda et al. 1994), and coupled general circulation models (GCMS; e.g., Kirtman et al. 1997; Stockdale et al. 1998; Saha et al. 2006). Statistical ENSO forecast techniques include linear methods such as multiple linear regression [e.g., El Niño–Southern Oscillation Climatology and Persistence (ENSO-CLIPER) Landsea and Knaff 2000], canonical correlation analyses (CCAs; Barnett et al. 1988; Barnston and Ropelewski 1992), and linear inverse models (LIMs; Penland and Sardeshmukh 1995b, hereafter PS95), where the latter uses concurrent and lagged covariance statistics to predict the evolution of the system and diagnose its dynamics. Nonlinear statistical methods, such as neural networks (Tangang et al. 1998), nonlinear principal component analysis (Mann 2001), and quadratic inverse models (Kondrashov et al. 2005), have also been used to forecast ENSO.

In contrast to the tropical Pacific, there have been far fewer attempts to forecast the North Pacific Ocean and often these are part of a broader effort to predict global ocean conditions. For example, the National Centers for Environmental Prediction (NCEP) has recently instituted a Climate Forecast System (CFS; Saha et al. 2006), which predicts global SSTs using a coupled GCM. A two-tier approach has also been developed to make 3-month forecasts of North Pacific upper-ocean conditions (Auad et al. 2004), in which an AGCM is driven by persistent initial SST anomalies (i.e., the SSTs are held fixed over the duration of the atmospheric model run) and then the anomalous flux fields from the AGCM are used to drive an OGCM. While this method has shown some success in some regions, the forecasts are of limited duration and skill due to the chaotic nature of the extratropical atmosphere and biases in the simulation of the extratropical ocean in the OGCM. Previous statistical methods used to predict global SSTs, have also had limited success in forecasting extratropical SSTs on seasonal and longer time scales (Landman and Mason 2001). Given that there are well-understood oceanic processes that evolve over seasonal and longer time scales, a targeted effort should be able to improve forecasts of the North Pacific Ocean.

Pronounced decadal fluctuations occurred over the North Pacific during the twentieth century, which Mantua et al. (1997) termed the Pacific decadal oscillation (PDO) based on transitions between relatively stable states of the dominant pattern of North Pacific SST anomalies. The PDO consists of anomalies of one sign in the central-west Pacific between approximately 35° and 45°N, ringed by anomalies of the opposite sign, the most prominent pattern on interannual as well as decadal time scales (Zhang et al. 1997; Vimont 2005). The decadal SST transitions were accompanied by widespread changes in the atmosphere, ocean, and marine ecosystems (e.g., Miller et al. 1994; Trenberth and Hurrell 1994; Benson and Trites 2002). The associated SST anomalies extend over the entire basin and are symmetric about the equator (Zhang et al. 1997; Garreaud and Battisti 1999; Enfield and Mestás-Nunez 1999), leading some to term the phenomenon the interdecadal pacific oscillation (IPO; Power et al. 1999; Folland et al. 2002), suggesting that forecasts of the PDO (or IPO) should include information from the Pacific as a whole.

The mechanisms proposed to explain Pacific decadal
climate variability, including the PDO, can be classified as extratropical or tropically forced. Extratropical mechanisms for the PDO include intrinsic atmospheric variability that force the ocean through surface heat fluxes (e.g., Frankignoul and Hasselmann 1977), large-scale patterns of heat fluxes plus advection by the mean ocean gyres (Saravanan and McWilliams 1998; Wang and Chang 2004), and by wind-driven adjustments in the ocean circulation via Rossby waves (Frankignoul et al. 1997; Jin 1997; Neelin and Weng 1999). If the SSTs induced by the altered circulation exert a strong positive feedback on the atmospheric, air-sea interaction combined with fluctuations in the extratropical ocean gyres may result in self-sustaining decadal oscillations (Latif and Barnett 1994, 1996; Robertson 1996; Kwon and Deser 2007). Atmospheric teleconnections from the equatorial Pacific can impact North Pacific SSTs primarily through variability in the surface fluxes and Ekman transport (Trenberth 1990; Graham et al. 1994; Deser et al. 2004; Alexander et al. 2002; Seager et al. 2004).

Several recent studies have used statistical analyses to reconstruct the annually averaged PDO and determine the processes that underlie its dynamics. Newman et al. (2003b) found that for annual average anomalies (July–June) the PDO is well modeled as the sum of atmospheric forcing represented by white noise, forcing due to ENSO, and memory of SST anomalies in the previous year. The latter results in part from the “re-emergence mechanism,” where ocean temperature anomalies created in winter are sequestered below the mixed layer in summer and then reentrained into the mixed layer in the following fall and winter (Alexander and Deser 1995; Alexander et al. 1999; Deser et al. 2003). Schneider and Cornuelle (2005) found that the annually averaged PDO could be reconstructed based on a first-order autoregressive model and forcing associated with variability in the Aleutian low (essentially internal atmospheric noise), ENSO, and wind-driven Rossby waves in the North Pacific Ocean. On interannual time scales, ENSO and the Aleutian low were about equally important at determining the PDO while on decadal time scales, all three processes, including ocean dynamics, were of equal importance. Newman (2007) expanded upon the heuristic PDO model (Newman et al. 2003b) by constructing a LIM from tropical and extratropical annual SST anomalies that indicated that three patterns explained most of the variability of the PDO: ENSO-like patterns with periods of 5 and 25 yr, and a pattern that reflects very low-frequency variability that also resembles the trend over the last 50 yr. While Schneider and Cornuelle (2005) and Newman (2007) found modes that had decadal periods, all of the modes were damped on subdecadal time scales, implying that statistical predictions of the PDO are possible on seasonal-to-interannual but not decadal time scales.

