

Impact of Annual Cycle on ENSO Variability and Predictability

SANG-IK SHIN,^{a,b} PRASHANT D. SARDESHMUKH,^{a,b} MATTHEW NEWMAN,^{a,b} CECILE PENLAND,^b
AND MICHAEL A. ALEXANDER^b

^aCIRES, University of Colorado Boulder, and NOAA Physical Sciences Laboratory, Boulder, Colorado

^bNOAA Physical Sciences Laboratory, Boulder, Colorado

(Manuscript received 22 April 2020, in final form 22 September 2020)

ABSTRACT: Low-order linear inverse models (LIMs) have been shown to be competitive with comprehensive coupled atmosphere–ocean models at reproducing many aspects of tropical oceanic variability and predictability. This paper presents an extended cyclostationary linear inverse model (CS-LIM) that includes the annual cycles of the background state and stochastic forcing of tropical sea surface temperature (SST) and sea surface height (SSH) anomalies. Compared to a traditional stationary LIM that ignores such annual cycles, the CS-LIM is better at representing the seasonal modulation of ENSO-related SST anomalies and their phase locking to the annual cycle. Its deterministic as well as probabilistic hindcast skill is comparable to the skill of the North American Multimodel Ensemble (NMME) of comprehensive global coupled models. The explicit inclusion of annual-cycle effects in the CS-LIM improves the forecast skill of both SST and SSH anomalies through SST–SSH coupling. The impact on the SSH skill is particularly marked at longer forecast lead times over the western Pacific and in the vicinity of the Pacific North Equatorial Countercurrent (NECC), consistent with westward propagating oceanic Rossby waves that reflect off the western boundaries as eastward propagating Kelvin waves and influence El Niño development in the region. The higher CS-LIM skill is thus associated with the improved representation of both ENSO phase-locking and Pacific NECC variations. These improvements result from explicitly accounting for not only the annual cycle of the background state, but also that of the stochastic forcing.

KEYWORDS: El Niño; Seasonal forecasting; Climate variability; Seasonal cycle

1. Introduction

Numerous studies have recognized the important role of tropical sea surface temperature (SST) variations in global climate variability, and progress in seasonal to interannual climate predictions around the globe has followed largely from this recognition (Goddard et al. 2001; Barnston et al. 2005). Anomalies associated with tropical El Niño–Southern Oscillation (ENSO) influence climate anomalies worldwide through atmospheric teleconnections (Ropelewski and Halpert 1986, 1987, 1989; Halpert and Ropelewski 1992; Klein et al. 1999; Alexander et al. 2002; Goddard and Dilley 2005). Improving ENSO predictions has therefore been an important goal of the seasonal forecasting community (see, e.g., Jin et al. 2008; Barnston et al. 2012, 2015; and references therein).

Despite significant progress, however, the ENSO prediction skill of comprehensive current climate models remains comparable to that of vastly simpler empirical models. For example, Newman and Sardeshmukh (2017) showed that the prediction skill of their simple linear inverse model (LIM) was very similar to that of the comprehensive coupled models comprising the North

American Multimodel Ensemble (NMME; Kirtman et al. 2014) used operationally for seasonal predictions. Indeed, they interpreted this very similar skill, and the fact that it was also very similar to the potential skill of their LIM, to conclude that one may now be near the intrinsic predictability limit of tropical SSTs. Given its practical implications, such a conclusion obviously warrants closer scrutiny. One way to challenge it would be to provide convincing evidence that the newest versions of the coupled models consistently beat the LIM's predictions. Another way would be to show that the potential skill estimated using a LIM is too low, possibly due to the LIM's low-dimensionality and assumption of linearity.

On the other hand, the LIMs could also be improved. LIMs are low-dimensional empirical–dynamical models derived from the observed zero-lag and time-lag covariances of climate anomalies, and have been used extensively in many climate diagnostic and predictability studies (e.g., Penland and Sardeshmukh 1995, hereafter PS95; Penland 1996, hereafter P96; Johnson et al. 2000b; Xue et al. 2000; Kondrashov et al. 2005; Penland and Matrosova 2006; Newman 2007; Alexander et al. 2008; Hawkins and Sutton 2009; Compo and Sardeshmukh 2010; Shin et al. 2010; Newman et al. 2011a,b; Vimont 2012; Vimont et al. 2014; Capotondi and Sardeshmukh 2015; Chen et al. 2016; Huddart et al. 2016; Newman et al. 2016; Newman and Sardeshmukh 2017; and many others). In this paper, we consider one way to improve LIMs by explicitly accounting for the seasonal variation of their parameters. We call such LIMs cyclostationary LIMs (CS-LIMs).

CS-LIMs may be partly motivated by the fact that the variance of tropical SST anomalies in the east-central equatorial Pacific, where ENSO anomalies attain their largest magnitude, has a pronounced annual cycle with a minimum in boreal spring and a maximum in boreal winter, even if the anomalies

Denotes content that is immediately available upon publication as open access.

Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JCLI-D-20-0291.s1>.

Corresponding author: Sang-Ik Shin, sangik.shin@noaa.gov

DOI: 10.1175/JCLI-D-20-0291.1

© 2020 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).

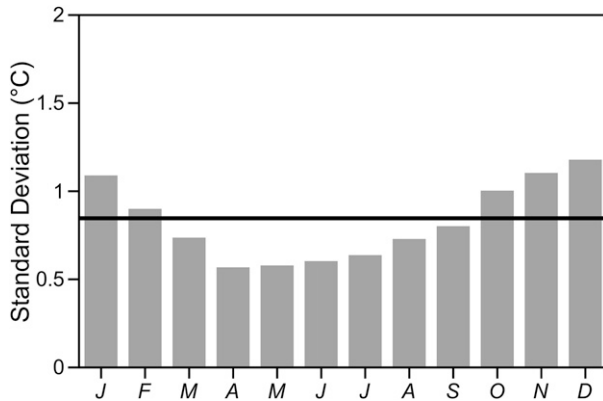


FIG. 1. Monthly standard deviations of Niño-3.4 (5°S–5°N; 170°–120°W) averaged SST anomalies during the period 1961–2010 derived from the HadISST dataset (Rayner et al. 2003). The thick black line shows the annual mean.

themselves are defined as departures from the climatological annual cycle (Fig. 1). The winter variance maximum coincides with the timing of peak warm and cold ENSO events. This synchronization is called “ENSO phase locking” and is evident both in observations (e.g., Rasmusson and Carpenter 1982) and in forward (as opposed to inverse) models of varying complexity (Tziperman et al. 1997, 1998; Harrison and Vecchi 1999; Ham and Kug 2014). In the context of the linear inverse models considered here, this raises a basic question as to what extent the annual cycle of the SST anomaly variance arises from the annual cycle of the predictable SST dynamics versus the annual cycle of the stochastic forcing of the SSTs. P96 presented evidence that the annual cycle of the stochastic forcing may be dominant in this regard. Some subsequent studies (Vimont et al. 2003; Chang et al. 2007) also emphasized the role of extratropical atmospheric stochastic forcing, mediated through the Pacific meridional mode (Chiang and Vimont 2004), in ENSO phase-locking. However, other studies, using simple theoretical models with seasonally varying parameters as well as seasonally varying Markov models (Blumenthal 1991; Johnson et al. 2000a,b; Xue et al. 2000; Thompson and Battisti 2000, 2001; Jin et al. 2019; Liu et al. 2019), emphasized the role of seasonal variations in the deterministic system dynamics. None of these studies quantified the relative roles of the seasonal variations of the predictable SST dynamics and unpredictable stochastic forcing in ENSO phase-locking, or in determining the seasonal variation of SST predictability. The CS-LIM presented here is especially well suited for such a diagnosis.

Section 2 gives a brief overview of tropical LIM construction and discusses some additional technical aspects that need to be considered when constructing a CS-LIM. Detailed intercomparisons of stationary and cyclostationary LIMs whose anomaly state vector comprise only tropical SST anomalies (“SST-only LIMs”) are provided in section 3. Section 4 investigates the extent to which the relative roles of the annual cycles of the predictable dynamics and unpredictable stochastic forcing inferred from the SST-only CS-LIM are also reproduced in large ensembles of coupled climate model simulations, specifically the preindustrial

climate simulations in the archive of phase 5 of the Coupled Model Intercomparison Project (CMIP5; Taylor et al. 2012) and the Large Ensemble simulations performed with the NCAR Community Earth System Model version 1 (CESM-LE; Kay et al. 2015). Section 5 presents an extended coupled SST–SSH CS-LIM whose anomaly state vector comprises both tropical SST and SSH anomalies, and compares its seasonally varying deterministic and probabilistic forecast skill with that of the NMME. A summary and concluding remarks follow in section 6.

2. Linear inverse model

In a linear inverse modeling framework (e.g., PS95), a model of evolving tropical climate anomalies is formulated in EOF space as

$$\frac{d\mathbf{x}}{dt} = \mathbf{L}\mathbf{x} + \mathbf{S}\boldsymbol{\eta}, \quad (1)$$

where \mathbf{x} is an N -component system state vector comprising the principal components (i.e., the time-evolving N EOF coefficients) of the climate anomalies, $\mathbf{L}(t)$ is an $N \times N$ deterministic feedback matrix, and all unpredictable nonlinear system dynamics are approximated as an N -component stochastic forcing vector $\mathbf{S}\boldsymbol{\eta}$, where $\mathbf{S}(t)$ is an $N \times M$ forcing amplitude matrix and $\boldsymbol{\eta}$ is an M -component vector of M independent unit-amplitude white noises. The amplitude matrix \mathbf{S} can be estimated from the covariance budget equation associated with (1), also known as the Lyapunov equation or the fluctuation–dissipation relation (FDR; e.g., Penland and Matrosova 1994), as

$$\frac{d\mathbf{C}(0)}{dt} = \mathbf{L}\mathbf{C}(0) + \mathbf{C}(0)\mathbf{L}^T + \mathbf{Q}, \quad (2)$$

where $\mathbf{Q} = \mathbf{S}\mathbf{S}^T$ and $\mathbf{C}(0)$ is the zero-lag covariance matrix of \mathbf{x} . Note that \mathbf{Q} must be positive semidefinite, but not necessarily diagonal.

Two types of stationarity assumptions are often used to reduce the complexity of (1). One is that the first and second statistical moments of \mathbf{x} are time-independent (i.e., stationary; e.g., PS95; P96; Newman et al. 2011a,b; Newman and Sardeshmukh 2017), and the other is that those moments depend on the annual cycle (i.e., cyclostationary; e.g., OrtizBeviá 1997; Johnson et al. 2000b; Kondrashov et al. 2005; Chen et al. 2016). Under the stationary assumption, the linear system matrix can be determined via an error minimization procedure (e.g., PS95) as

$$\mathbf{L}^{\text{ST}} = \tau^{-1} \ln[\mathbf{C}(\tau)\mathbf{C}(0)^{-1}], \quad (3)$$

where $\mathbf{C}(\tau) = \langle \mathbf{x}(t + \tau)\mathbf{x}^T(t) \rangle$ is the lag-covariance matrix of \mathbf{x} at time lag τ . Unless stated otherwise, we used a training lag $\tau = \tau_0 = 1$ month in this study. Then, using (2), \mathbf{Q}^{ST} can be estimated as

$$\mathbf{Q}^{\text{ST}} = -[\mathbf{L}^{\text{ST}}\mathbf{C}(0) + \mathbf{C}(0)(\mathbf{L}^{\text{ST}})^T]. \quad (4)$$

In the cyclostationary case, two approaches for estimating \mathbf{L} have been suggested, referred to as “phase-averaged” and “fixed-phase” approaches (OrtizBeviá 1997). In the former, it is assumed that $\mathbf{L}^{\text{CS}}(t) = \mathbf{L}_0 + \mathbf{L}_1 \cos\omega t + \mathbf{L}_2 \sin\omega t$, where $\omega = 2\pi/T$

and $T = 12$ months (in our problem). The constant matrices \mathbf{L}_0 , \mathbf{L}_1 , and \mathbf{L}_2 can then be determined by solving a linear system of equations derived in section 3 of [OrtizBeviá \(1997\)](#). This phase-averaged approach is relatively less vulnerable to sampling uncertainties because of its use of the entire empirical record; however, it necessarily introduces an ambiguity of phase ([OrtizBeviá 1997](#)). In the fixed-phase approach, \mathbf{L}^{CS} is estimated by applying the error minimization procedure (3) separately to each calendar month j , that is,

$$\mathbf{L}_j^{\text{CS}} = \tau_0^{-1} \ln[\mathbf{C}_j(1) \mathbf{C}_j(0)^{-1}], \quad j = 1, \dots, 12, \quad (5)$$

where $\mathbf{C}_j(1)$ and $\mathbf{C}_j(0)$ are the lag-1 and lag-0 covariance matrices of \mathbf{x} for month j . This fixed-phase approach retains the correct annual phase; however, it tends to amplify sampling uncertainties by dividing the empirical data series into 12 shorter subseries.

Having estimated \mathbf{L}^{CS} using either of the above approaches, \mathbf{Q}^{CS} can be estimated via (2) as

$$\mathbf{Q}_j^{\text{CS}} = \frac{\mathbf{C}_{j+1}(0) - \mathbf{C}_{j-1}(0)}{2\Delta t} - [\mathbf{L}_j^{\text{CS}} \mathbf{C}_j(0) + \mathbf{C}_j(0)(\mathbf{L}_j^{\text{CS}})^T], \quad (6)$$

where $\Delta t = 1$ month and $\mathbf{C}(0)$ is cyclostationary; that is, $\mathbf{C}_j(0) = \mathbf{C}_{j+T}(0)$.

[P96](#) assumed that the major cyclostationary dependence is in the stochastic forcing and not in the deterministic dynamics of the system (i.e., in \mathbf{Q} and not in \mathbf{L}). This approach was also used by [Alexander et al. \(2008\)](#), [Newman et al. \(2011b\)](#), and others. In this case (hereafter CSQ-LIM), \mathbf{L}^{ST} is used to estimate the cyclostationary $\hat{\mathbf{Q}}_j^{\text{CS}}$ for each month as

$$\hat{\mathbf{Q}}_j^{\text{CS}} = \frac{\mathbf{C}_{j+1}(0) - \mathbf{C}_{j-1}(0)}{2\Delta t} - [\mathbf{L}^{\text{ST}} \mathbf{C}_j(0) + \mathbf{C}_j(0)(\mathbf{L}^{\text{ST}})^T]. \quad (7)$$

3. SST-only LIMs

We first consider annual-cycle effects in the simplest LIM framework, in which the state vector \mathbf{x} comprises only SST anomalies. We call such LIMs ‘‘SST-only LIMs.’’ We constructed this LIM using monthly SST data from the HadISST dataset ([Rayner et al. 2003](#)) over the tropics (25°S–25°N) for the period 1961–2010. To determine the monthly anomalies on a 1° latitude \times 1° longitude grid, the grand mean and the climatological annual cycle were removed at each grid point. An EOF analysis of the resulting anomalous SSTs was then performed, and the dominant 15 EOFs explaining about 80.4% of total SST anomaly variance over the domain were retained. All subsequent calculations were performed in this 15-dimensional EOF space (i.e., by defining the state vector \mathbf{x} as a 15-component vector representing the time-varying amplitudes of the 15 EOF patterns), and then transformed to geographical space for presentation and interpretation.

a. Stability of the linear system

The linear system matrix \mathbf{L} must be dissipative to satisfy the stability of the model (1), regardless of the stationarity

assumptions used. Under statistical stationarity, Eq. (1) without forcing, $d\mathbf{x}/dt = \mathbf{L}^{\text{ST}}\mathbf{x}$, has general eigensolutions of the form

$$\mathbf{x}_\alpha = \mathbf{u}_\alpha \exp(\beta_\alpha t) c_\alpha, \quad (8)$$

where \mathbf{u}_α is the (normalized) α th eigenvector of \mathbf{L}^{ST} , β_α is the corresponding eigenvalue, and c_α is the eigenvector amplitude at time $t = 0$. These eigenvector solutions are also referred to as the principal oscillation patterns (POPs; e.g., [von Storch et al. 1995](#)) of the system. System stability requires that all β_α have negative real parts, and the stationary LIM (hereafter ST-LIM) constructed using 15 EOFs meets this condition.

Under cyclostationarity with period T , Eq. (1) without forcing, $d\mathbf{x}/dt = \mathbf{L}^{\text{CS}}\mathbf{x}$, has general solutions given by Floquet theory (e.g., [Grimshaw 1993](#)) of the form

$$\mathbf{x}_\kappa = \mathbf{p}_\kappa \exp(\mu_\kappa t) d_\kappa, \quad (9)$$

where \mathbf{p}_κ is the κ th cyclic eigenvector with period T , the time-independent μ_κ is the corresponding complex Floquet exponent, and d_κ is the eigenvector amplitude at time $t = 0$. These solutions are also referred as cyclostationary POPs in the literature (e.g., [Blumenthal 1991](#); [von Storch et al. 1995](#)). Stability of the system in this case requires that all Floquet exponents have negative real parts.

In the phase-averaged approach, the Floquet exponents can be determined numerically (e.g., [Wang and Hale 2001](#)) by integrating $d\mathbf{A}/dt = \mathbf{L}^{\text{CS}}\mathbf{A}$ over one period T from an initial condition $\mathbf{A}(0) = \mathbf{I}$ to form a nonsingular fundamental matrix $\mathbf{A}(T)$. The time-independent Floquet multipliers (ρ_κ) are the eigenvalues of the monodromy (characteristic) matrix $\mathbf{M} = \mathbf{A}(T)$, and the Floquet exponents are determined as $\mu_\kappa = \ln(\rho_\kappa)/T$. In the fixed-phase approach, the monodromy matrix of discretized data can be estimated as $\mathbf{M} = \mathbf{G}_T^{\text{CS}} \mathbf{G}_{T-1}^{\text{CS}} \cdots \mathbf{G}_2^{\text{CS}} \mathbf{G}_1^{\text{CS}}$, where $\mathbf{G}_j^{\text{CS}} = \exp(\mathbf{L}_j^{\text{CS}})$ is the 1-month linear system propagator at month j , and $\mathbf{G}_j^{\text{CS}} = \mathbf{G}_{j+T}^{\text{CS}}$. Note that overall system stability requires that the magnitudes of the Floquet multipliers (the eigenvalues of \mathbf{M}) are less than 1. However, the magnitudes of the eigenvalues of some of the constituent \mathbf{G}_j^{CS} can be greater than 1 without violating this requirement. When this happens, one may say that the corresponding eigenmode is ‘‘temporarily unstable’’ during some phases of its cyclostationary evolution. For future reference, we note that this is true of the dominant ENSO eigenmode in both our ‘‘SST-only’’ CS-LIM (not shown) and the ‘‘coupled SST–SSH’’ CS-LIM (see [Fig. 13d](#)).

