

Predictability of 2-year La Niña events in a coupled general circulation model

Pedro N. DiNezio^{1,3} · Clara Deser² · Yuko Okumura² · Alicia Karspeck³

Received: 15 August 2016 / Accepted: 2 February 2017
© Springer-Verlag Berlin Heidelberg 2017

Abstract The predictability of the duration of La Niña is assessed using the Community Earth System Model Version 1 (CESM1), a coupled climate model capable of simulating key features of the El Niño/Southern Oscillation (ENSO) phenomenon, including the multi-year duration of La Niña. Statistical analysis of a 1800 year long control simulation indicates that a strong thermocline discharge or a strong El Niño can lead to La Niña conditions that last 2 years (henceforth termed 2-year LN). This relationship suggest that 2-year LN maybe predictable 18 to 24 months in advance. Perfect model forecasts performed with CESM1 are used to further explore the link between 2-year LN and the “Discharge” and “Peak El Niño” predictors. Ensemble forecasts are initialized on January and July coinciding with ocean states characterized by peak El Niño amplitudes and peak thermocline discharge respectively. Three cases with different magnitudes of these predictors are considered resulting in a total of six ensembles. Each “Peak El Niño” and “Discharge” ensemble forecast consists of 30 or 20 members respectively, generated by adding a infinitesimally small perturbation to the atmospheric initial conditions unique to each member. The forecasts

show that the predictability of 2-year LN, measured by the potential prediction utility (PPU) of the Niño-3.4 SST index during the second year, is related to the magnitude of the initial conditions. Forecasts initialized with strong thermocline discharge or strong peak El Niño amplitude show higher PPU than those with initial conditions of weaker magnitude. Forecasts initialized from states characterized by weaker predictors are less predictable, mainly because the ensemble-mean signal is smaller, and therefore PPU is reduced due to the influence of forecast spread. The error growth of the forecasts, measured by the spread of the Niño -3.4 SST index, is independent of the initial conditions and appears to be driven by wind variability over the south-eastern tropical Pacific and the western equatorial Pacific. Analysis of observational data supports the modeling results, suggesting that the “thermocline discharge” and “Peak El Niño” predictors could also be used to diagnose the likelihood of multi-year La Niña events in nature. These results suggest that CESM1 could provide skillful long-range operational forecasts under specific initial conditions.

Keywords ENSO · El Nino · La Nina · prediction · discharge

School of Ocean and Earth Science and Technology publication number XXXX.

✉ Pedro N. DiNezio
pdn@hawaii.edu

¹ Department of Oceanography, School of Ocean and Earth Science and Technology, University of Hawaii at Manoa, 1000 Pope Road, Honolulu, HI 96822, USA

² National Center for Atmospheric Research, Boulder, CO, USA

³ University of Texas Institute for Geophysics, Austin, TX, USA

1 Introduction

Year to year fluctuations in the sea-surface temperature (SST) of the tropical Pacific Ocean have dramatic impacts on weather and climate throughout the world. These variations are mainly driven by the El Niño/Southern Oscillation (ENSO) phenomenon, which is characterized by an initial warming of the central and eastern Pacific that peaks in boreal winter, known as El Niño, typically followed by anomalous cooling on the subsequent year, known as La

Niña. After this sequence of warm and cold SST anomalies the tropical Pacific returns to near normal conditions. Much of the research on prediction and predictability of ENSO has focused on the *onset* of El Niño (Latif et al. 1998; Kirtman et al. 2001), mainly because predicting the onset of La Niña events is comparatively trivial since they generally follow El Niño.

According to linear ENSO theory, the onset and decay of El Niño and La Niña events are governed by wind-driven changes in the depth of the equatorial thermocline, the boundary separating the warm surface waters from deep and colder waters (Cane and Zebiak 1985; Neelin et al. 1998). Weaker trade winds associated with an El Niño event drive an initial relaxation of east-west tilt of the equatorial thermocline, resulting in a deeper thermocline in the east and a shallower thermocline in the west. This response is followed by a zonal-mean shoaling that peaks a few seasons later once Sverdrup balance is established. This anomalously shallow thermocline initiates the demise of El Niño because it increases the entrainment and upwelling of cold subsurface waters. During this transition phase the upper ocean is characterized by a reduction in heat content, which is typically referred as “discharge” of equatorial heat content. Conversely, stronger trade winds at the peak of La Niña drive a delayed zonal-mean *deepening* of the thermocline characterized by increased upper ocean heat content. During this “recharge” phase the anomalously deeper thermocline diminishes the cooling effect of upwelling initiating the demise of La Niña and transition into El Niño.