We are interested in constructing an empirical model that can predict Pacific SSTs on seasonal time scales and be used to diagnose how different ocean regions impact these predictions. LIMs have been used to study and predict many aspects of the climate system (e.g., Farrell and Ioannou 1995; Zhang and Held 1999; Winkler et al. 2001). Furthermore, LIMs have been used extensively to make forecasts of SST anomalies in the tropical Pacific associated with ENSO (Penland and Magorian 1993; PS95), and in the tropical Indian and Atlantic Oceans as well (Penland and Matrosova 1998, hereafter PM98) that are skillful at lead times of up to ~15 months. These forecasts (available online at http://www.cdc.noaa.gov/forecast1/Frcst.html) are competitive with dynamical forecasts methods, especially for forecast leads of 6 months and greater. In the present study, we will expand the LIM-based predictions to include North Pacific SSTs, with a focus on forecasting the PDO. We will also use LIM in conjunction with composite analyses to understand the physical processes that influence the development of ENSO events and the evolution of the PDO.

2. Methods and data

a. LIM

Inverse modeling can be defined as the extraction of dynamical properties of a system from its observed statistics. It can be contrasted with forward modeling in which one defines the system dynamics from the equations for heat, mass, and momentum with specified parameters. LIM assumes that a system can be separated into a linear deterministic portion and a nonlinear portion that can be represented by white noise–random fluctuations that may be correlated in space but not over time. The deterministic portion, the part that is used in making forecasts, is estimated using eigenfunction analysis of matrices whose elements are estimated from multiple linear regression (e.g., von Storch et al. 1988) and is related to other statistical methods. For example, LIM-generated forecasts are identical to those from CCA when all canonical patterns are superposed (DelSole and Chang 2003). By including a stochastic component, LIM also allows one to test several aspects of the forecasts, including its error characteristics. Here we provide only a brief description of LIM, as the underlying assumptions, the procedures used to derive it, and its strengths and weaknesses have been described in a wide array of studies (e.g., Penland
Linear inverse modeling assumes that the relevant dynamics can be written in the form of a linear stochastic differential equation:

$$\frac{dx}{dt} = Bx + \xi,$$

where $x$ is the state of the system, $B$ is the evolution operator, and $\xi$ is the stationary white noise. For LIM to be an appropriate forecast method with stationary statistics, $B$ must be dissipative (i.e., its eigenvalues must have negative real parts). In LIM, one assumes that Eq. (1) is valid, resulting in the relation

$$C(\tau) = \mathbf{G}(\tau)C(0),$$

where

$$\mathbf{G}(\tau) = \exp(B\tau),$$

and $C(0)$ and $C(\tau)$ are the covariance matrices of vector $x$ at lags 0 and $\tau$. In a forecasting context, $\mathbf{G}(\tau)x(t)$ represents the “best” forecast (in a least squares sense) of $x(t + \tau)$ given $x(t)$. The matrices $B$ and $\mathbf{G}$ can then be determined from observational estimates of $C(0)$, and $C(\tau_0)$ at some lag $\tau_0$:

$$B = \tau_0^{-1} \ln[C(\tau_0)C(0)^{-1}].$$

Unlike multiple linear regression, in the LIM framework $B$ and $\mathbf{G}$ are independent of the choice of $\tau_0$. How well this holds in practice provides one measure of the efficacy of using LIM for a given system (PS95; Penland 1996).

Once $B$ has been computed, the statistics of the noise forcing ($\xi$) can be determined from the fluctuation–dissipation relationship (FDR),

$$BC(0) + C(0)B^T + \langle \xi\xi^T \rangle dt = 0,$$

essentially a budget of the second-order moments of the system. The angle brackets in Eq. (5a) denote an ensemble mean estimated here as a time average. In this study, we use a slight generalization of Eqs. (1)–(5a): the statistics of $\xi$ are allowed to be periodic. The LIM formalism still holds provided that all averages in Eqs. (2)–(4) are interpreted as integrals over the annual cycle. The time dependence of the noise statistics are estimated using the annually varying equal-time covariance matrix of $x$ in the time-dependent FDR (e.g., Penland and Matrosova 1994):

$$B(xx^T) + (xx^T)B^T + \langle \xi\xi^T \rangle dt = \frac{d(xx^T)}{dt}.$$  

Here we have defined the state vector, $x$, only in terms of SST anomalies. The linear influence of other variables on SST evolution, such as surface fluxes and oceanic advection, however, are implicitly included in $B$.

b. Data

To maintain the continuity with previous LIM-based SST prediction studies (e.g., PS95; PM98), we use the collection of historical records from ship observations from the Comprehensive Ocean–Atmosphere Data Set (COADS; Woodruff et al. 1987; Worley et al. 2005). The initial data consist of monthly mean SSTs for the period of January 1951–December 2000 over the Pacific Ocean between 30°S and 60°N. The data, which were originally on 2° latitude × 2° longitude grid, were consolidated onto a 4° × 10° grid, with the requirement that at least four of the original 10 grid boxes have observations in any given month. SST data were smoothed using a 3-month running mean. Anomalies were calculated relative to the 1951–2000 climatology in each 4° × 10° grid square. The 3-month coarse graining does not allow some dynamical processes, such as tropical instability waves, to be explicitly resolved, but rather allows their effect to be approximated as stochastic forcing.