[Figure 2](#) shows the sensitivity of the maximum real part of the Floquet exponents of the phase-averaged and fixed-phase CS-LIMs to the EOF truncation of the training data. Interestingly, the phase-averaged CS-LIM is unstable if more than 3 EOFs are used, whereas the fixed-phase CS-LIM remains stable using as many as 16 EOFs. We therefore decided to use the fixed-phase approach in this study, motivated by the need to retain enough SST variance in the truncated EOF space to diagnose the cyclostationary dynamics of a large fraction of tropical SST variability.

b. Numerical integration

The stationary stochastic forcing covariance matrix \mathbf{Q}^{ST} estimated using (4) is positive definite. However, the matrices

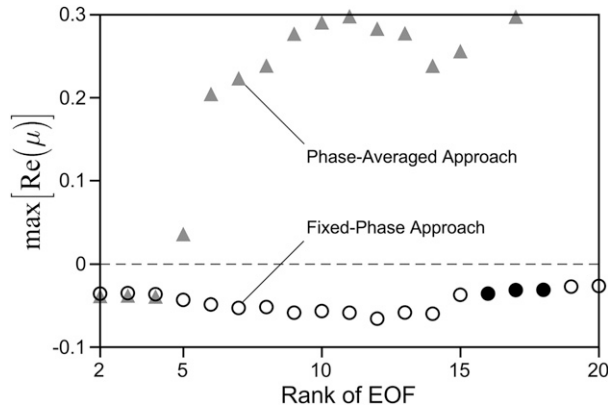


FIG. 2. The maximum real part of Floquet exponents (μ) estimated using the phase-averaged (gray triangles) and the fixed-phase (black circles) approaches and EOF truncations of the data ranging from 2 to 20 EOFs. The filled black circles indicate cases in which one or more of the 12 monthly linear system matrices used in the fixed-phase approach are imaginary.

\mathbf{Q}^{CS} and $\hat{\mathbf{Q}}^{\text{CS}}$ estimated from (6) and (7) have four to five negative eigenvalues, possibly due to the limited sample sizes available for their estimation. As in other LIM studies (e.g., P96), we adjusted these matrices by setting their negative eigenvalues to zero, and then rescaling the positive ones to conserve the total variance of the forcing (i.e., the trace of the original matrix). Examining the variance maps (i.e., the diagonal elements of the \mathbf{Q} matrices transformed to geographical space) of the adjusted and original forcing reveals only minor differences (see Fig. S1 in the online supplemental material).

Having estimated the linear system matrices (\mathbf{L}^{ST} and \mathbf{L}^{CS}) and stochastic forcing covariance matrices (\mathbf{Q}^{ST} , \mathbf{Q}^{CS} , and $\hat{\mathbf{Q}}^{\text{CS}}$), we numerically integrated the tropical SST-only LIMs with a time step $\Delta t \sim 1$ day (see PS95 for the details). Thus, in the cyclostationary case, we used

$$\begin{aligned} \mathbf{y}(t + \Delta t) &= [\mathbf{I} + \mathbf{L}^{\text{CS}}(t)\Delta t]\mathbf{y}(t) + \sqrt{\Delta t}\mathbf{S}^{\text{CS}}(t)\mathbf{r}(t), \\ \mathbf{x}(t + \Delta t/2) &= [\mathbf{y}(t + \Delta t) + \mathbf{y}(t)]/2 \end{aligned} \quad (10)$$

with $\mathbf{x}(0) = \mathbf{y}(0) = 0$, where \mathbf{L}^{CS} and \mathbf{S}^{CS} are discrete matrices (constant within each month) that are periodic in time (period = 12 months), and \mathbf{r} is a random Gaussian noise vector whose components have zero mean and unit variance and are generated using the Mersenne twister (Matsumoto and Nishimura 1998) at each time step. The stochastic forcing amplitude matrices were specified as $\mathbf{S}^{\text{CS}} = \Psi^{\text{CS}}(\mathbf{q}^{\text{CS}})^{1/2}$, where Ψ^{CS} and \mathbf{q}^{CS} are the discrete periodic matrices of the eigenvectors and eigenvalues of \mathbf{Q}^{CS} , respectively. We integrated three models, CS-LIM (\mathbf{L}^{CS} and \mathbf{Q}^{CS}), CSQ-LIM (\mathbf{L}^{ST} and $\hat{\mathbf{Q}}^{\text{CS}}$), and ST-LIM (\mathbf{L}^{ST} and \mathbf{Q}^{ST}), for 80 100 years from initial conditions $\mathbf{x}(0) = 0$, and used the last 80 000 years for analysis.

c. ENSO phase locking

We divided each 80 000-yr LIM run into 1600 50-yr segments and determined the 1600 sample standard deviations and power spectra of the simulated Niño-3.4 SST anomalies. The median and 5th and 95th percentiles of the standard deviations

and power spectra are shown in Fig. 3. All three LIMs reproduce the main features of the observed Niño-3.4 spectrum, also shown in the figure, with a peak near 4-yr periods. In general, the LIM spectra are smoother than observed, not only because the LIMs were constructed in a truncated EOF space but also because the observed spectrum is noisier due to limited sampling. It is also evident that the CS-LIM and CSQ-LIM spectra match the observed spectrum somewhat better than does the ST-LIM spectrum. All three LIMs, however, exaggerate the power at frequencies lower than ENSO frequencies.

The seasonally varying standard deviations of observed monthly Niño-3.4 SST anomalies are indicated by the black dots in the lower panels of Fig. 3. The standard deviations of the ST-LIM are by definition constant and cannot explain the observed annual cycle. On the other hand, introducing cyclostationarity either in the CSQ-LIM ($\hat{\mathbf{Q}}^{\text{CS}}$) or the CS-LIM (\mathbf{L}^{CS} and \mathbf{Q}^{CS}) clearly improves the seasonal variation of ENSO-related SST anomalies such that the largest (weakest) variability occurs in boreal winter (summer). Note also that the CS-LIM correctly locates the variance minimum in boreal spring, whereas the CSQ-LIM shifts it to summer and fall. This suggests that the cyclostationarity of the SSTs does not arise solely from the cyclostationarity of the stochastic forcing.

ENSO phase locking is further assessed in Fig. 4. This was done, as in P96, by calculating the pattern correlations of the observed leading SST EOF pattern (not shown) with the monthly LIM-generated SST anomaly patterns in each one of the 1600 50-yr segments of the simulated SST time series, and counting the number of warm (cold) events in each calendar month in which the pattern correlation was higher (lower) than 0.75 (-0.75). The CS-LIM simulates ENSO phase locking better than the other LIMs, since the warm and cold events tend to peak primarily in winter as observed. The CSQ-LIM has only weak phase locking, and the ST-LIM cannot represent phase locking at all.

One could argue that even the weak phase-locking in the CSQ-LIM may be exaggerated, since the seasonal variation of \mathbf{Q} estimated by specifying \mathbf{L}^{ST} in (7) may implicitly include the seasonal variation of \mathbf{L} through the observed $d\mathbf{C}(0)/dt$ term. On the other hand, the estimated seasonal variation of \mathbf{L}^{CS} could also be partly due to an implicit contribution of the seasonal variation of \mathbf{Q}^{CS} to $\mathbf{C}(0)$ in (5). To get a better sense of these implicit effects, we decomposed our estimated \mathbf{L}^{CS} and \mathbf{Q}^{CS} matrices into time-mean and seasonally varying parts as

$$\begin{aligned} \mathbf{L}^{\text{CS}} &= \overline{\mathbf{L}^{\text{CS}}} + \sum_{n=1}^{\infty} [\mathbf{a}_n \cos(n\omega t) + \mathbf{b}_n \sin(n\omega t)], \\ \mathbf{Q}^{\text{CS}} &= \overline{\mathbf{Q}^{\text{CS}}} + \sum_{n=1}^{\infty} [\mathbf{c}_n \cos(n\omega t) + \mathbf{d}_n \sin(n\omega t)], \end{aligned} \quad (11)$$

and suppressed their seasonal variations by setting $\mathbf{Q}^{\text{CS}} \approx \overline{\mathbf{Q}^{\text{CS}}}$ and $\mathbf{L}^{\text{CS}} \approx \overline{\mathbf{L}^{\text{CS}}}$. Note that $\mathbf{Q}^{\text{CS}} \neq \mathbf{Q}^{\text{ST}}$ and $\mathbf{L}^{\text{CS}} \neq \mathbf{L}^{\text{ST}}$. We then performed two additional 80 100-yr integrations similar to those described above, one with suppressed seasonal variations of \mathbf{Q} (\mathbf{L}^{CS} and $\overline{\mathbf{Q}^{\text{CS}}}$; CSL*-LIM) and the other with suppressed seasonal variations of \mathbf{L} ($\overline{\mathbf{L}^{\text{CS}}}$ and \mathbf{Q}^{CS} ; CSQ*-LIM). The result for CSL*-LIM in Fig. 5a is qualitatively similar to that for the CS-LIM and observations in Fig. 3, albeit with an overall

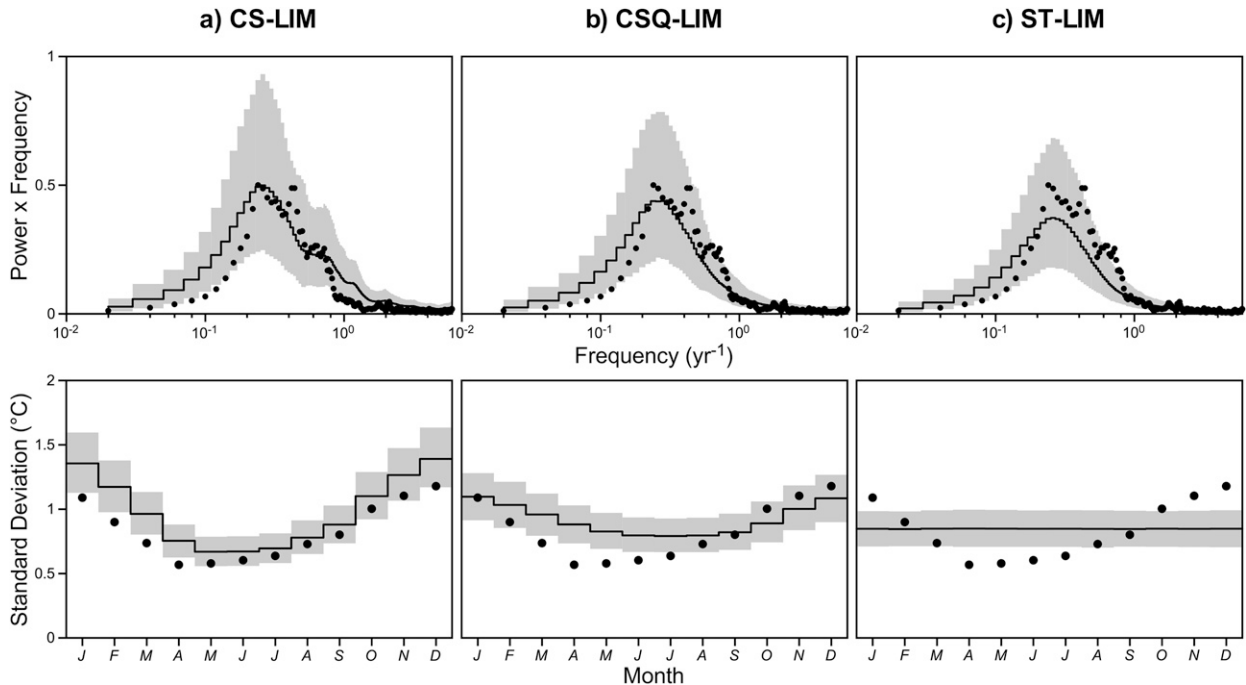


FIG. 3. (top) Power spectra and (bottom) interannual standard deviations of monthly Niño-3.4 (5°S – 5°N ; 170° – 120°W) SST anomalies derived from 80 000-yr SST-only (a) CS-LIM, (b) CSQ-LIM, and (c) ST-LIM runs. Thick black lines show the median of 1600 values derived from 1600 50-yr segments and gray shading the 90% confidence interval. The black dots show values estimated from the 50-yr HadISST (1961–2010; Rayner et al. 2003) dataset.

positive variance bias. On the other hand, the result for CSQ*-LIM in Fig. 5b is even weaker than that for CSQ-LIM in Fig. 3, with hardly any ENSO phase locking.

This is not the only way to account for the implicit contribution of the seasonal variation of \mathbf{L}^{CS} in the estimation of \mathbf{Q}^{CS} . Nonetheless, it suggests a weak role for \mathbf{Q}^{CS} and strong role for \mathbf{L}^{CS} in ENSO phase locking, consistent with many previous studies (Johnson et al. 2000a,b; Thompson and Battisti 2000, 2001; Jin et al. 2019; Liu et al. 2019). The seasonal variations of \mathbf{Q}^{CS} are not entirely unimportant, however, as evident from the variance bias introduced in Fig. 5a by neglecting them. Also, the relative importance of \mathbf{L}^{CS} and \mathbf{Q}^{CS} in ENSO predictability, as opposed to their relative importance in ENSO variability, is a different matter that we have not addressed here, although it is worth noting that stationary LIMs that ignore seasonal variations of \mathbf{L} (e.g., Newman and Sardeshmukh 2017) are well able to capture the seasonally varying prediction skill as well as potential predictability of tropical SST anomalies.

d. The 2015–16 El Niño event

The ability of the CS-LIM to represent ENSO phase-locking results in improved hindcasts of ENSO evolution. This is clearly evident in hindcasts of the 2015–18 period, which lies outside the CS-LIM’s training period (1961–2010). Figure 6 shows 64-member ensemble hindcasts of Niño-3.4 SST anomalies during the 2015–16 El Niño event, generated using the CS-LIM and ST-LIM. Hindcasts initialized in June and December of 2015 are shown separately to assess the LIMs’ ability to capture the developing and decaying phases of the event. In both phases, the median CS-LIM hindcast captures the observed trajectory of the

Niño-3.4 SST anomalies better than the median ST-LIM hindcast, which decays monotonically from the initial conditions. The CS-LIM also represents the observed evolution better beyond 12 months.

4. Annual cycles of \mathbf{L} and \mathbf{Q} in climate models

To further assess the importance and robustness of the annual cycles of \mathbf{L} and \mathbf{Q} , we examined them in long preindustrial runs of five climate models available in the CMIP5 archive (piControl; see Table 1), and also in 40-member ensemble simulations of the 1961–2010 period generated with the NCAR-CESM1 model (CESM-LE). The CMIP5 piControl simulations were generated with fixed preindustrial forcings associated with fixed atmospheric concentrations of trace gases and aerosols, as well as unperturbed land use (Taylor et al. 2012). The CESM-LE simulations were generated with prescribed historical forcings for the twentieth century and the representative concentration pathway 8.5 (RCP8.5) forcings for the twenty-first century (Kay et al. 2015). In this study, we used only CMIP5 piControl simulations longer than 500 years, and only the last 500 years of those simulations. For the CESM-LE simulations, we used the simulations of the 1961–2010 period, to be consistent with the period of the HadISST observations used to construct our LIMs. For reference, the standard deviations of Niño-3.4 SST anomalies in the model simulations are compared with the observed standard deviation in the HadISST dataset in Table 2. More detailed information on the representation of ENSO in these and other CMIP5 piControl simulations may be found in Bellenger et al. (2014).

All model simulations were interpolated to a common 2° latitude \times 2° longitude grid prior to the analysis. The mean biases were corrected

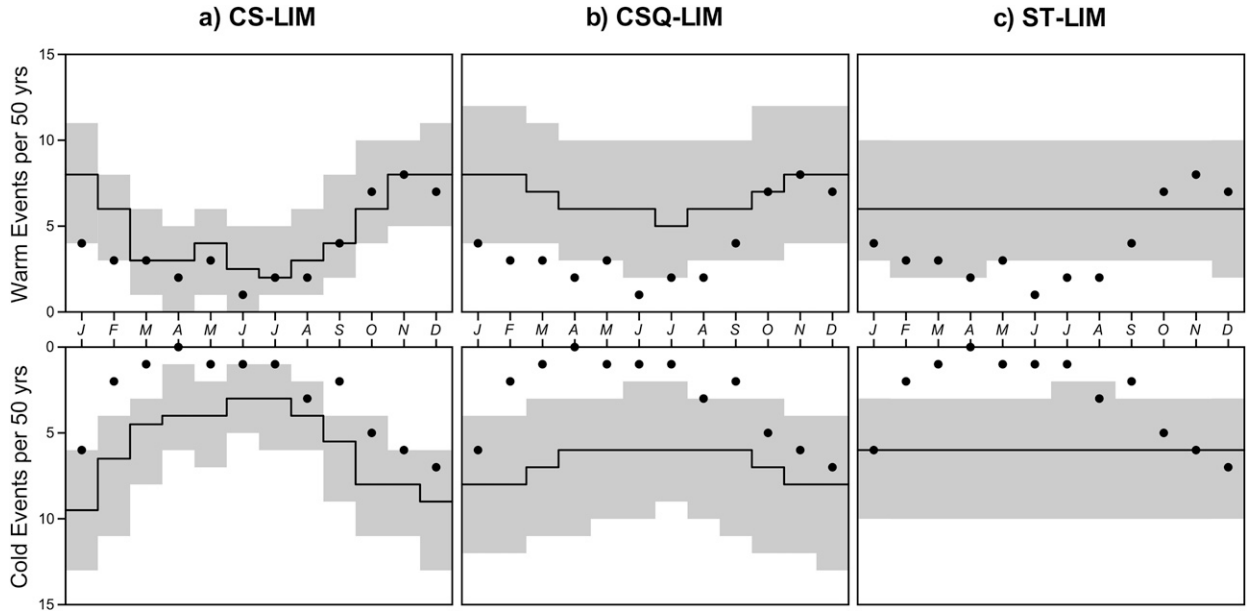


FIG. 4. (top) Number of times the pattern correlation of the simulated SST anomalies and the observed EOF-1 pattern (figure not shown) was higher than 0.75 in 50-yr segments of 80 000-yr runs of the SST-only (a) CS-LIM, (b) CSQ-LIM, and (c) ST-LIM. Thick black lines denote the median of the 1600 values obtained from the 1600 50-yr segments and gray shading the 90% confidence interval. (bottom) As at top, but for pattern correlation lower than -0.75 . Note that that ordinate is reversed from that in the top panel. The black dots show values estimated from the 50-yr HadISST (1961–2010; Rayner et al. 2003) dataset.

for each ensemble member by subtracting the ensemble grand mean and climatological annual cycle at each grid point. An EOF analysis was then performed, using the covariance matrix \mathbf{C} of the n -member ensemble of m -month runs, and the principal components of the k th ensemble member, \mathbf{x}_k , determined as

$$\mathbf{C} = \frac{1}{m(n-1)} \sum_{k=1}^n \mathbf{x}_k \mathbf{x}_k^T = \mathbf{E} \boldsymbol{\sigma} \mathbf{E}^T,$$

and

$$\mathbf{x}_k = \mathbf{E}^T \mathbf{X}_k, \tag{12}$$

where the matrix \mathbf{X}_k represents the anomaly data in the k th ensemble member whose element X_{ab} denotes SST anomalies at spatial location a in month b . For CMIP5 piControl, $n = 5$ and $m = 500 \times 12 = 6000$. For CESM-LE, $n = 40$ and $m = 50 \times 12 = 600$. The matrix $\boldsymbol{\sigma}$ is a diagonal matrix of the eigenvalues of \mathbf{C} and the corresponding column vectors of \mathbf{E} are the

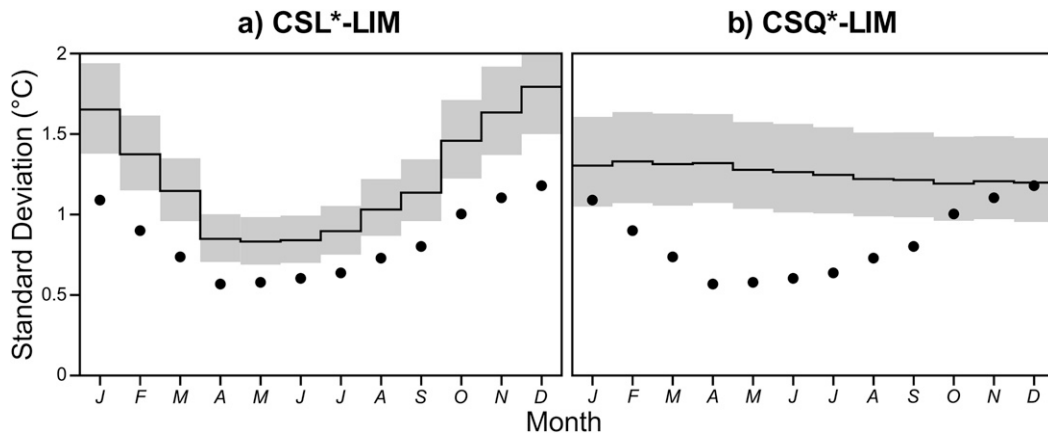


FIG. 5. Interannual standard deviations of monthly Niño-3.4 (5°S – 5°N , 170° – 120°W) SST anomalies derived from 80 000-yr SST-only (a) CSL*-LIM, and (b) CSQ*-LIM runs. Thick black lines show the median of 1600 values derived from 1600 50-yr segments and gray shading the 90% confidence interval. The black dots show values estimated from the 50-yr HadISST (1961–2010; Rayner et al. 2003) dataset.