Observed El Niño and La Niña events show striking departures from this linear oscillatory behavior. First, most El Niño events last a few seasons and then quickly transition into La Niña. In contrast, one out of two observed La Niña events last 2 years or longer (Kessler 2002; McPhaden and Zhang 2009; Ohba and Ueda 2009; Okumura and Deser 2010; Okumura et al. 2011; Deser et al. 2012). Moreover, very few La Niña events transition directly into El Niño as expected from oscillatory behavior. Instead, the great majority of La Niña events slowly decay, often-times taking several years of near-neutral conditions until the next El Niño event is triggered (Kessler 2002). These observational findings suggest a break down of the oscillatory coupling between SST and thermocline anomalies that are at the heart of linear ENSO theory.

Several mechanisms have been proposed to explain these asymmetries in the dynamics and duration of El Niño and La Niña. These mechanisms emphasize different asymmetries in the coupling between SST, wind, and thermocline depth during El Niño vs. La Niña. One mechanism argues that positive SST anomalies lead to a much stronger wind response and associated delayed thermocline response than negative SST anomalies of the same magnitude. This asymmetric wind response arises from the

non-linear response of atmospheric deep convection to SST anomalies and explains the quick and effective termination of El Niño events, as well as their consistent transition into La Niña. In contrast, negative SST anomalies drive a weaker wind response, which leads to a weaker thermocline recharge and an ineffective termination of La Niña events (e.g. Frauen and Dommenges 2010; Choi et al. 2013; DiNezio and Deser 2014). Spatial asymmetries in the patterns of the wind anomalies of El Niño and La Niña have also been linked to their asymmetric duration (e.g. Ohba and Ueda 2009; Okumura et al. 2011).

Other studies have focused on the role of thermocline depth anomalies on the dynamics of 2-year La Niña. Hu et al. (2013) proposed the following conditions for a second year La Niña: (1) a strong La Niña, which would interrupt the recharge phase by causing the subsurface ocean to remain cold, and (2) absence of downwelling equatorial Kelvin waves after the first year, which would otherwise lead to the demise of La Niña. Because thermocline depth anomalies lead SST anomalies by a few seasons, then these mechanisms could be used to predict the return or the demise of La Niña, and hence its duration. However, the mechanisms proposed by Hu et al. (2013) may not provide predictions with lead times longer than 1 year for the following reasons. The first condition, a strong La Niña during the first year, could provide information up to 12 months in advance, however is not supported by observations which show that some of the strongest events (e.g. the La Niña event of 1988/1989) have lasted only 1 year. The second condition would provide prediction lead times shorter than 1 year, which would be useful, but would not represent an improvement from the predictions provided by current seasonal climate forecast systems.

Longer term predictions have received less attention beyond the study of DiNezio and Deser 2014, (hereafter DD14). Based on analysis of climate model output and ocean reanalysis data, DD14 showed that the state of ENSO the year following the first peak of La Niña is related to the magnitude of the thermocline shoaling 6 months before the onset of La Niña that is, 18 months before the second year peak of La Niña. This relationship indicates that a strong thermocline shoaling or discharge prior to La Niña’s first peak will lead to La Niña conditions during the following year, i.e. a 2-year La Niña. DD14 propose that events with a larger discharge must persist longer because the recharge—the thermocline deepening following the peak of La Niña—is proportionally weaker and therefore ineffective at returning to thermocline to neutral conditions in just 1 year. As discussed above, this asymmetry occurs because of nonlinearities in both the wind response (Choi et al. 2013), but also because vertical thermocline displacements are less effective at influencing SSTs during the recharge phase relative to the discharge phase. During the recharge,

the thermocline deepens and becomes progressively decoupled from the ocean's surface layer becoming less effective at influencing SSTs. As a result, La Niña events initiated by a large discharge have to persist for an additional year in order to drive the integrated recharge heating required to return to neutral conditions.