As in many previous studies, LIM is constructed using a reduced space of empirical orthogonal functions (EOFs) where the state vector ($x$), consists of the amplitude of the corresponding principal components (PCs). No artificial noise was used to regularize the covariance matrix used to estimate the EOFs, but the EOF projection is itself a mild regularization of the system since dimensions corresponding to very small negative eigenvalues of the covariance matrix are eliminated by truncating the EOF expansion. Truncating the number of EOFs also reduces the amount of processing and eliminates some of the errors due to missing or imperfect data. Here, we have retained the 15 leading EOFs, which explain 65% of the total variance. For individual grid squares these EOFs explain >80% of the variance in portions of the central-eastern tropical Pacific and in the vicinity of 45°N, 150°W, >65% over most of the ENSO region and the northeast Pacific, >50% outside the tropical convergence zones, and >35% over nearly the entire domain (see online at http://www.cdc.noaa.gov/forecast1/descrpdo.html).

Following Penland (1996) and PM98, estimates of $\mathbf{G}(\tau_0)$, with $\tau_0 = 3$ months, were obtained using a jackknife approach, where $\mathbf{G}(\tau_0)$ is calculated from 45 yr of data in the 1951–2000 period and then used to forecast SSTs in the excluded 5-yr periods. We use six nonoverlapping verification periods during 1971–2000, when the data are most reliable. For example, $\mathbf{G}(\tau_0)$ was estimated using data from January–March (JFM) 1951–October–December (OND) 1971 and November–Janu-
ary (NDJ) 1976–OND 2000 and verified in NDJ 1971–OND 1976. The EOF basis computed from all 50 yr of data was used in the cross-validation procedure, which could result in shared information between prediction and verification periods. In practice, Penland and Sardeshmukh (1995a) found that the spatial patterns of the individual SSTA EOFs differed only slightly when verification periods comparable to those used here were excluded, while in a similar context Newman et al. (2003a) found that using an EOF basis that excluded the verification period had a very small impact on forecast skill. The forecast skill was assessed using the average of the statistics estimated during each verification period. Application of the “tau test” (Penland 1989; PS95) indicated that the forecast skill was robust for over one period, or a single stationary mode ($\tau_{\text{tau}}$).

c. Empirical normal modes

The eigenvector solutions to the deterministic portion of Eq. (1), that is,

$$dx/dt = Bx,$$  \hspace{1cm} (6)

were termed principal oscillation patterns (POPs) by Hasselmann (1988) and von Storch et al. (1988), who were concerned with finding a single pattern or pair of patterns that dominated the field of interest. Penland and Ghil (1993) used the term empirical normal modes (ENMs) to describe these eigenvectors to emphasize their collective behavior and to conform to terminology used in other scientific disciplines. ENMs can be expressed as

$$x_j(t) = (a_j \cos \omega_j t + b_j \sin \omega_j t) \exp(\sigma_j t),$$  \hspace{1cm} (7)

where $\omega_j/2\pi$ and $-1/2\sigma_j$ are the period and decay time, respectively, of the $j$th ENM; $a_j$ and $b_j$ are real vectors that satisfy the conditions $a_j \cdot b_j = 0$, $a_j \cdot a_j = 1$, and $b_j \cdot b_j \geq 1$ (with no implied summation over $j$); and $j$ increases as the decay time decreases (PS95). Since the data are described in terms of an EOF expansion truncated at 15 members, no more than 15 ENMs can be estimated. The ENMs take the form of either oscillatory paired modes that evolve from $a \rightarrow b \rightarrow a \rightarrow b \rightarrow a$ over one period, or a single stationary mode ($\omega_j = b_j = 0$).

Here we only briefly discuss the ENMs for the SST anomalies as they have been studied extensively for the tropics (30°N–30°S) by Penland and Magorian (1993), PS95, PM98, and Penland and Matrosova (2006). In general agreement with these studies, the evolution of ENSO-related SST anomalies in the tropical Pacific are dominated by a few (five in this analysis: $j = 3–6, 8$) normal modes. In addition, there are two stationary modes ($j = 1, 2$) that represent a global very low-frequency signal (nearly a trend) with a strong projection on both the tropical and North Pacific Ocean. The evolution of the sum of these modes is nearly identical to a single mode isolated in Penland and Matrosova (2006).

d. The PDO

Following Mantua et al. (1997) we first calculate the PDO from the leading EOF of North Pacific (20°–60°N) SSTAs, with the difference that we use 3-month running means on a 4° × 10° grid. The EOF pattern is very similar to that of Mantua et al., with anomalies of one sign in the central and west Pacific between approximately 35° and 45°N ringed by anomalies of the opposite sign (Fig. 1a). The corresponding time series, or principal component, is termed the “original PDO.” Since we are interested in using our inverse model to predict the PDO, we compute a “reconstructed PDO” time series by first reconstructing the SSTAs in geographical space using the 15 basinwide EOFs and then projecting these SSTAs in the North Pacific onto the original PDO EOF. The reconstructed PDO has a very close correspondence with the original [i.e., the correlation ($r$) between the two time series is 0.99 (Fig. 1b)]. Given the two are nearly identical, we will refer to the reconstructed values as the PDO from hereon.