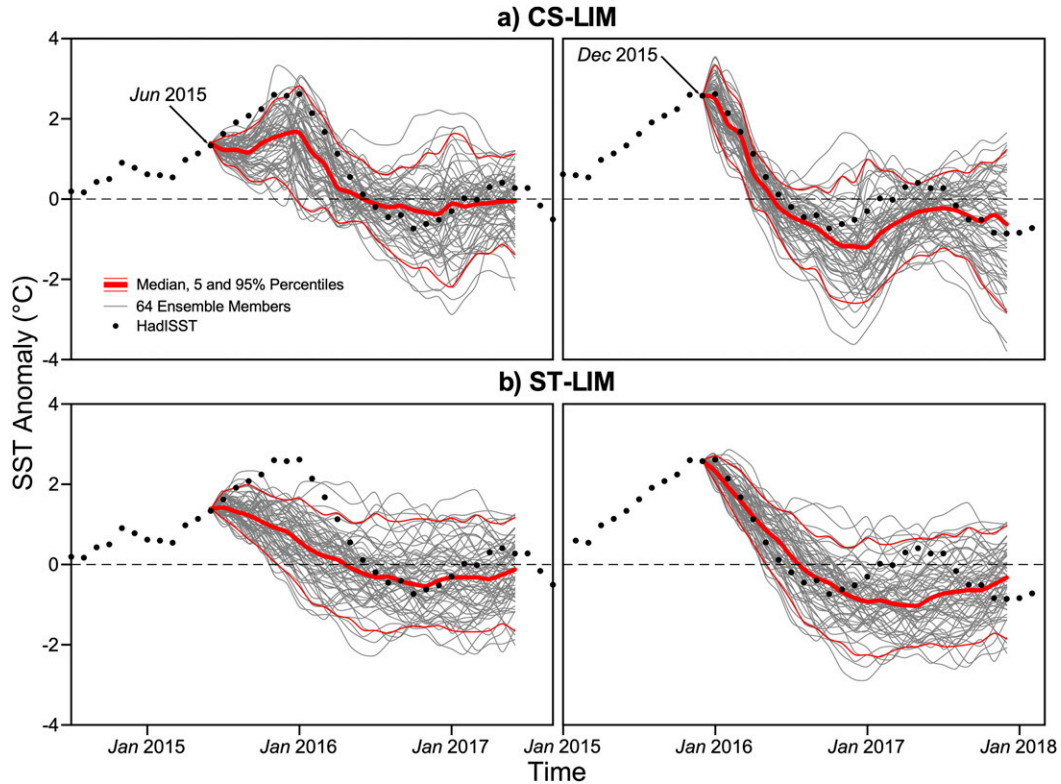


FIG. 6. The 64-member ensemble hindcasts (thin gray lines) of the 2015–16 El Niño event using the SST-only (a) CS-LIM and (b) ST-LIM. The hindcasts were initialized in (left) January 2015 and (right) December 2015. Thick red lines show the median of the ensemble and thin red lines the 90% confidence interval. Black dots show the observed SST anomalies estimated from the 50-yr HadISST (1961–2010; Rayner et al. 2003) dataset.

corresponding eigenvectors of \mathbf{C} . For consistency with the observational analysis, we retained the dominant 15 EOFs to construct CS-LIMs for each ensemble member as well as for all the members combined. These long model runs and large ensemble simulations provide a unique opportunity to assess the robustness of the seasonal cycles of \mathbf{L} and \mathbf{Q} inferred from a single 50-yr observational record.

An important feature of LIMs is their ability to identify optimal initial perturbations with the potential to grow over finite time intervals before decaying, as they eventually must if \mathbf{L} is stable. Such growth can be quantified via a singular value

decomposition (SVD) of the system propagator matrix \mathbf{G} (e.g., Farrell 1988 and PS95; also many others). Under the statistical stationarity assumption, the SVD of $\mathbf{G}^{\text{ST}}(\tau)$ yields a dominant pair of normalized singular vectors ($\mathbf{u}_1^{\text{ST}}, \mathbf{v}_1^{\text{ST}}$) associated with a dominant singular value λ_1 . In such a linear system, an optimal initial anomaly vector \mathbf{v}_1 grows maximally to the final anomaly vector $\lambda_1 \mathbf{u}_1$ at time $t = \tau$. In the cyclostationary case, the system propagator \mathbf{G}^{CS} and thus the dominant singular value depends on the starting calendar month of the forecast. For instance, for forecasts initialized in month j , the SVD of $\mathbf{G}_j^{\text{CS}}(\tau) = \mathbf{G}_{j+\tau-1}^{\text{CS}} \mathbf{G}_{j+\tau-2}^{\text{CS}} \cdots \mathbf{G}_{j+1}^{\text{CS}} \mathbf{G}_j^{\text{CS}}$ yields a dominant pair of normalized singular vectors ($\mathbf{u}_{j,1}, \mathbf{v}_{j,1}$) associated with the dominant singular

TABLE 1. Coupled climate model simulations used. All data are available in the World Climate Research Programme (WCRP) CMIP5 repository (<https://esgf-node.llnl.gov/projects/cmip5/>). We followed the nomenclature used in the repository for the models and modeling centers; N denotes ensemble size.

Model	Modeling center	Tropical ocean min (lat) \times lon	Years used	N	Reference
CMIP5 preindustrial control simulations					
CanESM2	CCCma	$0.9^\circ \times 1.4^\circ$	Last 500 years	1	Yang and Saenko (2012)
CCSM4	NCAR	$0.25^\circ \times 0.9^\circ$	Last 500 years	1	Gent et al. (2011)
ESM2M	NOAA GFDL	$0.33^\circ \times 1^\circ$	Last 500 years	1	Dunne et al. (2012)
ESM-LR	MPIM	Marsland et al. (2003)	Last 500 years	1	Raddatz et al. (2007)
CGCM3	MRI	$0.5^\circ \times 1^\circ$	Last 500 years	1	Yukimoto et al. (2011)
CESM Large Ensemble simulations					
CESM1	NCAR	$0.27^\circ \times 1.11^\circ$	50 years (1961–2010)	40	Kay et al. (2015)

TABLE 2. Standard deviations of monthly Niño-3.4 (5°S–5°N, 170°–120°W) SST anomalies derived from the coupled climate model simulations (see Table 1) and HadISST dataset (Rayner et al. 2003). For the CESM-LE, the interensemble spread (one standard deviation) is shown in parentheses.

CanESM2	CCSM4	ESM2M	ESM-LR	CGCM3	CESM-LE	Obs
0.91	1.06	1.15	0.80	0.59	1.10 (0.09)	0.84

value $\lambda_{j,1}$. In this case an optimal initial vector $\mathbf{v}_{j,1}$ grows maximally to the final vector $\lambda_{j,1}\mathbf{u}_{j,1}$ at forecast lead time $t = \tau$. This provides a simple and robust way to characterize the seasonal dependence of \mathbf{L}^{CS} by examining the seasonal dependence of the maximum singular values for growth over selected time intervals from each initialization month. Here we focused on growth over $\tau = 7$ and 12 months, guided by the major peaks of the maximum amplification curve (depicting $\lambda_{j,1}$ as a function of τ) of the SST-only CS-LIM (figure not shown).

In Fig. 7, all CMIP5 models show some seasonal variation of the maximum singular values, although their magnitude differs substantially among the models. For time intervals $\tau = 7$, the springtime maximum is well separated from the early fall minimum. Even for $\tau = 12$, a clear summer minimum is

evident. Similar seasonal variation is apparent in the CESM-LE ensemble. The individual ensemble members, each 50 years long, also show this behavior, although the singular values tend to be biased toward larger values than obtained from the combined ensemble equivalent to a $40 \times 50 = 2000$ -yr simulation.

Figure 8 quantifies the relative importance of the seasonal variations of \mathbf{L} and \mathbf{Q} in the seasonally varying FDR budget [Eq. (6)]. This was done for all the CMIP5 models combined (equivalent to a 2500-yr simulation) and all the CESM-LE members combined. For brevity, only the seasonal variations of the variance [that is, the traces of the matrix terms in (6)] are shown. Both the CMIP5 and CESM-LE simulations indicate that the seasonal variations of \mathbf{L} are relatively more important

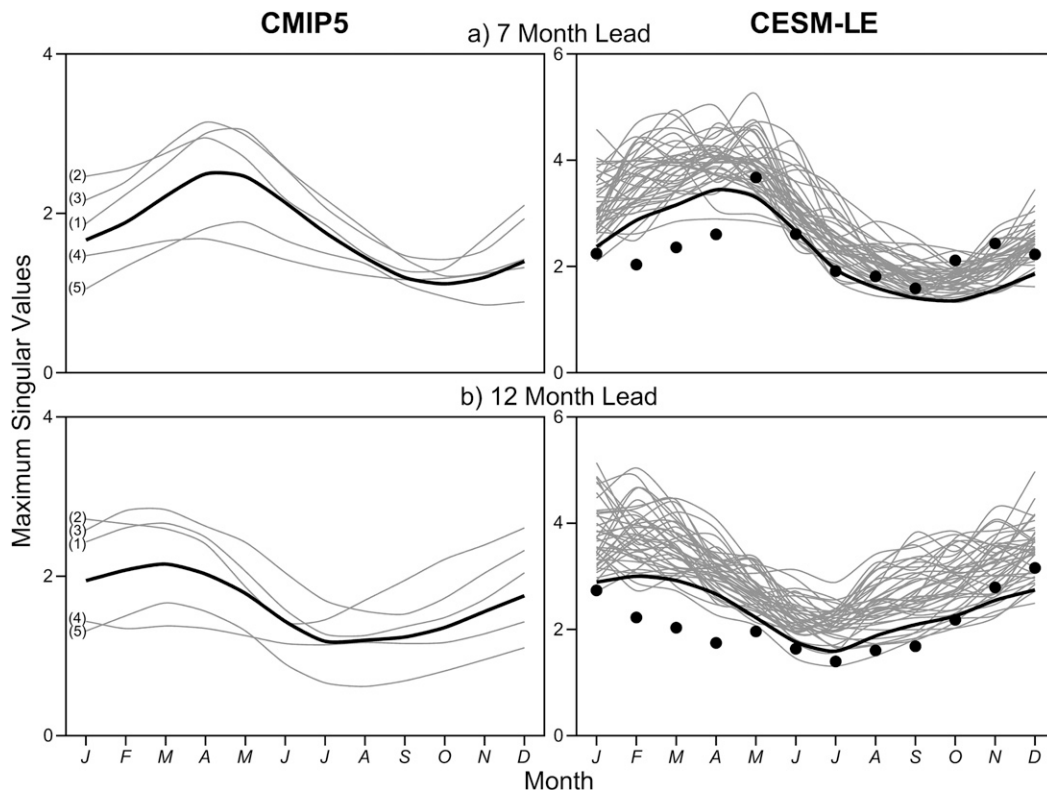


FIG. 7. Gray curves: maximum singular values of \mathbf{G}_j^{CS} at (a) 7-month and (b) 12-month leads for hindcasts initialized in month j , derived from (left) five 500-yr CMIP5 and (right) forty 50-yr CESM-LE simulations (see Table 1 for details). Thick black lines: values obtained using combined multimodel ensemble at left and all ensemble members at right. Black dots in the right panels show corresponding values derived from the 50-yr HadISST (1961–2010; Rayner et al. 2003) dataset. The CMIP5 models used are 1) CanESM2, 2) CCSM4, 3) ESM2M, 4) ESM_LR, and 5) CGCM3. Note that the scales are different for the CMIP5 and CESM-LE estimations, and the observed values are only shown for CESM-LE in which the time-varying external forcing in the model is consistent with the observed forcing.

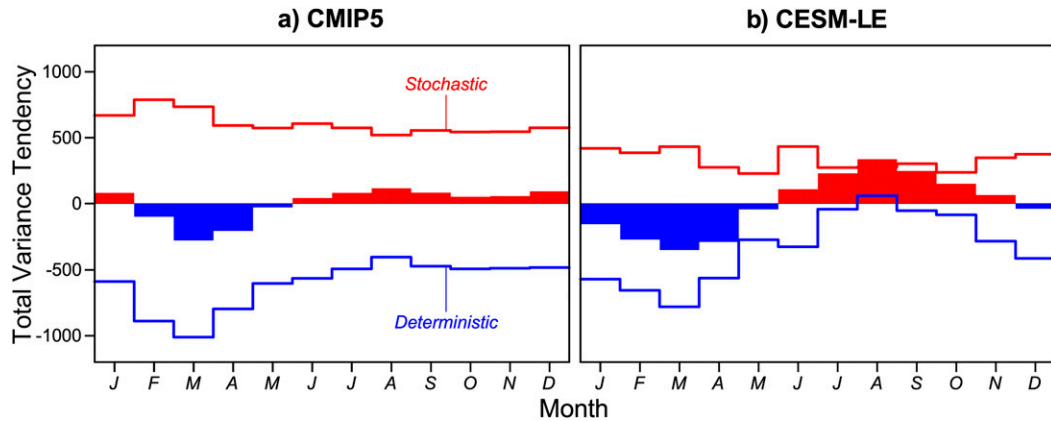


FIG. 8. Annual cycle of total variance tendency ($\text{tr}[d\mathbf{C}/dt]$) derived from (a) the 5-member CMIP5 multimodel ensemble and (b) the 40-member CESM-LE. The total variance tendency in each month is the line with filled bars. The thin blue and red lines show the contributions from the deterministic dynamics $[\mathbf{L}^{\text{CS}}\mathbf{C}(0) + \mathbf{C}(0)(\mathbf{L}^{\text{CS}})^T]$ and stochastic forcing (\mathbf{Q}^{CS}), respectively, to the total variance tendency.

in the early part of the year, whereas, to a lesser extent, those of \mathbf{Q} are more important in the later part. The seasonal variations are generally larger in the CESM-LE than in the CMIP5 simulations, for reasons that are unclear at present.

With regard to the observational CS-LIM, one is faced with the dilemma that while the seasonal variations of \mathbf{L} and \mathbf{Q} are important for improving tropical SST predictions, the limited length of our training dataset (about 50 years) leads to large sampling uncertainties in their estimation. Appendix A shows that accurate observational SST datasets longer than 100 years may be needed to reliably capture the seasonal variations of \mathbf{L} and \mathbf{Q} , although even a 50-yr record may be adequate for some purposes. One way to increase confidence in results obtained from a 50-yr record is to utilize additional information contained in other system variables not explicitly included in an SST-LIM, such as SSH, and exploit their dynamical links to SST. Appendix A shows that doing this does indeed increase confidence in our results. We consider such SST–SSH coupled LIMs next.

5. Coupled SST–SSH LIMs

a. Model construction and characteristics

It is now well recognized that extending an SST-only stationary LIM by explicitly including at least some aspects of ocean dynamics in the LIM's state vector leads to better representations and predictions of seasonal to interannual tropical SST variability (Johnson et al. 2000b; Xue et al. 2000; Hawkins and Sutton 2009; Newman et al. 2011a,b; Ault et al. 2013; Vimont et al. 2014; Capotondi and Sardeshmukh 2015; Chen et al. 2016; Newman and Sardeshmukh 2017). To assess whether this improvement is retained even in a similarly extended CS-LIM, we constructed an SST–SSH coupled CS-LIM using SST and SSH anomalies. The monthly SSTs used were derived from the HadISST dataset (Rayner et al. 2003), and the SSHs were derived from the ECMWF ocean reanalysis system ORAS4 (Balmaseda et al. 2013) over the tropics (25°S – 25°N) for the 1961–2010 period. Monthly anomalies on a 1° latitude \times 1°

longitude grid were determined by removing the grand mean and climatological annual cycle at each grid point. EOF analyses of the SST and SSH anomalies were then performed, and 11 (5) EOFs of SST (SSH) explaining 75.8% (58.3%) of the total variance were retained. As before, all subsequent calculations were performed in EOF space, and then transformed to geographical space for presentation and interpretation.