Although our mechanistic understanding of the dynamics of La Niña events has advanced considerably, it is not yet clear whether skillful prediction of the duration of La Niña is possible. This is a critical issue for the prediction of persistent drought conditions associated with La Niña in areas such as, the Southeastern United States, Texas, and California (Schubert et al. 2004; Seager et al. 2008; Hoerling et al. 2009; Cook et al. 2011; Trenberth et al. 1988; Hoerling and Kumar 2003; Seager 2007; Hoerling et al. 2013). Some of the proposed mechanisms, particularly those focusing on the role of the equatorial thermocline, suggest that there are precursor conditions that could be used to anticipate the evolution of La Niña. One observational study evaluating the evolution of SST anomalies and heat content, a proxy for thermocline depth, found that the transition from La Niña to El Niño may take longer than 1 year (Kessler 2002). The study of DD14 found that the depth of the thermocline six months before the onset of La Niña, i.e. at the peak of the discharge phase, is correlated ($r = 0.47$) with SST anomalies a year and a half later. This relationship could provide an 18-month lead time for predicting a 2-year La Niña event. However, the value of r suggests that there is a large fraction of variability that could be explained by other predictors or be caused by unpredictable atmospheric variability (Penland and Sardeshmukh 1995; McPhaden and Yu 1999; Thompson and Battisti 2001; Fedorov et al. 2003; Chiodi and Harrison 2015).

The aim of our study is to explore the predictability of the *duration* of La Niña using the Community Earth System Model Version 1 (CESM1), a climate model that simulates realistic ENSO characteristics, particularly in regards to the asymmetric duration of El Niño and La Niña. Our approach is idealized in the sense that we use CESM1 to predict its own simulated climate trajectories. The resulting “perfect model” forecasts allow us to explore the upper limits of predictability, that is, the skill of the model under the best possible conditions. Using this “perfect model” approach we explore the link between initial conditions and long-term prediction of the duration of La Niña. Particularly, we focus on identifying initial conditions that could allow skillful prediction of 2-year La Niña more than 1 year in advance. First, we explore the predictors of La Niña duration in a long control simulation performed with CESM1. We then perform a series of “perfect model” forecasts from selected events simulated in the CESM1 control simulation. We use these forecasts to explore how long the forecasts remain skillful, mainly by focusing on 24 and

18-month lead time forecasts initialized at the peak of El Niño and the peak discharge, respectively. Last, we explore the effect of other modes of variability on error growth and associated predictability loss.

2 Data and methods

2.1 Coupled general circulation model

The CESM1 is a numerical model of the global coupled climate system consisting of coupled atmosphere and ocean global general circulation models (GCMs) and comprehensive land, cryosphere, and biogeochemistry models. Our forecasts were performed with all components of the model (ocean, atmosphere, cryosphere and land) were configured at nominal 1° latitude–longitude resolution. Land and ocean biogeochemistry modules were active, but configured so that their associated carbon fluxes do not affect atmospheric CO_2 and hence climate. A brief overview of CESM1 and salient improvements to the atmospheric and oceanic model components is given below. Further details about CESM1 maybe found in Hurrell et al. (2013), Kay et al. (2015), and in a special collection of the Journal of Climate (see <http://journals.ametsoc.org/page/CCSM4/CESM1>).

2.2 Long pre-industrial control simulation

We use a multi-century control simulation performed with CESM1 forced by constant external forcings set at pre-industrial levels. The control simulation is 2200 years in length, of which years 400–2200 are used in our study. The first 400 years exhibit climate drift due to initialization and are not considered in the analysis. The remaining 1800 years used in our analysis exhibit negligible climate drift. More details on the initialization and equilibration of the control simulation may be found in Kay et al. (2015).

2.3 Observational datasets

The following observational datasets are used to evaluate the realism of CESM1's ENSO simulation: SST from the Hadley Centre Sea Ice and SST (HadISST) dataset (Rayner et al. 2003) during 1880–2013 on a $1^\circ \times 1^\circ$ longitude–latitude grid; and (2) upper-ocean temperature from the European Centre for Medium-Range Weather Forecasts (ECMWF) ORA-S4 ocean reanalysis for the period 1958–2013 (Balmaseda et al. 2013). ORA-S4 assimilates temperature and salinity profiles, and along-track altimeter-derived sea-level anomalies on a $1^\circ \times 1^\circ$ longitude–latitude grid with progressively finer latitude resolution (0.3°) in the tropics, and 42 levels in the vertical (18 of which are in the

upper 200 m). The ORA-S4 reanalysis is driven by winds from the 40-year ECMWF Re-Analysis Project (ERA-40) until 1989, ERA-Interim from 1989–2010, and ECMWF NWP analysis thereafter. ORA-S4 also uses observed SST, sea-surface salinity, sea-ice, and global mean sea-level to correct biases in the heat and fresh-water budgets.