The time series of SSTA in the Niño-3.4 region (5°N–5°S, 170°–120°W), also shown in Fig. 1a, has a simultaneous correlation of 0.5 with the PDO (0.53 in the EOF-space reconstruction and 0.48 in the untruncated SST time series, all significant at the 95% level taking into account serial correlation), which suggests a moderately strong relationship between the PDO and SSTs in the ENSO region. This correlation is higher than the 0.38 value reported by Mantua et al. (1997) due to the reduction in noise by using 3-month running means rather than monthly values; differences in the datasets, periods of record, and ENSO indices may also affect the ENSO–PDO $r$ values obtained by the two studies.

3. LIM results

a. SST predictions

As a first step in predicting North Pacific SSTs, we evaluate forecasts of the PDO, given its relevance to the climate and ecosystem research communities. We assess the error in the LIM-based PDO forecasts by comparing them to other error estimates and prediction methods, including a theoretical error estimate, persistence, and an autoregressive model. Under LIM any
two states separated by a time interval \( \tau \) are related by

\[
x(t + \tau) = G(\tau)x(t) + \varepsilon,
\]

where \( \varepsilon \) is the error vector. The theoretically expected global error covariance, representing the effect of the unpredictable stochastic forcing but not uncertainties in estimating \( G(\tau) \) or inaccuracies in the initial conditions, is given by

\[
E(t, \tau) = \left( x(t + \tau)x^T(t + \tau) \right) - G(\tau)(x(t)x^T(t))G^T(\tau),
\]

where \( E(t, \tau) \) is solely a function of lead time \( \tau \) if the white noise forcing is stationary. Equation (8) provides a measure of the expected forecast error due to unpredictable dynamics. Thus, the similarity between the theoretical and the actual forecast errors provides a measure of the underlying dynamical assumptions used in the prediction model (i.e., that SST anomalies in the Pacific can be well described as a stable linear process driven by additive white noise; Penland and Matrosova 2001; Newman et al. 2003a).

The mean-square error variance of PDO forecasts from (i) LIM, (ii) persistence, and (iii) a univariate first-order autoregressive (AR1) method with a decay time of \( \sim 22 \) months, computed from \( 1/\ln \) of the 1-month lag autoregression value, as well as (iv) the theoretical error estimate \( [E(t, \tau)] \), all averaged over the annual cycle and normalized by the variance of the observed PDO time series, are shown in Fig. 2a. (Results from an AR2 model of the PDO are not shown as they had nearly identical forecast skill as the AR1 scheme.) Of these methods, LIM provides the best forecasts, especially at lead times of 5–15 months. AR1 models, like those used by Frankignoul and Hasselmann (1977) to predict mid-latitude SST anomalies at a given location, assume the dynamics to be local red noise (with a lag-1 autocorrelation between 0.0 and 1.0) and therefore cannot predict anomaly growth. Figure 2a indicates that even though the PDO has low-frequency variability its evolution is better predicted by LIM, where anomalies can grow via superposition of damped modes, then by an AR1 model (as discussed further in the following section). The actual forecast errors are also close to the theoretical uncertainty estimates, suggesting that the underlying dynamics in the LIM forecasts are reasonable.

Is the model skill in predicting the PDO due to information residing in the extratropical SSTs or does it arise from ENSO teleconnections and/or other signals originating from outside the North Pacific? While a LIM based solely on North Pacific SSTA could be developed to examine this question, the predictions would still be influenced by ENSO via its indirect influence on B. Instead, following Penland and Matrosova (2006), we construct a filter based on the empirical normal modes from the full Pacific domain to extract the “ENSO signal” and a “global trend” (see section 2c). These signals are identified by the strong projection of the relevant modal adjoints onto an “optimal initial structure” (discussed in section 3b), or alternatively the combination of normal modes whose constructive interference maximizes the field variance—the norm used here. The field obtained by combining the ENMs whose adjoints do not project strongly onto the OS are termed “the residual.” The normalized prediction errors and the correlation between the predicted and actual PDO time series from the complete LIM, the residual modes, and the residual plus the trend modes are shown in Figs. 2b,c, respectively. The filtered forecasts suggest that ENSO and the global trend provide much of the predictability for the PDO. For example, the correlation between the observed and the PDO forecasts from the full LIM, the residual + trend and the residual drops below 0.4 at leads of approximately 15, 10, and 4 months, respectively. Thus, we focus on the predictions from the complete LIM from hereon.

Predictions of the PDO time series at four lead times: 3, 6, 9, and 12 months (thick lines) and the corresponding observed PDO time series (thin lines) are presented in Fig. 3 for the period NDJ 1971–OND 2000. The forecasts are generated for 5-yr nonoverlapping data seg-
ments, which results in gaps of 3, 6, 9, and 12 months in the corresponding time series. The correlation between the predicted and observed PDO for all seasons during 1971–2001 is 0.81, 0.67, 0.55, and 0.44 at leads of 3, 6, 9, and 12 months, respectively. These values are significant at the 95% level and are comparable to LIM-based forecasts of SSTs in the Niño-3 region (e.g., see PS95).

The LIM forecast skill is explored further in Fig. 4a, which shows the correlation between the predicted and observed PDO time series as a function of the initial 3-month season and the length of the forecast. The correlation values decrease nearly uniformly with respect to the initial forecast time for leads of up to 6 months, but still exceed 0.6 regardless of which season they were initiated. For longer forecasts, those commenced during the cold season, particularly in late winter, tend be more skillful than those that begin in summer. For example, the correlations between the observed and predicted PDO are approximately 0.6, 0.2, and 0.4, for 12-month forecasts initiated in JFM, JAS, and NDJ, respectively. The correlation between the observed and LIM-based forecast of SST anomalies in Niño-3.4 are shown in Fig. 4b. The ENSO correlation skill tends toward a local maximum for forecasts that verify in DJF, which can result in forecasts at longer leads having greater skill. For example, the skill of predictions initialized in MJJ increases slightly for leads from 5 to 8 months, corresponding to forecasts that verify in SON and DJF, respectively. Overall, the PDO and ENSO forecasts are of comparable skill, although the latter generally exhibit higher correlations for \( \tau < 6 \) months, particularly for forecasts that begin in spring and summer.