Considering the pros and cons of the phase-averaged and fixed-phase approaches in CS-LIM construction (see section 2), we combined the two approaches by using three-month running means of the lag-1 and lag-0 covariance matrices in (5) and (6) centered on each calendar month to estimate \mathbf{L}_j^{CS} and \mathbf{Q}_j^{CS} . Such a hybrid approach restricts the phase-averaging effect to a season rather than the whole year, so that the coupled CS-LIM retains most of the phase-dependent features such as ENSO phase locking. (Note that phase-averaging over 12 months in this way would yield the stationary LIM, by design.) The 3-month phase averaging also effectively increases the sample size and thus reduces sampling uncertainties in CS-LIM construction. The phase averaging used here should be distinguished from a prior smoothing of the raw data, or a subsequent smoothing of the estimated \mathbf{L}^{CS} and \mathbf{Q}^{CS} . One can show that for prior smoothing of the raw data, a 12-month smoothing does not converge to the stationary LIM, whereas a subsequent smoothing of \mathbf{L}^{CS} and \mathbf{Q}^{CS} alters system stability and creates imbalances in the FDR.

Retaining 11 SST and 5 SSH EOFs, the SST–SSH coupled linear system is stable (i.e., all the Floquet exponents have negative real parts). However, the stochastic forcing covariance matrices again require minor adjustments to render them positive semidefinite. Similar to the adjustments made in the SST-only LIMs, for each month j we set the negative eigenvalues of \mathbf{Q}_j^{CS} to zero and then rescaled the positive ones to conserve the total variance (i.e., the trace) of the original matrix. Maps of the original and adjusted stochastic forcing variances (not shown) reveal that the adjustments in the SST components of \mathbf{Q}^{CS} are very minor. They are somewhat larger in the SSH components of \mathbf{Q}^{CS} , but only in regions of weak forcing.

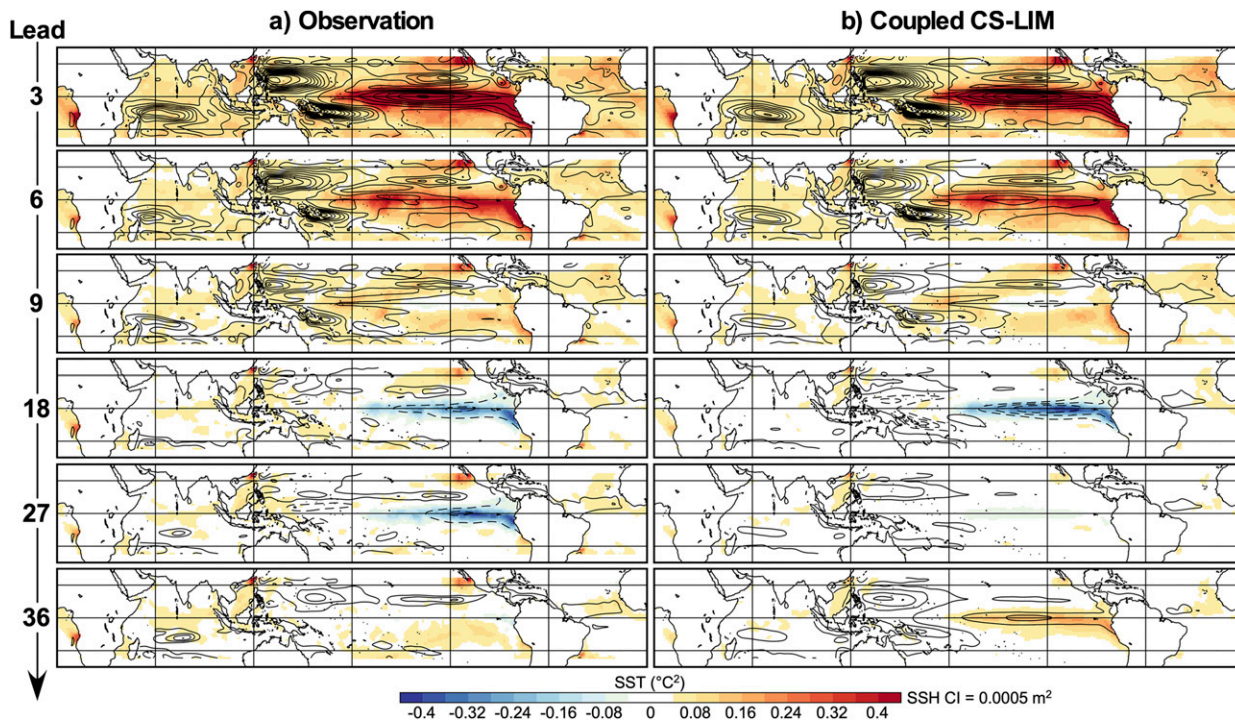


FIG. 9. (a) Observed and (b) coupled CS-LIM derived lag-covariances of SST (color shaded) and SSH (contoured) anomalies at 3-, 6-, 9-, 18-, 27-, and 36-month lags averaged over all months. Note that the observed lag covariances are derived from the full (not EOF-truncated) gridded SST (HadISST; Rayner et al. 2003) and SSH (ORAS4; Balmaseda et al. 2013) datasets.

If the linear approximation is valid, the coupled CS-LIM should be able to reproduce the observed lag-covariance statistics at much longer lags than the 1-month training lag (a.k.a. PS95's “Tau-test”; see also Newman et al. 2011b). The lag-covariance matrices for lags from calendar month j were calculated as $\mathbf{C}(j + \tau) = (\mathbf{G}_{j+\tau-1}^{\text{CS}} \mathbf{G}_{j+\tau-2}^{\text{CS}} \cdots \mathbf{G}_{j+1}^{\text{CS}} \mathbf{G}_j^{\text{CS}}) \mathbf{C}(j)$, bearing in mind that the 12 \mathbf{L}^{CS} do not commute with each other. The observed and “predicted” coupled CS-LIM lag-autocovariances of the SST and SSH anomalies at 3-, 6-, 9-, 18-, 27-, and 36-month lags averaged over all months are compared in Fig. 9. Note that the observed lag-autocovariances were estimated using the untruncated rather than the EOF-truncated anomalies. The coupled CS-LIM reproduces the main features of the observations, as in the stationary case (Newman et al. 2011b). Still, it tends to overestimate anomaly persistence, especially over the eastern equatorial Pacific for lags of up to a year. At the 27 (36)-month lag, the coupled CS-LIM also underestimates (overestimates) the magnitude of the negative (positive) lag-autocovariance over the eastern equatorial Pacific.

To assess annual cycle effects in the coupled CS-LIM, we also constructed and compared a coupled SST–SSH ST-LIM with an SST-only ST-LIM over the same domain (25°S–25°N) and using the same observations. The EOF truncations of these ST-LIMs were chosen to maximize their cross-validated ensemble-mean SST and SSH hindcast skills (presented below in section 5d) among all possible pairs of truncations, constrained by the requirements that 1) \mathbf{L}^{ST} must be stable, and 2) \mathbf{Q}^{ST} must be positive semidefinite. This procedure yielded EOF

truncations of 11 SST and 9 SSH EOFs for the coupled ST-LIM, and 17 EOFs of SST for the SST-only ST-LIM, as optimal choices.

The maximum amplification (MA) curves, depicting the maximum possible growth of SST anomalies over the domain (in the L2 SST-norm) over different time intervals, are shown for the coupled CS-LIM, coupled ST-LIM, and SST-only ST-LIM in Fig. 10. Interestingly, the shapes of the coupled CS-LIM curves (12 thin lines) are quite different from the corresponding stationary case (thick gray and black lines): they have secondary maxima and much larger amplification factors. These larger factors generally imply increased predictability (Newman et al. 2011b) and depend on the initial month of the forecast. While the curves for the ST-LIMs have a single peak at $\tau = 7$, the coupled CS-LIM curves show primary and secondary peaks for initial anomalies that grow to maximum amplitudes in December. For example, optimal February initial conditions can lead to a peak at $\tau = 10$ (primary) and $\tau = 22$ (secondary), and optimal August conditions can peak at $\tau = 4$ (secondary) and $\tau = 16$ (primary). The largest amplification occurs with February initial conditions, at lag $\tau = 10$.

The first two optimal initial structures (OS) in February and maximally evolved final structures (FS) in December (for $\tau = 10$) in the coupled CS-LIM are shown in Fig. 11. The leading OS of SST is very similar to that in the SST-only ST-LIM (e.g., PS95) and also closely resembles the well-known meridional mode pattern over the tropical Pacific (e.g., Chiang and Vimont 2004) that evolves in 10 months into the mature phase

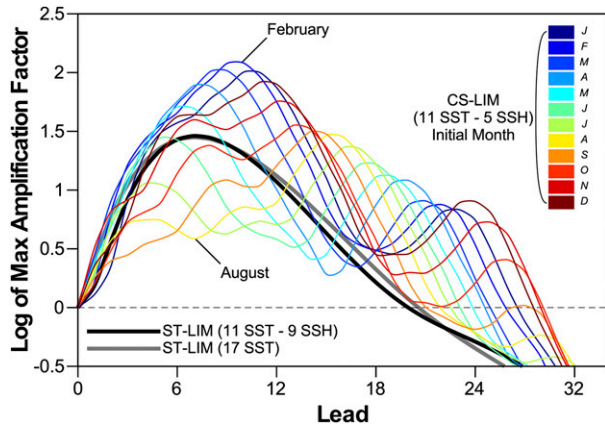


FIG. 10. Logarithm of the maximum anomaly amplification factors (singular values of \mathbf{G}) as a function of forecast lead time, in the coupled CS-LIM for 12 initializing months (12 thin colored lines), coupled ST-LIM (thick black line), and SST-only ST-LIM (thick gray line).

of El Niño in December. The leading OS and FS patterns of SSH are consistent with the delayed oscillator (Schopf and Suarez 1988; Battisti and Hirst 1989) as well as the recharge oscillator (e.g., Jin 1997) mechanisms of ENSO [reviewed recently by Battisti et al. (2018)]. We do not intend here to argue for the greater relevance of one mechanism over the other, but simply to show that our coupled CS-LIM captures the essence of ENSO dynamics. Our empirically derived leading optimal initial structures of SST and SSH are also generally similar to those derived by Thompson and Battisti (2000; see their Fig. 4) from a simple dynamical linear model (Battisti 1988). The subsequent evolution to $\tau = 22$ shows the tropical anomalies flipping sign and evolving into a La Niña-like condition. On the other hand, the second leading OS and FS of SST and SSH merely show a local amplification of the SST anomalies in the central equatorial Pacific and deepening of the thermocline, consistent with Newman et al. (2011a). The evolution from the

second OS to leads to a canonical El Niño pattern at $\tau = 22$, suggesting a possibility for the coupled CS-LIM to capture ENSO development over two years, given the “right” initial conditions (e.g., Kessler 2002).

The optimal structures of the ST-LIM do not depend on season. In such a model, the system has the potential to grow into a mature ENSO pattern in any month as long as the initial SST anomaly pattern projects strongly onto the optimal structures. However, the optimal structures of the coupled CS-LIM leading to a mature ENSO pattern exist only for particular months. The pattern correlations of the OS in February for $\tau = 10$ with the OSs in other months at τ values ranging from 1 to 22 months are shown in Fig. 12. The correlations are high mainly for winter-to-spring initial months, especially for SST, consistent with the mature phase of ENSO occurring primarily in December.

b. ENSO eigenmodes of the coupled cyclostationary system

Table 3 shows the decay times $-1/\text{Re}(\mu)$ and oscillation periods $2\pi/\text{Im}(\mu)$ associated with the Floquet exponents in (9) of the eigenmodes of the coupled SST-SSH CS-LIM. While all the modes are stable, some are “temporarily unstable” in the sense described in section 3a.

To identify the most ENSO-relevant eigenmodes, we estimated the contribution of each mode to the Niño-3.4 SST power spectrum by filtering it out of the data and recomputing the spectrum. Figures 13a and 13b highlight the dominance of eigenmode 4/5 with a period of 38.6 months and a decay time of 19.4 months in this regard. The modal amplitude time series has a correlation of 0.86 with the Niño-3.4 SST anomaly time series, and its power spectrum closely matches the Niño-3.4 SST power spectrum, although its peak is too high. Figure 13c shows that the maximum potential for SST anomaly growth from February into a mature ENSO event 10 months later is almost entirely associated with this single mode. Consistent with this, the mode’s corresponding adjoint SST and SSH patterns (which are optimal initial patterns for exciting the mode) are highly correlated with the optimal initial SST and

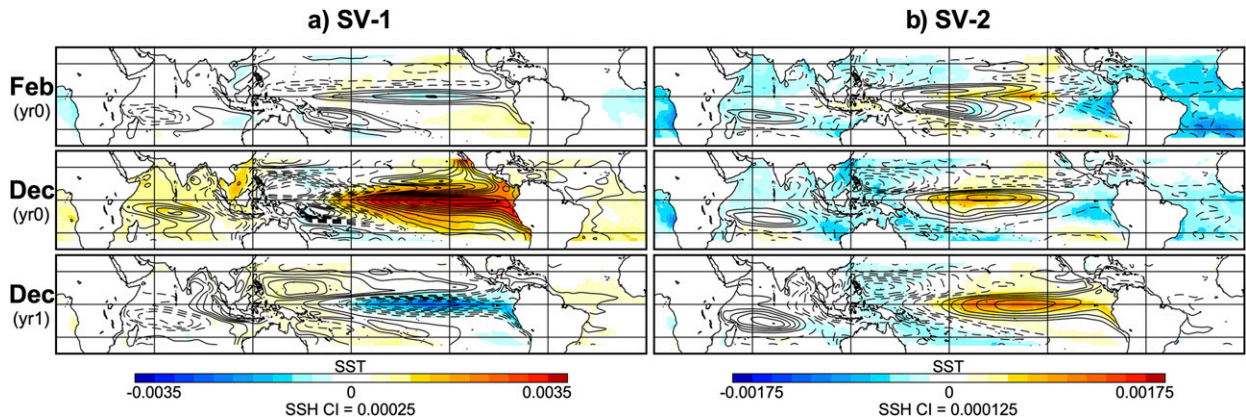


FIG. 11. (a) (top) Leading optimal initial patterns of SST (shaded) and SSH (contoured) anomalies in February for generating maximum rms SST-anomaly magnitude over the entire tropical domain over the next 10 months, and the evolved anomaly patterns in the coupled CS-LIM after (middle) 10 and (bottom) 22 months. The contour interval is arbitrary, but the same in all maps. Negative (zero) values of SSH are dashed (suppressed). (b) As in (a), but for the second leading optimal and evolved anomaly patterns.

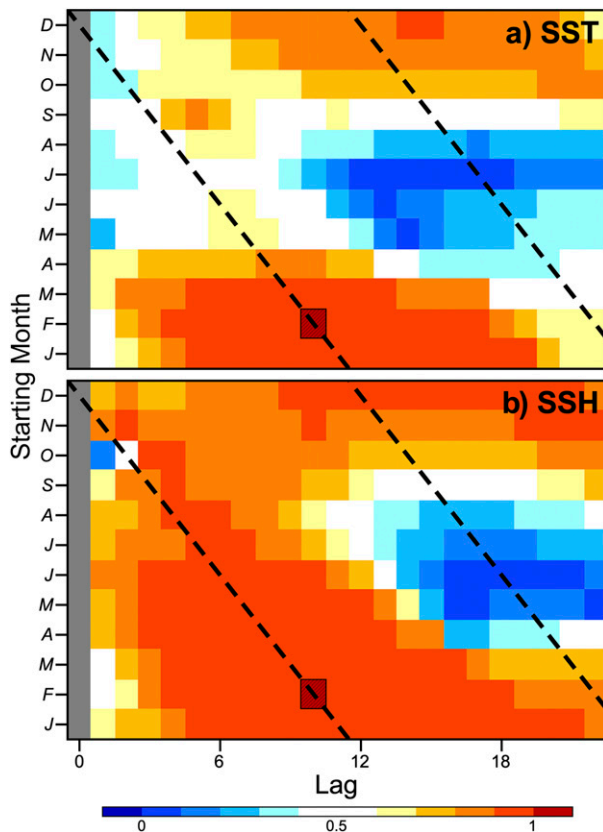


FIG. 12. Pattern correlations of the leading initial optimal structures in February in the coupled CS-LIM shown in Fig. 11 for growth over $\tau = 10$ months with the leading initial optimal structures in all months and for growth over all τ ranging from 1 to 22 months, for (a) SST and (b) SSH. Thick dashed lines indicate the month of December at final evolution.

SSH patterns in February (shown in the top left panel of Fig. 11) at 0.93 and 0.98, respectively. This eigenmode may thus be objectively designated as the dominant ENSO mode of the system. Thompson and Battisti (2000, 2001) also reported a similarly dominant ENSO mode (with a period of 44.4 months

TABLE 3. Empirically derived eigenmode time scales of the coupled CS-LIM. The dominant ENSO mode is indicated by the filled square. Asterisks (*) indicate that these modes are temporarily unstable in some phases of the annual cycle.

Mode	Decay time (months)	Period (months)	ENSO mode
1*	74.4	∞	
2/3*	20.5	519.4	
4/5*	19.4	38.6	■
6	7.6	∞	
7/8	7.2	90.3	
9	5.9	∞	
10/11	4.1	74.1	
12	3.8	∞	
13/14	3.5	52.8	
15/16	2.0	1901.3	

and a decay time of 22.6 months) in their simple dynamical cyclostationary model of ENSO.

The seasonally evolving pattern [i.e. the cyclic eigenvector $\mathbf{p}_{4/5}$ in (9)] of this dominant ENSO mode is shown in Fig. S2 of the online supplemental material. Its evolution from month j to month $j + 1$ can be expressed as $\mathbf{p}_{j+1} = c_j \mathbf{G}_j^{\text{CS}} \mathbf{p}_j$, where the arbitrary amplitude factor $c_j = r_j^{-1} \exp[-\text{Im}(\mu)i]$ can be chosen to normalize the eigenvector in each month, and μ is the Floquet exponent in (9). Specifying the damping rate as $r_j = \|\mathbf{G}_j^{\text{CS}} \mathbf{p}_j\|$ in each month ensures that $\mathbf{p}_{j+1}^\dagger \mathbf{p}_{j+1} = 1$ if $\mathbf{p}_j^\dagger \mathbf{p}_j = 1$ [Here \dagger denotes complex conjugate transpose; see Eq. (37) of von Storch et al. (1995)]. Figure 13d shows the seasonally varying damping rate of the mode. The dominance of the mode may be partly due to its being “temporarily unstable” in the late boreal summer and fall (August–October). A dominant ENSO eigenmode has also been found in some simple dynamical cyclostationary models of ENSO (e.g., Thompson and Battisti 2000).