2.4 ENSO indices

The Niño-3.4 SST index is computed by averaging the SST anomalies over the central Pacific (120°W – 170°W , 5°S – 5°N) for both the CESM1 and HadISST V1.1. Thermocline depth, Z_{TC} , is computed as the depth of the maximum vertical temperature gradient in the upper 500 m using the three-dimensional output from CESM1 and the ORA-S4 reanalysis. The zonal-mean thermocline index, \bar{Z}'_{TC} , is defined as the Z_{TC} anomalies averaged across the equatorial Pacific (140°W – 80°W , 5°S – 5°N). SST and Z_{TC} anomalies are computed relative to the 1958–2013 climatology, the common period between the ORAS-4 and HadISST datasets.

2.5 Simulated ENSO

CESM1 simulates many features of ENSO observed in nature, particularly the asymmetries in amplitude and duration between El Niño and La Niña. The Niño-3.4 SST index, an index commonly used to capture SST variability associated with ENSO, shows frequent multi-year La Niña events, which generally follow strong El Niño events both in observations (Fig 1a) and the CESM1 control (Fig 1b). We define 2-year La Niña (2-year LN) events as those with a Niño-3.4 SST index less than -0.5 standard deviations for two consecutive DJF seasons. Such persistent La Niña events in the model occur with a 45% frequency similar to observed events according the HadISST dataset which shows 14 out of 29 (44%) 2-year LN during the 1870–2015 period. Note the more recent 1958–2013 period had a total of 8 2-year LN out of 11 events (75%). The frequency of occurrence of *simulated* 2-year LN also exhibits remarkable multi-decadal variations, with values ranging from 6 to 73% over 50-year periods.

The spectrum of the simulated Niño-3.4 index shows a broad peak in the 2–7 years band similar to observations and previous versions of the model (Fig. 1c). On average, the maximum power is approximately 50% larger in CESM1 compared to HadISST, albeit it appears to be more realistic than in CCSM4, which simulated a spectral peak twice as large as observed. The stronger ENSO variability in CESM1 results is also reflected in a standard deviation of the Niño-3.4 SST index of 0.91 K compared to 0.76 K for the detrended HadISST1.1 data during the 1870–2015 period. Observations show centennial changes in ENSO amplitude, with values of 0.70 K for