The LIM-based forecasts provide useful information for SST evolution over much of the Pacific, as indicated by the geographical distribution of correlations between the model forecasts at lead times of 3, 6, 9, and 12 months and observations in each \( 4^\circ \times 10^\circ \) grid square (Fig. 5). The correlations exceed 0.6 for almost the entire domain for 3-month forecasts and 0.4 over much of the eastern tropical Pacific and western tropical Pacific north of the equator for forecasts out to 12 months. In the North Pacific, regions of higher skill occur in the the southern Bering Sea, the central basin (near \( 30^\circ N, 170^\circ W \)), and along the coast of North America. SST anomalies in these regions are influenced by the atmospheric bridge via teleconnections associated with ENSO (e.g., Alexander et al. 2002) and for nearshore regions by coastally trapped waves (e.g., Enfield and Allen 1980; Ryan and Noble 2002; Strub and James 2002).

Given that the PDO forecasts issued in winter appear to be the most skillful beyond \( \approx 6 \) months (Fig. 4), the

---

**Fig. 2.** (a) Mean-square forecast error, normalized by mean-square amplitude of the verification based on predictions from LIM, a univariate AR1 process, persistence, and a theoretical estimate of the error (see text for details). (b) Same as in (a), but for filtered forecasts from the residual (modes 7, 9–15), the residual plus the trend (modes 1, 2, 7, and 9–15), and the complete LIM [repeated from (a)]. (c) The correlation between the observed and predicted PDO time series from the residual, trend + residual, and complete LIM.
correlation between the observed SSTs and the corresponding 12-month forecast initiated in JFM over the Pacific domain are shown in Fig. 6. The model forecasts exhibit significant skill \( r > 0.4 \) in (i) the central subtropics of the Southern Hemisphere, (ii) the central equatorial Pacific with an extension toward central America, (iii) the vicinity of the Maritime Continent, (iv) in the central North Pacific, and (v) the subarctic Pacific extending from Canadian coast to the Kamchatka Peninsula. The latter two areas closely coincide with regions of strong loadings for the leading EOF of North Pacific SST anomalies (i.e., the PDO pattern; cf. Fig. 1a).

### b. Model dynamics: Anomaly growth and the optimal structure

Within the LIM framework, variability is modeled as a multivariate process, whose properties differ from univariate systems in two key ways: (i) the spatial structure of the patterns can evolve over time and (ii) even though the amplitude of the patterns are damped, constructive interference between patterns is still possible enabling transient energy growth (e.g., Farrell 1988). Here we define energy growth as the increase in SSTA variance integrated over the whole domain, where the energy changes by a factor corresponding to an eigen-
The fastest growth in the absence of forcing occurs when the initial condition projects strongly onto the leading eigenvector \( \phi_1 \) of matrix \( G^T G \) whose basis is orthonormal (in contrast to \( B \)); the growth rate of \( \phi_1 \) is given by the leading eigenvalue \( \gamma_1 \). The \( \gamma_1 \) values as a function of the forecast time \( \tau \), referred to as the “maximum amplification curve” (Borges and Sardeshmukh 1995; PS95), are shown in Fig. 7. The maximum growth rate of approximately 2.2 occurs at \( \tau \) of 6–7 months, and growth \( (\gamma_1 > 1) \) is limited to \( \sim 14 \) months. The general form of the curve is similar to those based on SST anomalies in the tropical Indo-Pacific (PS95) and the global tropics (PM98); although here the amplification is weaker and the period over which growth occurs is shorter than in PS95 and PM98.

The SSTA pattern that undergoes the maximum growth, the optimal (initial) structure, occurs when \( \mathbf{x}(0) \) strongly projects on \( \phi_1(\tau) \). When the optimal structure (OS) for \( \phi_1(7) \) (Fig. 8a) is specified as the initial condition in Eq. (6), the SSTA field evolves into the pattern shown in Fig. 8b 7 months later. The latter closely resembles the mature ENSO state, including the bridge-related SST anomalies in the North Pacific. The extratropical portion of Fig. 8b also strongly resembles the PDO pattern (see Fig. 1a). The close correspondence between the North Pacific SST forecasts initiated with the OS and the observed PDO is demonstrated in Fig. 9, which shows the pattern correlation between the two as a function of forecast time. The correlation is \( \sim 0.2 \) at \( \tau = 0 \), but increases to more than 0.9 for \( \tau \) between 6 and 9 months. In addition, if we construct a monthly time series of the pattern correlation between the OS and SSTA at zero lag (a measure of the projection of the OS onto the data) the temporal correlation between this time series and the observed PDO time series 7 months later is 0.6 (significant at the 95% level). Although there is clearly skill in predicting the PDO based on the optimal structure, the amplitude of the anomalies centered near 35°N, 160°W in the LIM forecasts (Fig. 8b) are somewhat weaker than in the PDO pattern and in most analyses of ENSO-related North Pacific SST anomalies.