There is also a dominant ENSO eigenmode of the SST-only CS-LIM discussed in section 3, which is also temporarily unstable in late boreal summer and fall (figure not shown). However, it is more stable overall than the ENSO eigenmode of the SST–SSH coupled CS-LIM, with a decay time scale of 13.0 months instead of 19.4 months, and is not as dominant, with its amplitude time series correlating with the Niño-3.4 SST time series at 0.75 instead of 0.86. This suggests that the explicit inclusion of persistent SSH dynamics in the SST–SSH coupled CS-LIM, rather than cyclostationarity per se, is more important for enhancing the persistence of the ENSO eigenmode. Consistent with this, the ENSO eigenmode of the SST–SSH coupled ST-LIM (not shown) also has a long decay time scale of 20.7 months. However, this mode is also not as dominant as the ENSO mode of the SST–SSH coupled CS-LIM, with its amplitude time series correlating with the Niño-3.4 SST time series at 0.75 instead of 0.86.

The sampling uncertainties of cyclostationary LIMs derived from a limited 50-yr record are relatively greater than those of the corresponding stationary LIMs. We acknowledge that this may compromise accurate estimation of some of the SST–SSH coupled CS-LIM’s parameters (appendix A) and yield an overly dominant ENSO mode. The fact that the “contribution” of this mode to the observed Niño-3.4 SST variance in Fig. 13a exceeds the actual Niño-3.4 variance might also raise a flag in this regard. However, such an exceedance would not necessarily be problematic in itself if the time series of some other modes are negatively correlated with that of the ENSO mode in the nonnormal tropical climate system with non-orthogonal eigenmodes (PS95). Nevertheless, the fact that the CS-LIM’s total power spectrum in Fig. 14a, which accounts for all eigenmodes, also exceeds the observed power in the ENSO frequency band does suggest that the dominance of the ENSO mode in our SST–SSH coupled CS-LIM may be somewhat exaggerated.

The dominance of this single eigenmode in our coupled CS-LIM is also at odds with the combined importance of several eigenmodes highlighted in previous LIM and simple dynamical model studies of ENSO (such as PS95; Newman et al. 2011b; Xie and Jin 2018; and many others) that ignored cyclostationarity and/or explicitly coupled SST and ocean dynamics. PS95

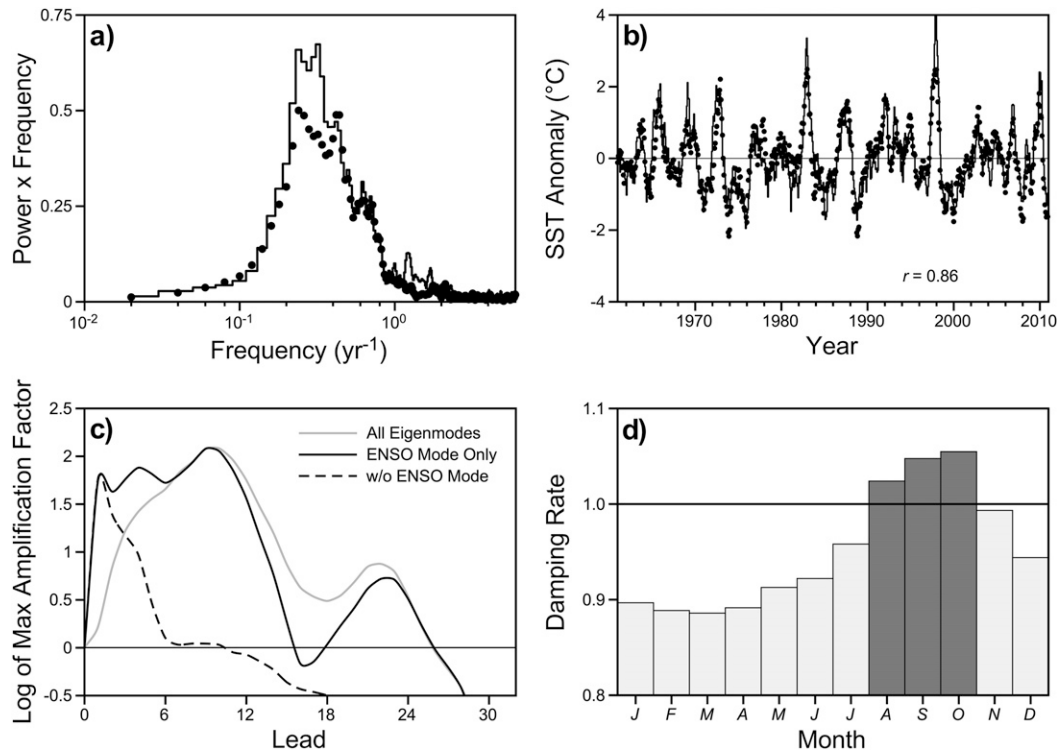


FIG. 13. (a) Power spectra of monthly Niño-3.4 (5°S–5°N; 170°–120°W) SST anomalies reconstructed by using the dominant ENSO eigenmode (black line; mode 4/5 in Table 3). The black dots indicate the spectra from the 50-yr HadISST (1961–2010; Rayner et al. 2003) dataset. (b) As in (a), but for the time series of Niño-3.4 SST anomalies. (c) Logarithm of the maximum amplification factors (MA) as a function of forecast lead time in the coupled CS-LIM initialized in February. The gray, solid black, and dashed lines represent the MA derived from the \mathbf{G}^{CS} with all eigenmodes, the ENSO mode only, and without the ENSO mode, respectively. (d) Seasonal variations of the damping rate of ENSO mode.

ignored both cyclostationarity and coupling in their SST-only ST-LIM, and found that three eigenmodes were important. We constructed an additional SST-only ST-LIM* similar to theirs, and also an SST-only CS-LIM*, from the HadISST dataset for 1961–2010 (where the asterisk indicates that we used 3-month running-mean SSTs instead of monthly SSTs as in all the other LIMs in this study) and obtained nearly identical results to theirs in both cases (not shown). This strongly suggests that cyclostationarity by itself does not lead to the dominance of a single ENSO eigenmode. On the other hand, explicit coupling of the SSTs with ocean dynamics (represented by either SSH or thermocline depth) also does not guarantee the dominance of a single eigenmode, considering that the coupled SST-thermocline depth LIM of Newman et al. (2011b) and the simple coupled dynamical model of Xie and Jin (2018) both had at least two distinct ENSO-related eigenmodes. Xie and Jin (2018) indeed interpreted this as providing a dynamical basis for the “ENSO diversity” that is currently at the forefront of ENSO research (Capotondi et al. 2015; Timmermann et al. 2018). Another confounding result in this context is from Berner et al. (2020), who examined ENSO in the NCAR CESM-LE simulations of the 1920–2100 period. They found a single dominant ENSO mode even in an SST-only ST-LIM derived from the simulated

SSTs, and whose dominance increased with global warming. Reconciling such disparate observational and model results with a single dominant cyclostationary ENSO mode in reality, as suggested but not conclusively established by our coupled SST–SSH CS-LIM, is an interesting issue whose detailed investigation is beyond the scope of this study, but is worth pursuing in the future.

c. Numerical integrations

Having estimated \mathbf{L}^{CS} and \mathbf{Q}^{CS} , we numerically integrated the coupled CS-LIM for 80 100 years from a zero anomaly initial condition $\mathbf{x}(0) = 0$. We then divided the last 80 000 years into 1600 50-yr segments and determined the 1600 sample standard deviations and power spectra of the simulated Niño-3.4 SST anomalies. The median and 5th and 95th percentiles of these quantities are shown in Fig. 14, in a similar format to Fig. 3. The results are very similar to those in Fig. 3, except that the coupled CS-LIM, compared to the SST-only CS-LIM, has a higher 4-yr spectral peak and lower power at frequencies lower than ENSO, and it captures better the seasonal variation of the standard deviation. The ENSO phase locking is similar to that in the SST-only CS-LIM (not shown).

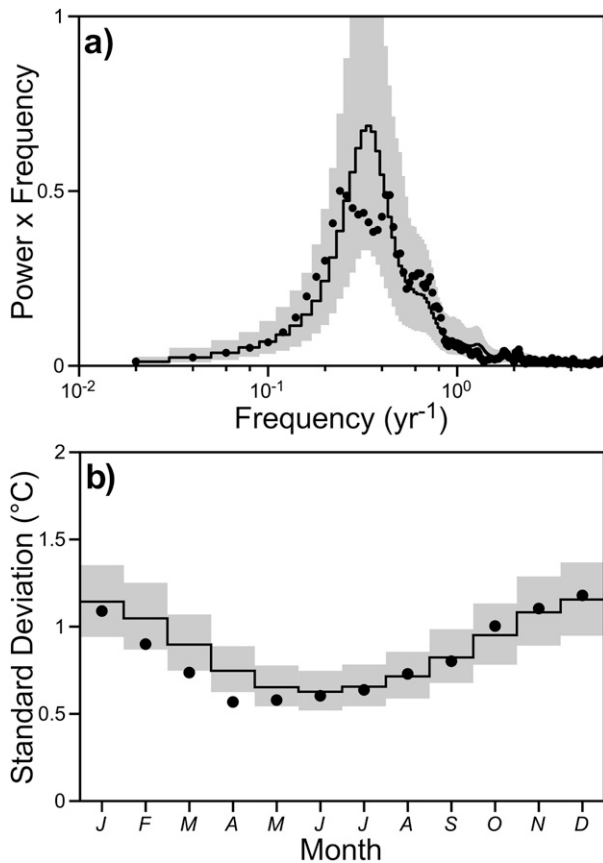


FIG. 14. (top) Power spectra and (bottom) interannual standard deviations of monthly Niño-3.4 (5°S – 5°N ; 170° – 120°W) SST anomalies derived from 80 000-yr coupled CS-LIM. Thick black lines show the median of 1600 values derived from 1600 50-yr segments and gray shading the 90% confidence interval. The black dots show values estimated from the 50-yr HadISST (1961–2010; Rayner et al. 2003) dataset.

d. Hindcast skill in the 1982–2010 period

We compared the hindcast skill of our LIMs with the skill of the nine models used in the NMME forecasting system (Table 4) over the 1982–2010 period. The NMME hindcasts were bias corrected for each model and each ensemble member by subtracting the mean difference between the hindcasts and the observations for each target month and each forecast lead time (e.g., Barnston et al. 2015). The multimodel ensemble mean (hereafter NMME-mean) hindcasts were then determined as the grand mean of all the individual ensemble-member hindcasts of all nine models.

For the coupled CS-LIM (ST-LIM), the cross-validated hindcast skills of SST and SSH were determined by subsampling the data: by successively removing one 5-yr segment (10% of the observations) at a time, re-estimating \mathbf{L}^{CS} (\mathbf{L}^{ST}) and \mathbf{Q}^{CS} (\mathbf{Q}^{ST}) using the remaining 45 independent years, and integrating the stochastic model [Eq. (10)] to generate 114-member ensembles of 1-yr hindcasts, initializing the LIMs in each month of the independent years. This procedure was

repeated for each nonoverlapping 5-yr segment in the 1961–2010 period. For a fair comparison of the LIM and NMME-mean hindcasts, the LIM hindcasts in the 1982–2010 period were subsampled in the same way following the NMME protocol. For the ensemble-mean hindcast, we re-estimated and used an effectively infinite-member ensemble-mean LIM hindcast, $\mathbf{x}_j(t + \tau) = (\mathbf{G}_{j+\tau-1}^{\text{CS}} \mathbf{G}_{j+\tau-2}^{\text{CS}} \cdots \mathbf{G}_{j+1}^{\text{CS}} \mathbf{G}_j^{\text{CS}}) \mathbf{x}_j(t)$ for starting month j and lead τ , instead of averaging the 114-ensemble members. Note that the NMME hindcasts were initialized on (or near) the first day of each month, while the LIM hindcasts were initialized with the monthly mean observations centered on the previous month. We therefore labeled the 1-month lead LIM hindcasts and 0.5-month lead NMME hindcasts as the “month 1” hindcasts, and so on for increasing hindcast lead times.

Local anomaly correlation (AC) skill and root-mean-square error (RMSE)-based skill score (e.g., Barnston et al. 2015) were used to assess the deterministic skill of the ensemble-mean hindcasts, and the relative operating characteristic (ROC; e.g., Mason 1982) curve was used to assess the probabilistic skill (see appendix B for details of the prediction skill measures used). For the deterministic skill, unless stated otherwise, we show the AC skill of the NMME-mean and the coupled CS-LIM. We also show the impacts of the “coupling effect” (coupled ST-LIM skill minus SST-only ST-LIM skill), and the “annual cycle effect” (coupled CS-LIM skill minus coupled ST-LIM skill) on the AC skill of the coupled CS-LIM. Results obtained using the RMSE-based skill score (shown in the supplemental material) are largely consistent with those obtained using the AC skill, and lead to the same conclusions.

1) DETERMINISTIC SKILL

The AC skills at months 6 and 12 of the coupled CS-LIM and NMME-mean are compared in Fig. 15 for SST and in Fig. 16 for SSH (see Figs. S3 and S4 of the supplemental material for the RMSE-based skill scores). For SST, the NMME-mean skill is somewhat higher than the coupled CS-LIM skill (Figs. 15a,b), except in the western Pacific warm pool region. This is because of the well-known erroneous tendency of virtually all climate models to extend ENSO too far west compared to observations, as also highlighted for the NMME models by Newman and Sardeshmukh (2017). The coupled CS-LIM does not have this deficiency.

Figures 15c and 15d show how the skill of the coupled CS-LIM is improved by incorporating both SST–SSH coupling (Fig. 15c) and cyclostationarity (Fig. 15d). The impact of coupling is especially marked over the Indian and Atlantic Oceans, leading to an increase in AC skill of as much as 0.4 at month 12. Over the central equatorial Pacific (Niño-3.4), coupling leads to about a 0.1 increase in skill. Introducing the annual cycle in \mathbf{L} leads to a further ~ 0.1 increase in skill over the central-eastern Pacific region of ENSO development and decay, especially at month 12 (Fig. 15d).

For SSH, the NMME-mean skill is slightly higher than the CS-LIM skill over the Indo-Pacific domain (Figs. 16a,b) except in the Atlantic where it is negative whereas the coupled CS-LIM skill is relatively high (>0.6). The SSH skill of the coupled CS-LIM is also improved by introducing cyclostationarity (Fig. 16c), especially over the western Pacific at month 12. Also

TABLE 4. The NMME models whose hindcast skill was assessed in this study. All models have SST hindcast datasets for the 1982–2010 period. Only four models provide SSH hindcasts for the same period. All data (except SSH from NOAA GFDL) are available on a 1° latitude \times 1° longitude grid at the Earth System Grid archive (<https://www.earthsystemgrid.org>). The SSH data from NOAA GFDL are available at GFDL’s data portal (<https://data1.gfdl.noaa.gov>). We followed the nomenclature of Kirtman et al. (2014) for models and modeling centers; N denotes ensemble size, and L denotes maximum hindcast lead in months.

Model	Modeling center	N	L	SSH	Reference
CCSM4	COLA and RSMAS	10	11.5	✓	Kirtman et al. (2014)
CM2p1-aer04	NOAA GFDL	10	11.5		Delworth et al. (2006)
CM2p5-FLOR-A06	NOAA GFDL	12	11.5		Vecchi et al. (2014)
CM2p5-FLOR-B01	NOAA GFDL	12	11.5	✓	Vecchi et al. (2014)
GMAO-062012	NASA GMAO	12	8.5		Borovikov et al. (2019)
CESM1	NCAR	10	11.5		Lawrence et al. (2012)
CFSv2	NOAA NCEP	28	9.5		Saha et al. (2014)
CanCM3	RMSC	10	11.5	✓	Merryfield et al. (2013)
CanCM4	RMSC	10	11.5	✓	Merryfield et al. (2013)

notable is the skill improvement along the region of the Pacific North Equatorial Countercurrent (NECC), located between approximately 2° and 7°N , at both 6- and 12-month lead times. As reviewed by Battisti et al. (2018), SSH anomalies along the NECC impact ENSO evolution, and are associated with westward propagating oceanic Rossby waves that reflect off the western boundaries as eastward propagating Kelvin waves into the central equatorial Pacific (e.g., Wyrtki 1973). The improved representation of dynamics in the NECC region in the coupled CS-LIM likely also contributes to its improved SST hindcast skill in Fig. 15.

Figure 17 provides a more detailed comparison of the seasonally dependent NMME-mean and coupled CS-LIM AC

skill, as well as the seasonally dependent impacts of SST–SSH coupling and cyclostationarity on the CS-LIM skill (see Fig. S5 of the supplemental material for the RMSE-based skill scores). The results presented are for the Niño-3.4 index (N34; e.g., Trenberth 1997), defined as the area averaged SST anomaly over the central and eastern equatorial Pacific (120° – 170°W , 5°S – 5°N). The seasonal variation of the coupled CS-LIM skill is generally similar to that of the NMME-mean skill. The NMME-mean skill is somewhat higher, especially in the summer verification months. The SST–SSH coupling (Fig. 17c) improves the CS-LIM skill in summer and fall (June–October) at longer than 3-month leads. This improvement, especially in

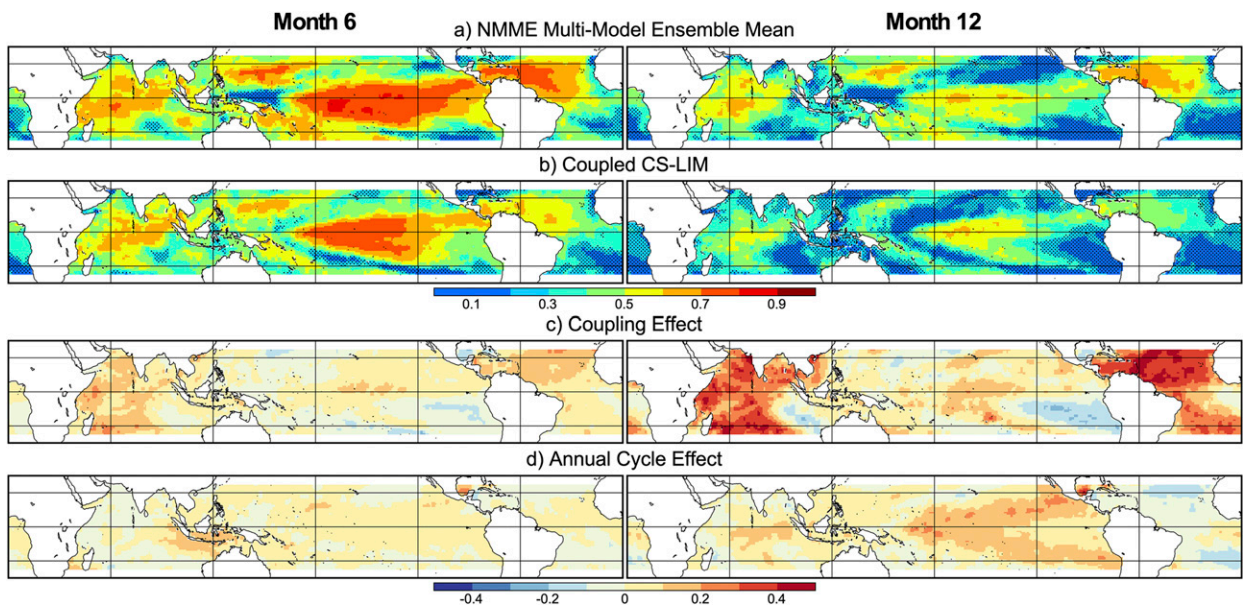


FIG. 15. (a) Local SST anomaly correlation skill of (left) month 6 and (right) month 12 NMME-mean hindcasts in the 1982–2010 period. (b) As in (a), but of the coupled CS-LIM hindcasts. (c) Impact on month 6 and month 12 coupled ST-LIM skill from the SST–SSH coupling (= coupled ST-LIM minus SST-only ST-LIM skill), and (d) impact on month-6 and month-12 coupled CS-LIM skill from including annual cycle effects (= coupled CS-LIM minus coupled ST-LIM skill). The sum of (c) + (d) quantifies the total impact on skill of including both SST–SSH coupling and annual cycle effects in the coupled CS-LIM. Regions in (a) and (b) of insignificant skill at the 95% confidence level are dotted. See Fig. S6 of the supplemental material for the local anomaly correlation skill of the SST-only and coupled ST-LIMs.