While the deterministic matrix \( B \) derived here is not seasonally dependent, the optimal structure may still preferentially occur at a given time of year. The seasonal dependence of the projection of the OS on the observed SST field is examined in Fig. 10, which shows the number of times out of the 50-yr record that the pattern correlation between the two fields exceeds 0.4, \(-0.4\), or \( |0.4| \) as a function of the 3-month season. The OS tends to occur more frequently in the first half of the year with a maximum in spring. Although the tendency for the OS to occur in spring is based on a fairly small sample, it is consistent with the evolution of the physical system (i.e., the maximum gain in the LIM forecast from the OS occurs over approximately 6–9 months; Fig. 7), thus, a forecast initiated in spring peaks during boreal winter, which is also when the mature phase of ENSO occurs in nature.

c. Physical interpretation of the optimal structure

While LIM has identified the optimal structure as the SSTA pattern that is most likely to grow into a large ENSO and PDO event, the physical processes for why this occurs are unclear, as the OS does not have strong loading in the equatorial Pacific and the anomalies in the North Pacific are nearly in quadrature with the PDO pattern. Recent studies, however, have suggested that extratropical–tropical interactions via subtropical SSTAs may be important for the excitation of ENSO
events and Pacific decadal variability (Vimont et al. 2001, 2003a,b; Anderson 2003, 2004). Large fluctuations in internal atmospheric modes over the North Pacific in winter, especially the North Pacific Oscillation, (NPO) impart an SST “footprint” on the underlying ocean via changes in the surface heat fluxes. The SST footprint, which peaks in spring and persists through summer in the subtropics, impacts the atmospheric cir-

![Image of maps showing SST anomaly correlations](image1)

**FIG. 5.** Maps of the correlation between the actual and predicted SST anomalies for predictions of 3, 6, 9, and 12 months, averaged over the 6 verification periods spanning 1971–2000. Only values significant at the 95% confidence level in each 4° × 10° grid square are shown.

![Image of maps showing 12-month forecasts](image2)

**FIG. 6.** Same as in Fig. 5, but for the correlations between the predicted and observed SSTAs for 12-month forecasts initialized in JFM.

![Image of growth rate curve](image3)

**FIG. 7.** The maximum amplification curve: the energy amplification of a forecast optimized to give the greatest growth at 7 months, where the growth rate is given by the leading eigenvalue ($\gamma_l$) of the matrix $G^T G$ and growth occurs for $\gamma_l > 1$. 
calculation including zonal wind stress anomalies that extend onto the equator. This seasonal footprinting mechanism (SFM) also influences the latitude of the tropical SST gradient and the intertropical convergence zone (ITCZ): the ITCZ is displaced toward (away from) the hemisphere with anomalously warm (cold) water and the associated winds flow across the equator from the negative toward and over the positive SST. The meridional displacements of the SST gradient and the ITCZ, termed the meridional mode (MM) by Chiang and Vimont (2004), has been well documented in the Atlantic over the past 30 yr but was only recently uncovered in the Pacific after accounting for the dominant ENSO signal. While, the existence of the SFM and MM appear to be independent of ENSO, as they occur in coupled models where only thermodynamic ocean–atmosphere interaction is permitted, the equatorial wind associated with these modes may stochastically force the equatorial Pacific Ocean and thus be an effective generator of interannual and decadal ENSO variability (Vimont et al. 2001, 2003a; Chiang and Vimont 2004; Chang et al. 2007).

Thus, the question arises: is there a relationship between the optimal structure and the SFM/MM? The SST anomalies associated with the OS, SFM, and MM all peak in the northern subtropics during late winter and spring. Vimont et al. (2003a) and Chang et al. (2007) also mention that the SSTA pattern associated with the SFM and MM during boreal spring resemble the OS. Here, we derive ENSO precursors relative to the optimal structure and see if they are consistent with the SFM and MM analyses. Specifically, we construct composites based on when the spatial correlation between OS and the observed SSTA field is \( r > 0.4 \), \( r > 0.4 \) (light gray bar), and \( r < -0.4 \) (cross-hatched bar). Positive (negative) correlations indicate conditions that are favorable for the development of El Niño (La Niña) events.

Thus, the question arises: is there a relationship between the optimal structure and the SFM/MM? The SST anomalies associated with the OS, SFM, and MM all peak in the northern subtropics during late winter and spring. Vimont et al. (2003a) and Chang et al. (2007) also mention that the SSTA pattern associated with the SFM and MM during boreal spring resemble the OS. Here, we derive ENSO precursors relative to the optimal structure and see if they are consistent with the SFM and MM analyses. Specifically, we construct composites based on when the spatial correlation between OS and the observed SSTA field is \( r > 0.4 \) and \( < -0.4 \) during MAM. The difference between these + and − composites for 3-month seasons of SST/sea level pressure (SLP) and vector winds/net surface heat flux are shown in Fig. 11 for DJF, MAM, JJA, and NDJ.
corresponding to lags of −3, 0, +3, and +7 months. All of the variables are obtained from NCEP reanalysis for the period 1951–2000 and displayed on a 2.5° latitude × 2.5° longitude grid.

In the winter prior to when the OS peaks (lag −3), the SLP difference or anomaly field exhibits a meridional dipole between 170°E and 110°W, that resembles the NPO (e.g., Walker and Bliss 1932; Rogers 1981, 1990) and is consistent with the first phase of the SFM.

The SST field during this time has negative anomalies in the western Pacific between about 20° and 30°N and is ringed by positive anomalies. The low persists in the North Pacific and the positive SST anomalies amplify over the next 3 months. The SST anomalies during winter and spring appear to result from surface heat fluxes, including the subtropics (approximately 10°–25°N) where the flow along the southern flank of the low opposes the mean trades reducing the upward net heat
The positive SSTAs that extend from Baja California to the equator, 160°E (MAM, lag 0), are closely aligned with those in the SFM and MM (Vimont et al. 2003a; Chiang and Vimont 2004; Chang et al. 2007), which Vimont et al. (2003b) and Chiang and Vimont (2004) have shown influence the winds over the tropical Pacific. Feedback between the tropical winds, evaporation, and sea surface temperature [(WES) feedback], can act to amplify the SFM/MM climate anomalies, particularly in late spring and summer (Chang et al. 2007). From DJF (lag −3) through the following JJA (lag +3) the westerly winds intensify over the central and western equatorial Pacific, which can efficiently force downwelling equatorial Kelvin waves and lead to a mature El Niño event during the subsequent winter (NDJ, lag +7).

The time series of the springtime optimal structure, as given by the pattern correlation between the observed and OS SST pattern in each MAM from 1951 to 2000, along with indices associated with the SFM and MM, are shown in Fig. 12. The OS varies over a wide range of time scales, including decadal (e.g., the OS is generally positive for 1980–97). The wintertime NPO, as indicated by the second EOF of North Pacific SLP during DJF (not shown), strongly resembles the meridional dipole shown in Fig. 11a; the temporal correlation between the associated principal component (PC2; Fig. 12a) and the OS is 0.39 (significant at the 95% level) but increases to ~0.5 when the NPO is computed for FMAM. A second measure of the wintertime atmospheric circulation, the SLP anomaly in DJF over the subtropical North Pacific (15°–30°N, 170°–140°W, box in Fig. 11a) is also shown in Fig. 12. The correlation between the subtropical SLP and the OS time series is 0.63, but drops to 0.4 if the SLP region is shifted north to 25°–45°N. These results suggest, that while the NPO may be a precursor to the OS (and the SFM/MM), the key feature is the anomalous circulation on the southern flank of the NPO, namely the strength and position of the subtropical high.

The optimal structure also exhibits significant temporal relationships with tropical SFM/MM indices, including a measure of the meridional mode, the zonal winds in the western equatorial Pacific during summer, and the SSTs in the Niño-3.4 region in the following winter. The MM time series was computed by first determining the dominant covarying pattern between the SST and both components of the NCEP reanalysis surface winds in MAM over the Pacific from 20°N to 20°S during non-ENSO (neutral) years. Then the temporal variability was derived by projecting the monthly zonal wind stress anomaly onto the leading zonal wind stress pattern. Finally, the monthly time series was averaged

![Fig. 12. The time series of the optimal structure along with several indices associated with the SFM and the MM. (a) The PC2 of SLP over the North Pacific in winter, which corresponds to the temporal variability of the NPO pattern; (b) SLP anomalies over the subtropical North Pacific in winter; (c) OS time series, given by the pattern correlation between the optimal structure and the observed SSTAs in spring; (d) the MM time series derived from the dominant covarying pattern between the SST and the surface winds; (e) zonal equatorial wind anomalies in summer; (f) Niño-3.4 SST anomalies in the following winter. The correlation between the OS and the other time series is given above the corresponding time series (all are significant at the 95% level). All time series have been normalized by their respective std dev.](image)
over JFMAM to obtain the MM index shown in Fig. 12. The correlation between the OS and MM indices is 0.52 ($r = 0.61$ when the MM is derived from the 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40) for the period 1958–2000. The correlations between the OS and the equatorial winds ($0^\circ$–$10^\circ$N, $140^\circ$E–160$^\circ$W, box in Fig. 11f) in JJA and SSTAs in the Niño-3.4 region in NDJ are 0.75 and 0.73, respectively. Figure 12 indicates that there is temporal coherence between the optimal structure and components of the SFM/MM, particularly the circulation over the North Pacific in the previous winter, the SSTAs in the subtropics in spring, the winds in the western equatorial Pacific in the subsequent summer, and SSTAs in the ENSO region in the following winter.

4. Discussion and conclusions

A linear inverse model, the statistical forecast method used by Penland and collaborators to predict tropical SST anomalies, has been extended here to include the North Pacific. LIM assumes that the evolution of the system can be represented by linear dynamics forced by white noise. The model predictors and subsequent forecasts are based on 3-month running mean Pacific SST anomalies between $30^\circ$S and $60^\circ$N. The prediction coefficients in this study are not seasonally dependent. However, the forecast skill varies with the initial conditions and with the stochastic disruption of the prediction, both of which are a function of the time of year. The model exhibits significant skill over much of the Pacific for two to three seasons in advance and up to a year in some locations, particularly for forecasts initialized in boreal winter. We also investigated the model’s ability to predict the Pacific decadal oscillation (PDO), the time series giving the amplitude and sign of the leading pattern of North Pacific SST anomalies. The correlation between the predicted and observed PDO for all seasons during 1971–2001 is 0.81, 0.64, 0.55, and 0.44 at leads of 1, 2, 3, and 4 seasons, respectively, while the correlation value is $\sim 0.6$ for winter-to-winter predictions. These values are significant at the 95% level and are comparable to LIM-based forecasts of SSTs in the Niño-3.4 region.

Within LIM SST anomalies can grow due to constructive interference of the empirically determined modes of the system, even though the individual modes are damped over time. For the Pacific domain, the basinwide SST variance can grow for up to $\sim 14$ months, consistent with the skill of the actual predictions conducted here and the length of forecast skill of North Pacific SSTs estimated in other studies (Grötzner et al. 1999; Scott 2003; Schneider and Cornuelle 2005; Newman 2007). While the presence of decadal variability, especially at preferred time scales, suggests that skillful predictions might be possible at longer lead times, if the processes that have decadal periods are damped on interannual and shorter time scales, then the forecast skill will be limited to 1–2 yr.