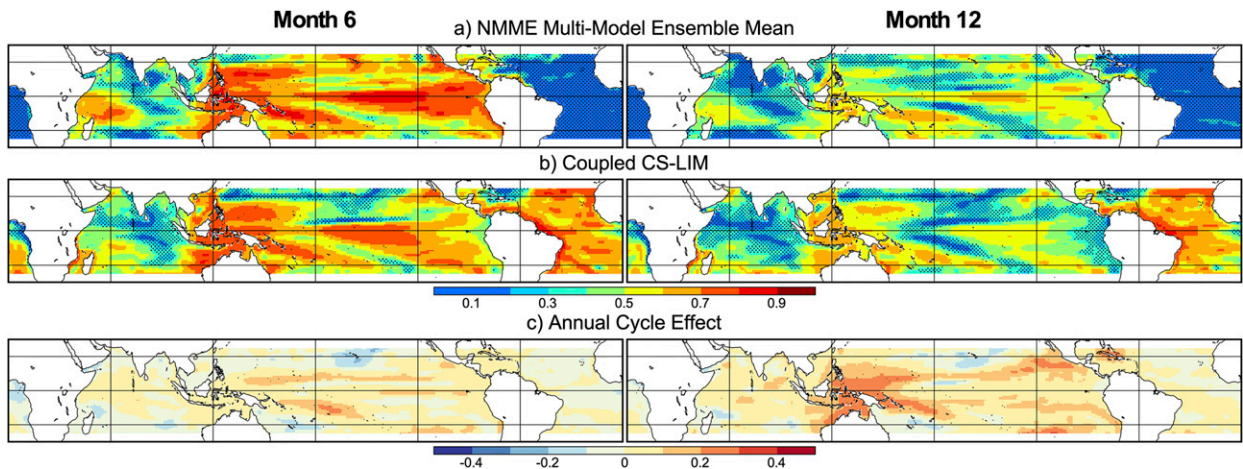


FIG. 16. (a),(b) As in Figs. 15a and 15b, but for the SSH anomaly hindcasts. (c) As in Fig. 15d, but for the impact of including annual cycle effects in the coupled CS-LIM on the SSH hindcast skill (= coupled CS-LIM minus coupled ST-LIM skill). See Fig. S7 of the supplemental material for the local anomaly correlation skill of the coupled ST-LIM.

summer, may be related to oceanic memory in the thermocline (equivalent to the SSH near the equator) of information from the previous winter. On the other hand, the beneficial impact of including the annual cycle in the CS-LIM (Fig. 17d) is greatest in months in which the N34 index typically attains its peak values. The impact on the N34 skill is thus largest in late winter (December–March) at longer than 6-month leads.

2) PROBABILISTIC SKILL

We also intercompared the cross-validated probabilistic skill of the 114-member ensemble SST hindcasts made using the coupled CS-LIM, coupled ST-LIM, and the NMME. The skill was assessed in terms of relative operating characteristic (ROC) curves for three N34 categories (below normal, neutral, and above normal), shown in Fig. 18 for the three prediction systems. At month 6, the ROC curves for the coupled CS-LIM are very similar to those of the NMME in the neutral and above-normal categories, while the NMME is clearly better than the coupled CS-LIM in the below-normal category. This is partly due to the inability of the CS-LIM to capture asymmetric aspects of ENSO (e.g., Okumura and Deser 2010). Recently, Martinez-Villalobos et al. (2019) attributed the asymmetry of El Niño (above normal) and La Niña (below normal) events to the correlated additive and multiplicative (CAM) noise forcing (Sardeshmukh and Penland 2015; Sardeshmukh et al. 2015) of ENSO events, which is ignored in the LIMs developed here.

The beneficial impact of accounting for annual cycle effects in the CS-LIM is clearly reflected in the higher ROC scores of the CS-LIM than the ST-LIM in the below-normal and above-normal categories at all leads longer than 3 months. This impact increases with forecast lead time. Interestingly, the CS-LIM ROC scores for these categories are slightly higher than even the NMME scores at leads longer than 10 months. In the neutral category, all three prediction systems show more or less similar ROC scores at all leads.

e. The 2015–16 El Niño event

The generally similar deterministic and probabilistic skill of the coupled CS-LIM and the NMME is evident even in

forecasts of strong El Niño events. Figure 19 compares ensemble forecasts of the strong 2015–16 event using the two forecast systems, initialized separately in July 2015 and January 2016 during the event’s developing and decaying phases. We stress again that this event was not included in the CS-LIM’s training period. The NMME simulations were initialized on (or near) the first day of July 2015 and January 2016, whereas the LIM hindcasts were initialized with the monthly mean observations centered on the previous months of June 2015 and December 2015.

In both the developing and decaying phases, the median ensemble forecasts of the CS-LIM and NMME are reasonably close to the observed trajectory of the N34 SST anomalies. The NMME spread tends to be larger than the CS-LIM spread, implying greater forecast uncertainty. The CS-LIM forecasts underpredict the observed warm anomalies during the developing phase and overpredict the cold anomalies during the decaying phase of the event. We should note, however, that the CS-LIM was constructed in a truncated EOF space of 11 SST EOFs explaining 75.8% of the SST variance over the entire domain, whereas the observed N34 SST evolution shown is in the untruncated space.

6. Summary and concluding remarks

In this study we developed cyclostationary linear inverse models (CS-LIMs) of monthly tropical SST and SSH anomalies from their observed zero-lag and 1-month lag covariance statistics over the 1961–2010 period. We constructed two versions of these LIMs, using only SST data (“SST-only” LIMs) and using both SST and SSH data (“coupled SST–SSH” LIMs). One of our major goals was to assess the impacts of the seasonal variations of the background state (encapsulated in the model’s deterministic feedback matrix \mathbf{L}) and stochastic forcing (encapsulated in the model’s stochastic forcing covariance matrix \mathbf{Q}) on the seasonally varying statistics and predictability of these variables. We used the SST-only LIMs for this purpose, comparing the variability and prediction skill

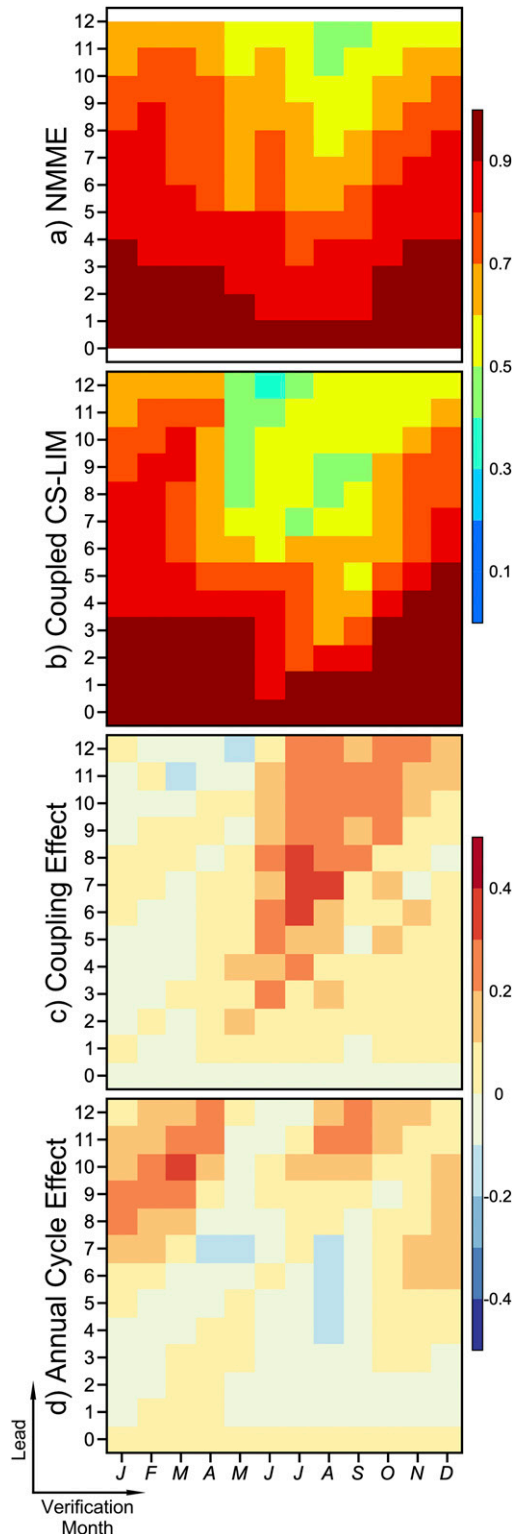


FIG. 17. Niño-3.4 SST index anomaly correlation skill in each verification month of the (a) NMME-mean and (b) coupled CS-LIM hindcasts, (c) the impact on skill of SST–SSH coupling (= coupled ST-LIM minus SST-only ST-LIM skill), and (d) the impact on skill of including annual cycle effects in the coupled CS-

of the SST-only CS-LIM with those of simpler SST-only LIMs in which the seasonal variations of \mathbf{L} and \mathbf{Q} were either completely ignored (ST-LIM) or were ignored in \mathbf{L} but retained in \mathbf{Q} (CSQ-LIM).

The spectra of Niño-3.4 SST anomalies in long integrations of the SST-only CS-LIM, ST-LIM, and CSQ-LIM reproduce the main features of the observed Niño-3.4 spectrum with a peak near 4-yr periods. The CS-LIM represents ENSO phase-locking better than the CSQ-LIM, which suggests that the seasonal variation of \mathbf{L} is more important than that of \mathbf{Q} . Further diagnosis performed after accounting for the implicit contribution of \mathbf{L} in the estimation of \mathbf{Q} suggests an even smaller role of \mathbf{Q} in ENSO phase locking, although it cannot be ignored. The ST-LIM cannot represent phase-locking at all.

SST-only CS-LIMs constructed from CMIP5 model and CESM-LE output suggest the importance of the seasonal variations of \mathbf{L} and \mathbf{Q} also in those models, although the details of the seasonally varying dominant singular values of the deterministic evolution matrix, which are important in determining seasonally varying predictability, differ among the models. In general the seasonal dependence of \mathbf{L} is relatively more important in the first half of the calendar year, whereas that of \mathbf{Q} is, to a lesser extent, more important in the second half. However, the magnitude and phase of the seasonal cycle of variance determined by the cyclic \mathbf{L} and \mathbf{Q} are again model dependent.

We compared the deterministic hindcast skill of the extended SST–SSH coupled CS-LIM with that of the comprehensive NMME prediction system over the 1982–2010 period at forecast leads of a month to a year. The CS-LIM is clearly improved by including both SST–SSH coupling and cyclostationarity, and is more skillful especially in the Pacific NECC region, which is of importance to ENSO. With these improvements, the CS-LIM skill becomes similar to the NMME skill and may indeed be used to benchmark the NMME performance.

We also compared the probabilistic hindcast skill of the coupled CS-LIM with that of the NMME, in terms of ROC scores of N34 anomalies in the above-normal, neutral, and below-normal categories. The scores for the coupled CS-LIM are generally comparable to the NMME scores, and higher than the scores for coupled ST-LIM at all leads beyond 3 months in the above-normal and below-normal categories. This highlights the importance of accounting for annual cycle effects in predicting extreme ENSO events. Still, the CS-LIM scores in the below-normal category are lower than the NMME scores between 3 and 9 months, perhaps due to the inability of the version of the CS-LIM developed here to represent asymmetric aspects of ENSO.

Collectively, our results indicate that a CS-LIM that includes the annual cycles of the background state and stochastic

←

LIM (= coupled CS-LIM minus coupled ST-LIM skill). See Fig. S8 of the supplemental material for the local anomaly correlation skills of the SST-only and coupled ST-LIMs.

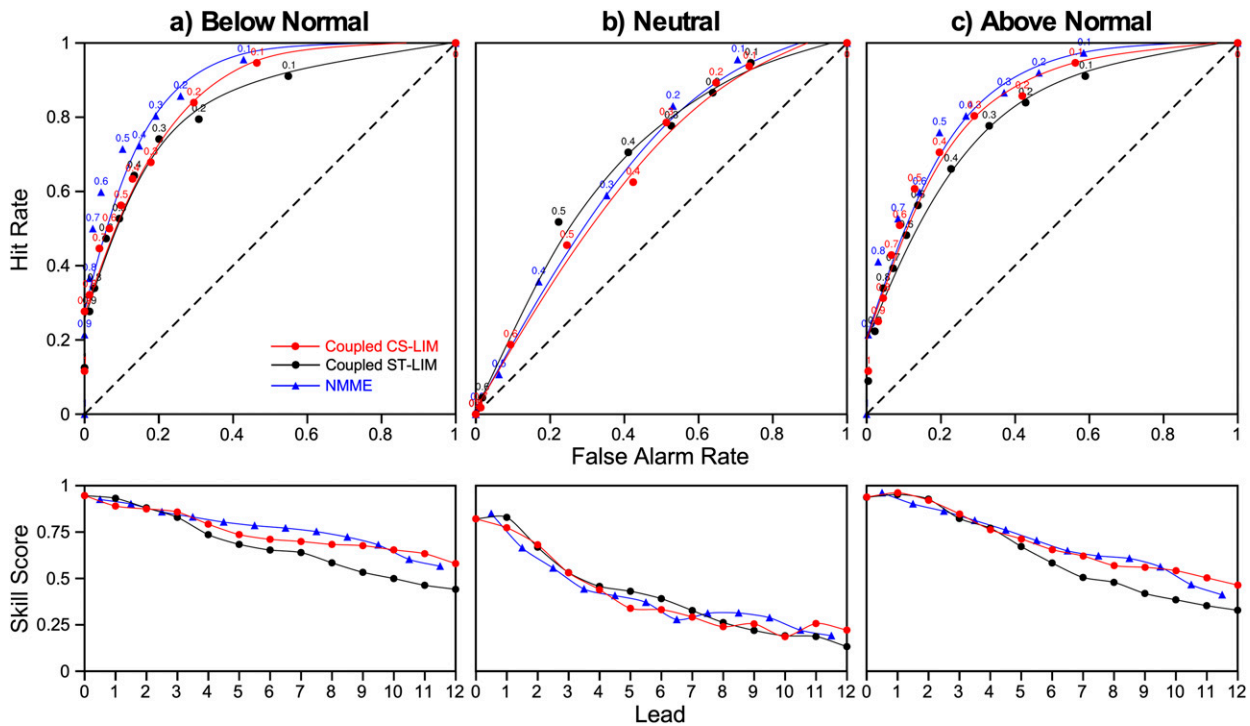


FIG. 18. Probabilistic skill measures of the NMME-mean (blue), coupled CS-LIM (red), and coupled ST-LIM skill (black) hindcasts of the N34 SST index in the 1982–2010 period. (top) ROC curves for the (a) below-normal category N34 index hindcasts of the NMME-mean at 5.5-month lead and the coupled CS-LIM and coupled ST-LIM (black) at 6-month leads, and (bottom) the ROC scores as a function of forecast lead time. (b),(c) As in (a), but for the neutral and above-normal N34 index categories, respectively.

forcing of tropical SST and SSH anomalies is better at representing the seasonal modulation of ENSO-related SST anomalies and their phase locking to the annual cycle than a traditional LIM that ignores such annual cycles. The explicit inclusion of these cycles in the SST–SSH coupled CS-LIM improves the forecast skill of both SST and SSH anomalies through improved SST–SSH coupling. The higher CS-LIM skill results from improving the representation of both ENSO phase-locking and Pacific NECC variations. These improvements result not only from explicitly accounting for the annual cycle of the background state, but also that of the stochastic forcing.

Our results suggest that the coupled CS-LIM may be a useful new tool to benchmark NMME forecasts of tropical SST and SSH anomalies at seasonal-to-interannual scales. They also suggest that the CS-LIM could be a useful component of a global two-tiered seasonal-to-interannual prediction system in which skillful CS-LIM forecasts of tropical SST and SSH would provide evolving boundary conditions to a global model. This is a topic of current research.

Acknowledgments. This work was supported by the Physical Sciences Laboratory of NOAA. SS and MN were also supported by NSF (Grant AGS-1624831) and DOE (Grant 0000238382). We wish to thank two reviewers, Dr. David Battisti and an anonymous reviewer, for their constructive reviews that greatly improved the manuscript. We also wish to thank Dr. Alex Voronovich for his internal review.

APPENDIX A

Sufficiency of Data Record Length to Build a Reliable CS-LIM

We define a “reliable” CS-LIM for this purpose as one whose seasonal variation of maximum monthly singular values, associated with optimal anomaly growth from initial anomalies in each month over the next 7 months (see section 3), is statistically significant. We addressed this issue using an 80 100-yr run of the SST-only CS-LIM (see section 3), dividing the last 80 000 years into 1600 50-yr segments. We constructed a CS-LIM from each segment and determined the maximum singular values of $\mathbf{G}_j^{\text{CS}}(\tau)$ for $\tau = 7$ months and each initialization month j . We then defined their spread (shaded in Fig. A1a) as the difference between the 5th and 95th percentiles of the 1600 values. We repeated this procedure using the 800 total 100-yr segments, 400 total 200-yr segments, and 200 total 400-yr segments of the 80 000-yr run. We also compared, in Fig. A1b, the 12-component “vectors” of the 12 monthly maximum singular values derived from the SST-only CS-LIM in each segment with the 12-component observational vector. Specifically, we computed the correlation coefficient (i.e., the normalized dot product) of the 12-component vectors. Figure A1b shows the probability density function (PDF) of these correlation coefficients, estimated separately for the 1600 total 50-yr segments, 800 total 100-yr segments, 400 total 200-yr segments, and 200 total 400-yr segments of the 80 000-yr run.

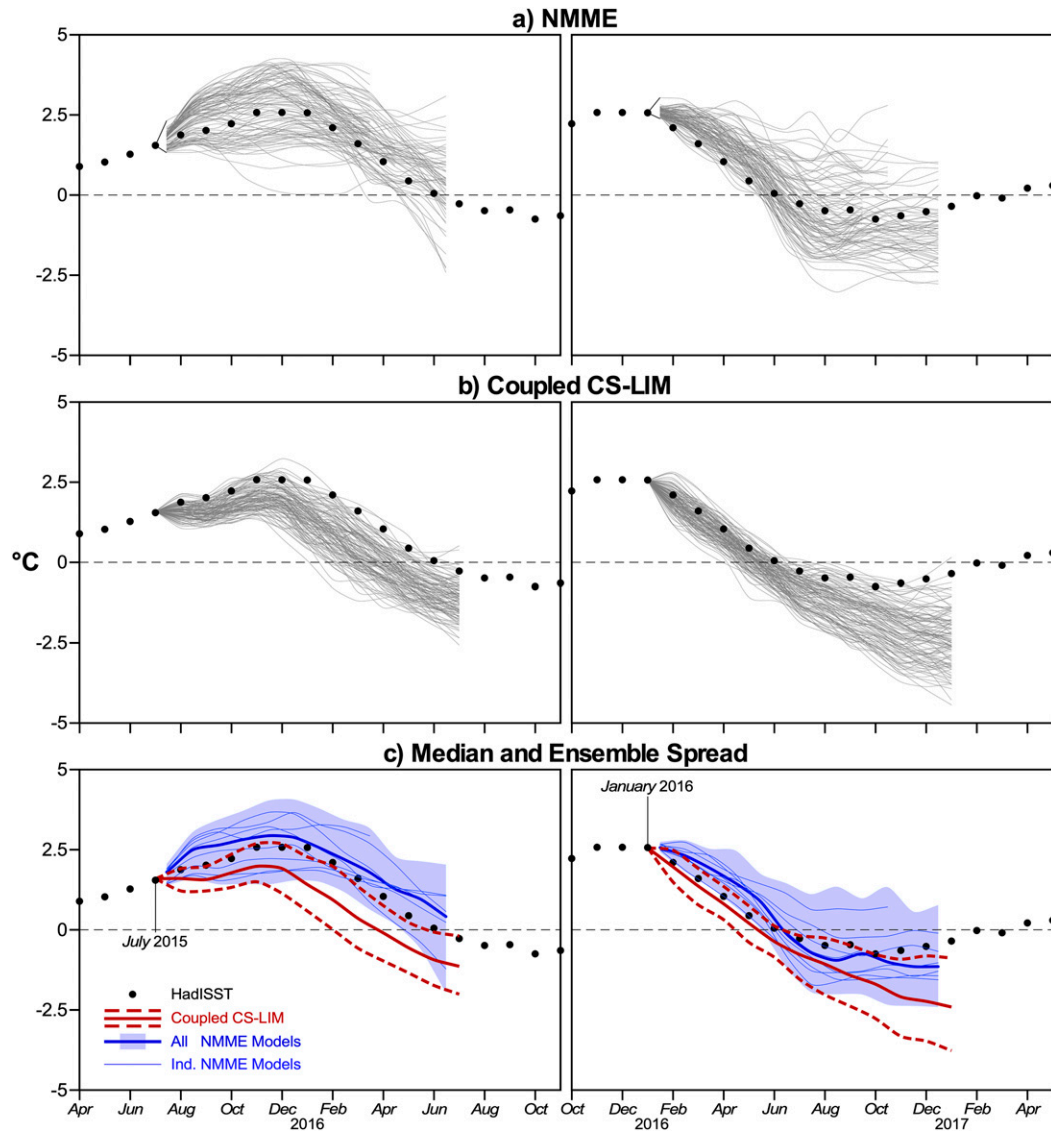


FIG. 19. (left) Ensemble hindcasts (thin gray lines) of the 2015–16 El Niño event using (a) the NMME (maximum of 114 members) and (b) the coupled CS-LIM (114 members), all initialized in July 2015. (c) Thick blue line shows the NMME median hindcast, light blue shading shows the ensemble spread (90% confidence interval). Thin blue lines are median hindcasts of the 9 individual NMME models (see Table 4). Thick red solid line shows the coupled CS-LIM ensemble median hindcast and the thick red dashed lines the ensemble spread (90% confidence interval). Black dots represent observed SST anomalies (HadISST; Rayner et al. 2003). (right) As at left, but for hindcasts initialized in January 2016.

Figure A1a compares the spreads of the maximum monthly singular values derived from the 50-, 100-, 200-, and 400-yr CS-LIM run segments with the values derived from the single 50-yr observational record. For the 50-yr SST-only CS-LIM segments, the spread is large. For longer segments the spread is tighter, and all the observed singular values tend to be located within the spread. We assess the seasonal variation of the observed singular values as significant if the value for any month lies outside the spread for at least one other month. This condition is clearly met for longer than 100-yr segments. It is nearly met even for 50-yr segments in that the observed value

for May lies outside the spread for all other months, and the observed values for all other months lie outside the spread for May. The largest increase in the ratio of the signal (i.e., the seasonal variation of the singular values) to noise (i.e., their spread) occurs when the segment length increases from 50 to 100 years. Figure A1a thus suggests that a 100-yr data record may be sufficient to build a reliable SST-only CS-LIM of the tropical climate system, although even a 50-yr record may be adequate for some purposes.

Figure A1b further supports this conclusion. For a 50-yr data record, the uncertainty of the correlation coefficients (i.e., the

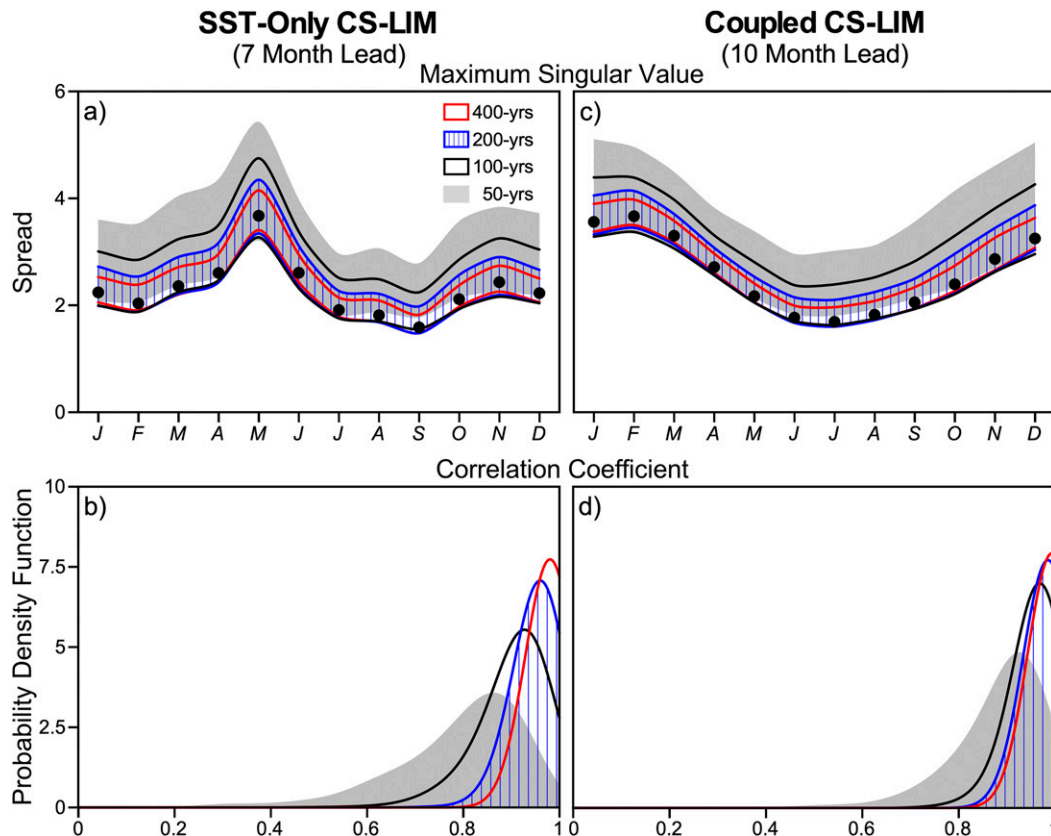


FIG. A1. (a) Estimated uncertainties (spreads) of the maximum singular values of \mathbf{G}_i^{CS} for initialization month j and 7-month lead derived from 50-, 100-, 200-, and 400-yr data records. The spreads were estimated using 1600 total 50-yr segments (gray shading), 800 total 100-yr segments (black lines), 400 total 200-yr segments (blue lines with hatching), and 200 total 400-yr segments (red lines) of an 80 000-yr SST-only CS-LIM run. Spread is defined as the 90% confidence interval. Black dots represent the observed maximum singular values estimated using the 50-yr HadISST dataset. (b) Estimated probability density functions (PDFs) of the correlations (i.e., the normalized dot product) of 12-component “vectors” of the 12 maximum monthly singular values derived from segments of varying lengths from the CS-LIM run with the 12-component vector derived from the HadISST observations. The PDFs were again estimated using 1600 total 50-yr segments (gray shading), 800 total 100-yr segments (black line), 400 total 200-yr segments (blue line with hatching), and 200 total 400-yr segments (red line) of the 80 000-yr run. The figure shows that 100-yr data records may be sufficient to reliably estimate the parameters of a CS-LIM of the tropical climate system, although even a shorter 50-yr record may be adequate for some purposes. See text for more explanation and interpretation. (c),(d) As in (a) and (b), but for an 80 000-yr SST–SSH coupled CS-LIM run and using singular values associated with optimal growth over 10 months. See text for more explanation.

width of their PDF) is large. For longer data records, the uncertainty is smaller. The estimated probabilities of obtaining a correlation with the “truth” (defined here as the 12-component “vector” of observed singular values) of larger than 0.8 (0.9) are 57.9% (19.6%), 91.9% (62.4%), 99.8% (95.8%), and 100% (100%) for 50-, 100-, 200-, and 400-yr records, respectively. Thus, Fig. A1b also suggests that a 100-yr data record may be sufficient to build a reliable SST-only CS-LIM of the tropical climate system, although even a 50-yr record may be adequate for some purposes.

We repeated this significance analysis using an 80 100-yr run of the SST–SSH coupled CS-LIM (see section 5), determining the maximum singular values associated with optimal ENSO amplification over $\tau = 10$ months (see Fig. 10). The results are shown in Figs A1c and A1d, in an identical format to Figs A1a

and A1b, respectively. The maximum singular values in Fig A1c for winter are well separated from those for other seasons for longer than 100-yr records, and reasonably well separated even for 50-yr records. The estimated probabilities (Fig. A1d) of obtaining a correlation with “truth” of larger than 0.8 (0.9) are 84.9% (55.4%), 99.5% (93.6%), 100% (100%), and 100% (100%) for 50-, 100-, 200-, and 400-yr records, respectively. Note that the probabilities are generally larger than in Fig A1b, likely due to the utilization of the extra information contained in SSH as well as the “hybrid” 3-month phase smoothing approach used to construct the SST–SSH coupled CS-LIM. Collectively, Figs. A1c and A1d suggest that a 100-yr data record may also be sufficient to build a reliable coupled CS-LIM of the tropical climate system, although even a 50-yr record may be adequate for some purposes.

APPENDIX B

Prediction Skill Measures

We defined the deterministic prediction skill as the skill of the ensemble-mean forecasts and measured it by the local anomaly correlation (AC) and RMSE-based skill scores. While AC measures the extent to which the prediction is temporally in phase with the observed anomaly, the RMSE-based skill score emphasizes the amplitude of the predicted anomaly in relation to that of the observed anomaly (e.g., Barnston et al. 2015). A positive (negative) RMSE-based skill score indicates that a prediction model performs better (worse) than a model which merely predicts observed climatological values. We estimated the statistical significance of our estimated AC using the two-tailed Student's *t* test, testing a null hypothesis of zero skill. We converted our AC values to Fisher's *z* statistic prior to applying the *t* test and used the effective degrees of freedom estimated from observations during 1982–2010.

We measured the probabilistic prediction skill by the ROC curve (e.g., Mason 1982). A ROC curve indicates the degree of correct probabilistic discrimination (hit rate vs false alarm rate) in a set of forecasts, where discrimination is defined as the ability to distinguish one categorical outcome from another, even if the forecast probabilities have biases or calibration problems. The area underneath the ROC curve is the ROC score. ROC scores above 0.5 reflect positive discrimination skill.

REFERENCES

- Alexander, M. A., I. Bladé, M. Newman, J. R. Lanzante, N.-C. Lau, and J. Scott, 2002: The atmospheric bridge: The influence of ENSO teleconnections on air–sea interaction over the global oceans. *J. Climate*, **15**, 2205–2231, [https://doi.org/10.1175/1520-0442\(2002\)015<2205:TABTIO>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<2205:TABTIO>2.0.CO;2).
- , L. Matrosova, C. Penland, J. D. Scott, and P. Chang, 2008: Forecasting Pacific SSTs: Linear inverse model predictions of the PDO. *J. Climate*, **21**, 385–402, <https://doi.org/10.1175/2007JCLI1849.1>.
- Ault, T. R., C. Deser, M. Newman, and J. Emile-Geay, 2013: Characterizing decadal to centennial variability in the equatorial Pacific during the last millennium. *Geophys. Res. Lett.*, **40**, 3450–3456, <https://doi.org/10.1002/grl.50647>.
- Balmaseda, M. A., K. Mogensen, and A. T. Weaver, 2013: Evaluation of the ECMWF ocean reanalysis system ORAS4. *Quart. J. Roy. Meteor. Soc.*, **139**, 1132–1161, <https://doi.org/10.1002/qj.2063>.
- Barnston, A. G., A. Kumar, L. Goddard, and M. P. Hoerling, 2005: Improving seasonal prediction practices through attribution of climate variability. *Bull. Amer. Meteor. Soc.*, **86**, 59–72, <https://doi.org/10.1175/BAMS-86-1-59>.
- , M. K. Tippett, M. L. L'Heureux, S. Li, and D. G. DeWitt, 2012: Skill of real-time seasonal ENSO model predictions during 2002–11: Is our capability increasing? *Bull. Amer. Meteor. Soc.*, **93**, 631–651, <https://doi.org/10.1175/BAMS-D-11-00111.1>.
- , —, H. M. van den Dool, and D. A. Unger, 2015: Toward an improved multimodel ENSO prediction. *J. Climate*, **54**, 1579–1595, <https://doi.org/10.1175/JAMC-D-14-0188.1>.
- Battisti, D. S., 1988: The dynamics and thermodynamics of a warming event in a coupled tropical atmosphere–ocean model. *J. Atmos. Sci.*, **45**, 2889–2919, [https://doi.org/10.1175/1520-0469\(1988\)045<2889:DATOAW>2.0.CO;2](https://doi.org/10.1175/1520-0469(1988)045<2889:DATOAW>2.0.CO;2).
- , and A. C. Hirst, 1989: Interannual variability in a tropical atmosphere ocean model: Influence of the basic state, ocean geometry and nonlinearity. *J. Atmos. Sci.*, **46**, 1687–1712, [https://doi.org/10.1175/1520-0469\(1989\)046<1687:IVIATA>2.0.CO;2](https://doi.org/10.1175/1520-0469(1989)046<1687:IVIATA>2.0.CO;2).
- , D. J. Vimont, and B. Kirtman, 2018: 100 years of progress in understanding the dynamics of coupled atmosphere–ocean variability. *A Century of Progress in Atmospheric and Related Sciences: Celebrating the American Meteorological Society Centennial, Meteor. Monogr.*, No. 59, Amer. Meteor. Soc., <https://doi.org/10.1175/AMSMONOGRAPHS-D-18-0025.1>.
- Bellenger, H., E. Guilyardi, J. Leloup, M. Lengaigne, and J. Vialard, 2014: ENSO representation in climate models: From CMIP3 to CMIP5. *Climate Dyn.*, **42**, 1999–2018, <https://doi.org/10.1007/s00382-013-1783-z>.
- Berner, J., H. Christensen, and P. D. Sardeshmukh, 2020: Does ENSO regularity increase in a warming climate? *J. Climate*, **33**, 1247–1259, <https://doi.org/10.1175/JCLI-D-19-0545.1>.
- Blumenthal, M. B., 1991: Predictability of a coupled ocean–atmosphere model. *J. Climate*, **4**, 766–784, [https://doi.org/10.1175/1520-0442\(1991\)004<0766:POACOM>2.0.CO;2](https://doi.org/10.1175/1520-0442(1991)004<0766:POACOM>2.0.CO;2).
- Borovikov, A., R. Cullather, R. Kovach, J. Marshak, G. Vernieres, Y. Vikhliav, B. Zhao, and Z. Li, 2019: GEOS-5 seasonal forecast system. *Climate Dyn.*, **53**, 7335–7361, <https://doi.org/10.1007/s00382-017-3835-2>.
- Capotondi, A., and P. D. Sardeshmukh, 2015: Optimal precursors of different types of ENSO events. *Geophys. Res. Lett.*, **42**, 9952–9960, <https://doi.org/10.1002/2015GL066171>.
- , and Coauthors, 2015: Understanding ENSO diversity. *Bull. Amer. Meteor. Soc.*, **96**, 921–938, <https://doi.org/10.1175/BAMS-D-13-00117.1>.
- Chang, P., L. Zhang, R. Saravanan, D. J. Vimont, J. C. H. Chiang, L. Ji, H. Seidel, and M. K. Tippett, 2007: Pacific meridional mode and El Niño–Southern Oscillation. *Geophys. Res. Lett.*, **34**, L16608, <https://doi.org/10.1029/2007GL030302>.
- Chen, C., M. A. Cane, N. Henderson, D. E. Lee, D. Chapman, D. Kondrashov, and M. D. Chekroun, 2016: Diversity, nonlinearity, seasonality, and memory effect in ENSO simulations and prediction using empirical model reduction. *J. Climate*, **29**, 1809–1830, <https://doi.org/10.1175/JCLI-D-15-0372.1>.
- Chiang, J. C. H., and D. J. Vimont, 2004: Analogous Pacific and Atlantic meridional modes of tropical atmosphere–ocean variability. *J. Climate*, **17**, 4143–4158, <https://doi.org/10.1175/JCLI4953.1>.
- Compo, G. P., and P. D. Sardeshmukh, 2010: Removing ENSO-related variations from the climate record. *J. Climate*, **23**, 1957–1978, <https://doi.org/10.1175/2009JCLI2735.1>.
- Delworth, T. L., and Coauthors, 2006: GFDL's CM2 global coupled climate models. Part I: Formulation and simulation characteristics. *J. Climate*, **19**, 643–674, <https://doi.org/10.1175/JCLI3629.1>.
- Dunne, J. P., and Coauthors, 2012: GFDL's ESM2 global coupled climate–carbon Earth system models. Part I: Physical formulation and baseline simulation characteristics. *J. Climate*, **25**, 6646–6665, <https://doi.org/10.1175/JCLI-D-11-00560.1>.
- Farrell, B., 1988: Optimal excitation of neutral Rossby waves. *J. Atmos. Sci.*, **45**, 163–172, [https://doi.org/10.1175/1520-0469\(1988\)045<0163:OEONRW>2.0.CO;2](https://doi.org/10.1175/1520-0469(1988)045<0163:OEONRW>2.0.CO;2).
- Gent, P. R., and Coauthors, 2011: The Community Climate System Model version 4. *J. Climate*, **24**, 4973–4991, <https://doi.org/10.1175/2011JCLI4083.1>.
- Goddard, L., and M. Dille, 2005: El Niño: Catastrophe or opportunity. *J. Climate*, **18**, 651–665, <https://doi.org/10.1175/JCLI-3277.1>.
- , S. J. Mason, S. E. Zebiak, C. F. Ropelewski, R. Basher, and M. A. Cane, 2001: Current approaches to seasonal to