The optimal structure, the initial pattern that LIM theory indicates should grow most rapidly with time, is clearly relevant to the prediction of Pacific SSTAs. The OS develops into a mature ENSO event 6–10 months later. The predicted SSTs in the North Pacific closely resemble ENSO-related anomalies due to the atmospheric bridge and coastal trapped waves. The former is due to ENSO-driven changes in the winds, air temperature, and humidity, and clouds that impact the underlying ocean via the surface heat, freshwater, and momentum fluxes, while the latter is generated either by equatorial Kelvin waves that propagate poleward upon reaching the coast or by wind stress changes over the eastern North Pacific. In addition, filtering the data using empirical normal modes of the full LIM system indicated that much of the PDO predictability resulted from modes associated with ENSO or a global trend, consistent with the findings of Livezey and Smith (1999). No filter, however, is perfect and some of the variability attributed to ENSO or global phenomena may originate in the North Pacific.

The tropical–extratropical connections may not be limited to ENSO time scales nor only directed from the equatorial to the North Pacific. For example, basinwide decadal variability is suggested by low-frequency normal modes that have strong loadings in both the tropics and extratropics as found here and in Newman (2007), as well as the strong coherence on decadal time scales between many tropical and North Pacific indices, including the PDO (Deser et al. 2004). In a LIM where the North Pacific and tropical Pacific SSTAs were designated as separate state vectors, including tropical–extratropical coupling was critical for obtaining realistic decadal variability in both regions (Newman 2007). The extratropics impacted the tropics both by the deterministic portion of the LIM, which may occur through either the atmosphere (Barnett et al. 1999; Pierce et al. 2000) or the ocean (Gu and Philander 1997; Kleeman et al. 1999; Yang et al. 2004), and by spatially coherent stochastic forcing, consistent with the seasonal footprinting mechanism.

Composite and time series analyses performed here revealed a close correspondence between the development of SST anomalies predicted by LIM and the evolution of the atmosphere–ocean system indicated by the SFM prior to ENSO/PDO events. During winters when the North Pacific oscillation, characterized by a merid-
ional SLP dipole over mid–high latitudes, is strong, the associated anomalous surface fluxes create an SST pattern in spring that strongly resembles the optimal structure. As in the SFM, the OS SSTA pattern is linked with anomalous westerly winds on the equator during spring and summer, which in turn can force oceanic Kelvin waves leading to ENSO events in the following winter. The optimal structure is also correlated with the meridional mode, a north–south dipole in tropical SST, precipitation, and winds, which has also been identified as a spring precursor for ENSO events. From our analyses, it is not possible to determine whether the SFM and/or MM would occur in the absence of ENSO, as suggested by Vimont et al. (2001, 2003a,b), Chiang and Vimont (2004), and Chang et al. (2007).

Other processes may also influence the optimal structure, its relationship to the SFM/MM, and how all three relate to subsequent ENSO events. While the OS is correlated with the NPO in the previous winter, it appears to be more strongly related to the atmospheric circulation along the southern edge of the NPO, suggesting that it is the subtropical atmosphere–ocean interactions that are critical to the formation of a subsequent ENSO and PDO events. Additional factors, which are at least partially independent from the OS/SFM/MM, impact the evolution of the equatorial Pacific Ocean. These include anomalous westerly winds in the central–western equatorial Pacific that are often associated with the Madden–Julian oscillation (Lau and Chan 1986, 1988; Moore and Kleeman 1999; Zavala-Garay 2008) and conditions in the Indian Ocean (PS95; Tourre and White 1997; Annamalai et al. 2005).

While tropical–extratropical teleconnections are clearly important for Pacific climate variability, skill in predicting North Pacific SSTAs, including the PDO, may also result from midlatitude phenomena. These include (i) thermal inertia, a process that occurs over the global oceans but can greatly enhance persistence in the extratropics during winter when the mixed layer is deep; (ii) the reemergence mechanism, where SST anomalies created in one winter are sequestered in the summer seasonal thermocline and are then reentrained into the mixed layer in the following fall and winter (Alexander and Deser 1995; Alexander et al. 1999; Deser et al. 2003); (iii) wind generated Rossby waves that propagate westward on interannual-to-decadal time scales and influence the strength and position of the ocean gyres and the SSTs in the Kuroshio region (e.g., Miller et al. 1998; Deser et al. 1999; Schneider and Miller 2001; Seager et al. 2001; Kwon and Deser 2007); and (iv) positive feedbacks between stratus clouds and SSTs (Norris and Leovy 1994; Norris et al. 1998; Park et al. 2006). In addition to the atmospheric bridge, the seasonality in forecast skill (see Fig. 4) may partly result from processes (i) to (iii), which most strongly impact SSTs during winter. The predictability from these processes, however, may not be fully realized without the explicit inclusion of the seasonal cycle and subsurface information in the forecast system.

Acknowledgments. Comments from Matt Newman and the two anonymous reviewers along with conversations with Gil Compo and Dan Vimont helped to significantly improve the manuscript. This work was supported by grants from NASA’s Interdisciplinary Science Program and NOAA’s Climate Variability and Predictability Program.

REFERENCES


Landsea, C. W., and J. A. Knaff, 2000: How much “skill” was