- interannual climate predictions. *Int. J. Climatol.*, **21**, 1111–1152, <https://doi.org/10.1002/joc.636>.
- Grimshaw, R., 1993: *Nonlinear Ordinary Differential Equations*. CRC Press, 328 pp.
- Halpert, M. S., and C. F. Ropelewski, 1992: Surface temperature patterns associated with the Southern Oscillation. *J. Climate*, **5**, 577–593, [https://doi.org/10.1175/1520-0442\(1992\)005<0577:STPAWT>2.0.CO;2](https://doi.org/10.1175/1520-0442(1992)005<0577:STPAWT>2.0.CO;2).
- Ham, Y.-G., and J.-S. Kug, 2014: ENSO phase-locking to the boreal winter in CMIP3 and CMIP5 models. *Climate Dyn.*, **43**, 305–318, <https://doi.org/10.1007/s00382-014-2064-1>.
- Harrison, D. E., and G. A. Vecchi, 1999: On the termination of El Niño. *Geophys. Res. Lett.*, **26**, 1593–1596, <https://doi.org/10.1029/1999GL900316>.
- Hawkins, E., and R. Sutton, 2009: Decadal predictability of the Atlantic Ocean in a coupled GCM: Forecast skill and optimal perturbations using linear inverse modeling. *J. Climate*, **22**, 3960–3978, <https://doi.org/10.1175/2009JCLI2720.1>.
- Huddart, B., A. Subramanian, L. Zanna, and T. Palmer, 2016: Seasonal and decadal forecasts of Atlantic sea surface temperatures using a linear inverse model. *Climate Dyn.*, **49**, 1833–1845, <https://doi.org/10.1007/s00382-016-3375-1>.
- Jin, E. K., and Coauthors, 2008: Current status of ENSO prediction skill in coupled ocean–atmosphere models. *Climate Dyn.*, **31**, 647–664, <https://doi.org/10.1007/s00382-008-0397-3>.
- Jin, F.-F., 1997: An equatorial recharge paradigm for ENSO: I. Conceptual model. *J. Atmos. Sci.*, **54**, 811–829, [https://doi.org/10.1175/1520-0469\(1997\)054<0811:AEORPF>2.0.CO;2](https://doi.org/10.1175/1520-0469(1997)054<0811:AEORPF>2.0.CO;2).
- Jin, Y., Z. Liu, and X. Rong, 2019: General seasonal phase-locking of variance and persistence: Application to tropical Pacific, North Pacific and global ocean. *Climate Dyn.*, **53**, 2825–2842, <https://doi.org/10.1007/s00382-019-04659-7>.
- Johnson, S. D., D. S. Battisti, and E. S. Sarachik, 2000a: Empirically derived Markov models and prediction of tropical Pacific sea surface temperature anomalies. *J. Climate*, **13**, 3–17, [https://doi.org/10.1175/1520-0442\(2000\)013<0003:EDMMAP>2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013<0003:EDMMAP>2.0.CO;2).
- , —, and —, 2000b: Seasonality in an empirically derived Markov model of tropical Pacific sea surface temperature anomalies. *J. Climate*, **13**, 3327–3335, [https://doi.org/10.1175/1520-0442\(2000\)013<3327:SIAEDM>2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013<3327:SIAEDM>2.0.CO;2).
- Kay, J. E., and Coauthors, 2015: The Community Earth System Model (CESM) large ensemble project: A community resource for studying climate change in the presence of internal climate variability. *Bull. Amer. Meteor. Soc.*, **96**, 1333–1349, <https://doi.org/10.1175/BAMS-D-13-00255.1>.
- Kessler, W. S., 2002: Is ENSO a cycle or a series of events? *Geophys. Res. Lett.*, **29**, 2125, <https://doi.org/10.1029/2002GL015924>.
- Kirtman, B. P., and Coauthors, 2014: The North American Multimodel Ensemble: Phase-1 seasonal-to-interannual prediction; Phase-2 toward developing intraseasonal prediction. *Bull. Amer. Meteor. Soc.*, **95**, 585–601, <https://doi.org/10.1175/BAMS-D-12-00050.1>.
- Klein, S. A., B. J. Soden, and N.-C. Lau, 1999: Remote sea surface variations during ENSO: Evidence for a tropical atmospheric bridge. *J. Climate*, **12**, 917–932, [https://doi.org/10.1175/1520-0442\(1999\)012<0917:RSSTVD>2.0.CO;2](https://doi.org/10.1175/1520-0442(1999)012<0917:RSSTVD>2.0.CO;2).
- Kondrashov, D., S. Kravtsov, A. W. Robertson, and M. Ghil, 2005: A hierarchy of data-based ENSO models. *J. Climate*, **18**, 4425–4444, <https://doi.org/10.1175/JCLI3567.1>.
- Lawrence, D. M., K. W. Oleson, M. G. Flanner, C. G. Fletcher, P. J. Lawrence, S. Levis, S. C. Swenson, and G. B. Bonan, 2012: The CCSM4 land simulation, 1850–2005: Assessment of surface climate and new capabilities. *J. Climate*, **25**, 2240–2260, <https://doi.org/10.1175/JCLI-D-11-00103.1>.
- Liu, Z., Y. Jin, and X. Rong, 2019: A theory for seasonal predictability barrier: Threshold, timing, and intensity. *J. Climate*, **32**, 423–443, <https://doi.org/10.1175/JCLI-D-18-0383.1>.
- Marsland, S. J., H. Haak, J. H. Jungclaus, M. Latif, and F. Röske, 2003: The Max-Planck-Institute global ocean/sea ice model with orthogonal curvilinear coordinates. *Ocean Modell.*, **5**, 91–127, [https://doi.org/10.1016/S1463-5003\(02\)00015-X](https://doi.org/10.1016/S1463-5003(02)00015-X).
- Martinez-Villalobos, C., M. Newman, D. J. Vimont, C. Penland, and J. D. Neelin, 2019: Observed El Niño–La Niña asymmetry in a linear model. *Geophys. Res. Lett.*, **46**, 9909–9919, <https://doi.org/10.1029/2019GL082922>.
- Mason, I., 1982: A model for assessment of weather forecasts. *Aust. Meteor. Mag.*, **30**, 291–303.
- Matsumoto, M., and T. Nishimura, 1998: Mersenne twister: A 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Trans. Model. Comput. Simul.*, **8**, 3–30, <https://doi.org/10.1145/272991.272995>.
- Merryfield, W. J., and Coauthors, 2013: The Canadian Seasonal to Interannual Prediction System. Part I: Models and initialization. *Mon. Wea. Rev.*, **141**, 2910–2945, <https://doi.org/10.1175/MWR-D-12-00216.1>.
- Newman, M., 2007: Interannual to decadal predictability of tropical and North Pacific sea surface temperatures. *J. Climate*, **20**, 2333–2356, <https://doi.org/10.1175/JCLI4165.1>.
- , and P. D. Sardeshmukh, 2017: Are we near the predictability limit of tropical Indo-Pacific sea surface temperatures? *Geophys. Res. Lett.*, **44**, 8520–8529, <https://doi.org/10.1002/2017GL074088>.
- , S.-I. Shin, and M. A. Alexander, 2011a: Natural variation in ENSO flavors. *Geophys. Res. Lett.*, **38**, L14705, <https://doi.org/10.1029/2011GL047658>.
- , M. A. Alexander, and J. D. Scott, 2011b: An empirical model of tropical ocean dynamics. *Climate Dyn.*, **37**, 1823–1841, <https://doi.org/10.1007/s00382-011-1034-0>.
- , and Coauthors, 2016: The Pacific decadal oscillation, revisited. *J. Climate*, **29**, 4399–4427, <https://doi.org/10.1175/JCLI-D-15-0508.1>.
- Okumura, Y. M., and C. Deser, 2010: Asymmetry in the duration of El Niño and La Niña. *J. Climate*, **23**, 5826–5843, <https://doi.org/10.1175/2010JCLI3592.1>.
- OrtizBeviá, M. J., 1997: Estimation of the cyclostationary dependence in geophysical data fields. *J. Geophys. Res.*, **102**, 13 473–13 486, <https://doi.org/10.1029/97JD00243>.
- Penland, C., 1996: A stochastic model of IndoPacific sea surface temperature anomalies. *Physica D*, **98**, 534–558, [https://doi.org/10.1016/0167-2789\(96\)00124-8](https://doi.org/10.1016/0167-2789(96)00124-8).
- , and L. Matrosova, 1994: A balance condition for stochastic numerical models with application to the El Niño–Southern Oscillation. *J. Climate*, **7**, 1352–1372, [https://doi.org/10.1175/1520-0442\(1994\)007<1352:ABCFSN>2.0.CO;2](https://doi.org/10.1175/1520-0442(1994)007<1352:ABCFSN>2.0.CO;2).
- , and P. D. Sardeshmukh, 1995: The optimal growth of tropical sea surface temperature anomalies. *J. Climate*, **8**, 1999–2024, [https://doi.org/10.1175/1520-0442\(1995\)008<1999:TOGOTS>2.0.CO;2](https://doi.org/10.1175/1520-0442(1995)008<1999:TOGOTS>2.0.CO;2).
- , and L. Matrosova, 2006: Studies of El Niño and interdecadal variability in tropical sea surface temperatures using a non-normal filter. *J. Climate*, **19**, 5796–5815, <https://doi.org/10.1175/JCLI3951.1>.
- Raddatz, T. J., and Coauthors, 2007: Will the tropical land biosphere dominate the climate–carbon cycle feedback during the twenty-first century? *Climate Dyn.*, **29**, 565–574, <https://doi.org/10.1007/s00382-007-0247-8>.

- Rasmusson, E. M., and T. H. Carpenter, 1982: Variations in tropical sea surface temperature and surface wind fields associated with the Southern Oscillation/El Niño. *Mon. Wea. Rev.*, **110**, 354–384, [https://doi.org/10.1175/1520-0493\(1982\)110<0354:VITSST>2.0.CO;2](https://doi.org/10.1175/1520-0493(1982)110<0354:VITSST>2.0.CO;2).
- Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.*, **108**, 4407, <https://doi.org/10.1029/2002JD002670>.
- Ropelewski, C. F., and M. S. Halpert, 1986: North American precipitation and temperature patterns associated with the El Niño–Southern Oscillation (ENSO). *Mon. Wea. Rev.*, **114**, 2352–2362, [https://doi.org/10.1175/1520-0493\(1986\)114<2352:NAPATP>2.0.CO;2](https://doi.org/10.1175/1520-0493(1986)114<2352:NAPATP>2.0.CO;2).
- , and —, 1987: Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation. *Mon. Wea. Rev.*, **115**, 1606–1626, [https://doi.org/10.1175/1520-0493\(1987\)115<1606:GARSPP>2.0.CO;2](https://doi.org/10.1175/1520-0493(1987)115<1606:GARSPP>2.0.CO;2).
- , and —, 1989: Precipitation patterns associated with the high index phase of the Southern Oscillation. *J. Climate*, **2**, 268–284, [https://doi.org/10.1175/1520-0442\(1989\)002<0268:PPAWTH>2.0.CO;2](https://doi.org/10.1175/1520-0442(1989)002<0268:PPAWTH>2.0.CO;2).
- Saha, S., and Coauthors, 2014: The NCEP Climate Forecast System version 2. *J. Climate*, **27**, 2185–2208, <https://doi.org/10.1175/JCLI-D-12-00823.1>.
- Sardeshmukh, P. D., and C. Penland, 2015: Understanding the distinctively skewed and heavy tailed character of atmospheric and oceanic probability distributions. *Chaos*, **25**, 036410, <https://doi.org/10.1063/1.4914169>.
- , G. P. Compo, and C. Penland, 2015: Need for caution in interpreting extreme weather statistics. *J. Climate*, **28**, 9166–9187, <https://doi.org/10.1175/JCLI-D-15-0020.1>.
- Schopf, P. S., and M. J. Suarez, 1988: Vacillations in a coupled ocean–atmosphere model. *J. Atmos. Sci.*, **45**, 549–566, [https://doi.org/10.1175/1520-0469\(1988\)045<0549:VIACOM>2.0.CO;2](https://doi.org/10.1175/1520-0469(1988)045<0549:VIACOM>2.0.CO;2).
- Shin, S.-I., P. D. Sardeshmukh, and K. Pegion, 2010: Realism of local and remote feedbacks on tropical sea surface temperatures in climate models. *J. Geophys. Res.*, **115**, D21110, <https://doi.org/10.1029/2010JD013927>.
- Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498, <https://doi.org/10.1175/BAMS-D-11-00094.1>.
- Thompson, C. J., and D. S. Battisti, 2000: A linear stochastic dynamical model of ENSO. Part I: Development. *J. Climate*, **13**, 2832–2883, [https://doi.org/10.1175/1520-0442\(2000\)013<2818:ALSDMO>2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013<2818:ALSDMO>2.0.CO;2).
- , and —, 2001: A linear stochastic dynamical model of ENSO. Part II: Analysis. *J. Climate*, **14**, 445–466, [https://doi.org/10.1175/1520-0442\(2001\)014<0445:ALSDMO>2.0.CO;2](https://doi.org/10.1175/1520-0442(2001)014<0445:ALSDMO>2.0.CO;2).
- Timmermann, A., and Coauthors, 2018: El Niño–Southern Oscillation complexity. *Nature*, **559**, 535–545, <https://doi.org/10.1038/s41586-018-0252-6>.
- Trenberth, K. E., 1997: The definition of El Niño. *Bull. Amer. Meteor. Soc.*, **78**, 2771–2777, [https://doi.org/10.1175/1520-0477\(1997\)078<2771:TDOENO>2.0.CO;2](https://doi.org/10.1175/1520-0477(1997)078<2771:TDOENO>2.0.CO;2).
- Tziperman, E., S. E. Zebiak, and M. A. Cane, 1997: Mechanisms of seasonal ENSO interaction. *J. Atmos. Sci.*, **54**, 61–71, [https://doi.org/10.1175/1520-0469\(1997\)054<0061:MOSEI>2.0.CO;2](https://doi.org/10.1175/1520-0469(1997)054<0061:MOSEI>2.0.CO;2).
- , M. A. Cane, and B. Blumenthal, 1998: Locking of El Niño peak time to the end of the calendar year in the delayed oscillator picture of ENSO. *J. Climate*, **11**, 2191–2199, [https://doi.org/10.1175/1520-0442\(1998\)011<2191:LOENOS>2.0.CO;2](https://doi.org/10.1175/1520-0442(1998)011<2191:LOENOS>2.0.CO;2).
- Vecchi, G. A., and Coauthors, 2014: On the seasonal forecasting of regional tropical cyclone activity. *J. Climate*, **27**, 7994–8016, <https://doi.org/10.1175/JCLI-D-14-00158.1>.
- Vimont, D. J., 2012: Analysis of the Atlantic meridional mode using linear inverse modeling: Seasonality and regional influences. *J. Climate*, **25**, 1194–1212, <https://doi.org/10.1175/JCLI-D-11-00012.1>.
- , J. M. Wallace, and D. S. Battisti, 2003: Seasonal footprinting in the Pacific: Implications for ENSO. *J. Climate*, **16**, 2668–2675, [https://doi.org/10.1175/1520-0442\(2003\)016<2668:TSEMIT>2.0.CO;2](https://doi.org/10.1175/1520-0442(2003)016<2668:TSEMIT>2.0.CO;2).
- , M. Alexander, and M. Newman, 2014: Optimal growth of central and east Pacific ENSO events. *Geophys. Res. Lett.*, **41**, 4027–4034, <https://doi.org/10.1002/2014GL059997>.
- von Storch, H., G. Bürger, R. Schnur, and J.-S. von Storch, 1995: Principal oscillation patterns: A review. *J. Climate*, **8**, 377–400, [https://doi.org/10.1175/1520-0442\(1995\)008<0377:POPAR>2.0.CO;2](https://doi.org/10.1175/1520-0442(1995)008<0377:POPAR>2.0.CO;2).
- Wang, X., and J. K. Hale, 2001: On monodromy matrix computation. *Comput. Methods Appl. Mech. Eng.*, **190**, 2263–2275, [https://doi.org/10.1016/S0045-7825\(00\)00243-7](https://doi.org/10.1016/S0045-7825(00)00243-7).
- Wyrtki, K., 1973: Teleconnections in the equatorial Pacific Ocean. *Science*, **180**, 66–68, <https://doi.org/10.1126/science.180.4081.66>.
- Xie, R., and F. F. Jin, 2018: Two leading ENSO modes and El Niño types in the Zebiak–Cane model. *J. Climate*, **31**, 1943–1962, <https://doi.org/10.1175/JCLI-D-17-0469.1>.
- Xue, Y., A. Leetmaa, and M. Ji, 2000: ENSO prediction with Markov models: The impact of sea level. *J. Climate*, **13**, 849–871, [https://doi.org/10.1175/1520-0442\(2000\)013<0849:EPWMMT>2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013<0849:EPWMMT>2.0.CO;2).
- Yang, D., and O. A. Saenko, 2012: Ocean heat transport and its projected change in CanESM2. *J. Climate*, **25**, 8148–8163, <https://doi.org/10.1175/JCLI-D-11-00715.1>.
- Yukimoto, S., and Coauthors, 2011: Meteorological Research Institute–Earth System Model version 1 (MRI–ESM1)—Model description. Tech. Rep. 64, Meteorological Research Institute, Tsukuba, Japan, 88 pp.