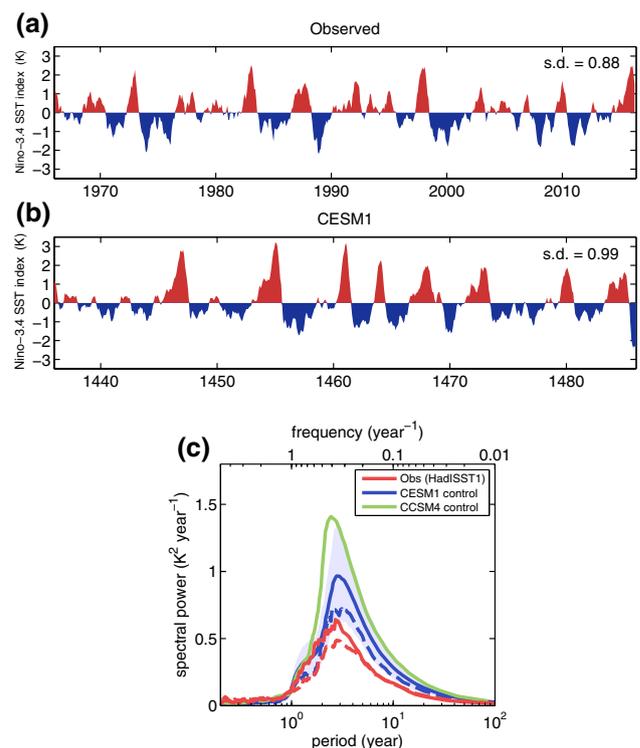


Fig. 1 **a** Observed and **b** simulated Niño-3.4 SST index corresponding to 50 year periods (1964–2015 for observations and 1436–1485 for CESM1). **c** Power spectrum [$\text{K}^2 \text{year}^{-1}$] of the Niño-3.4 SST index from detrended observations (HadISST1.1, *dashed red curve* for 1870–1969 and *solid red curve* for 1916–2015), CESM1 (model years 400–2200, *blue curve*), and CCSM4 (model years 1–1300; *green curve*). CESM1 and CCSM4 data are from control simulations performed with constant pre-industrial forcings. The *light-blue shading* depicts the 95% confidence intervals for the CESM1 power spectrum based on the individual spectra of 100-year segments. *Dashed blue line* shows the power spectra of CESM1’s simulated 100-year period with the smallest spectral peak. Note that to quantitatively compare the variance between the model and observations, the curves must be integrated over the frequency

the 1870–1969 period and of 0.77 K for the 1916–2015 period. CESM1 simulates centennial variations in amplitude ranging from 0.81 to 1.02 K (computed from overlapping 100 year periods, not shown). Although the model’s ENSO amplitude is larger than observed, it represents an improvement relative to CCSM4, which had a mean ENSO amplitude of 1.07 K (Deser et al. 2012). Last, CESM1 has a tendency to simulate El Niño events that start with weak Niño-3.4 SST warming (about 1 K) for a first year, attaining full amplitude 1 year later (Fig. 1b, e.g. El Niño events in 1540s and 1550s decades). While similar 2-year El Niño events have been observed (e.g. during 1931 and 2015), these events are more frequent in the CESM1 control, in which 1 in 3 strong El Niño events are preceded by warming the previous year.

2.6 Simulated 2-year La Niña

We evaluate the dynamics of the simulated 2-year LN by comparing the composite evolution SST and Z_{TC} anomalies against observations. The observed composite is based on 8 2-year events during the 1958–2013 period for which there is available sub-surface temperature data from ORAS4 reanalysis. The simulated composite is based on 156 events. The observed composite agrees with the analysis of (Okumura and Deser 2010) which showed that 2-year La Niña are common throughout the historical record. We identify key times during the evolution of 2-year LN events by their calendar month together with a superscript indicating the lag in months relative to January of the second year peak (Jan⁰), which is our target forecast date. Observed and simulated La Niña have their first peak during boreal winter, denoted Jan⁻¹², 1 year after the preceding El Niño, denoted Jan⁻²⁴ (Fig. 2, shading).

Both observations and CESM1 show that the thermocline shoals across the equatorial Pacific after the peak of El Niño with peak shoaling occurring around July (Jul⁻¹⁸) (Fig. 2, contours). This zonal mean shoaling of the thermocline (Fig. 2, purple contours), is a delayed response to

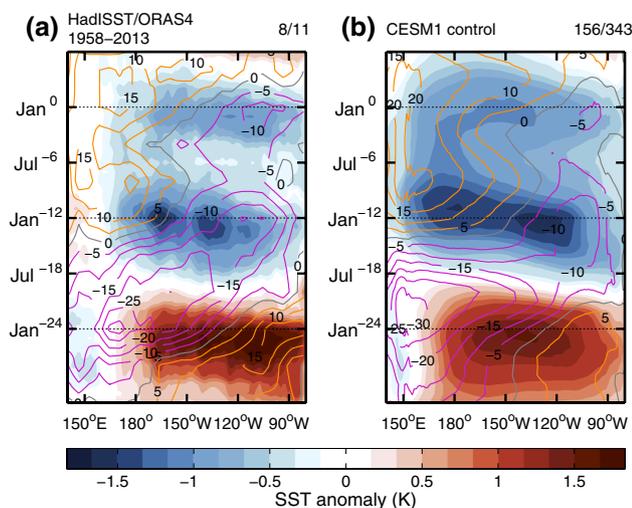


Fig. 2 Spatiotemporal evolution of simulated sea-surface temperature (SST) anomalies (*shading*) and thermocline depth anomalies (Z'_{TC} , contours) for **a** observed and **b** simulated La Niña (LN) events lasting 2 years (2-year LN). Observational data are HadISST1.1 (Rayner et al. 2003) and ORAS4 (Balmaseda et al. 2013) respectively. Simulated data are from the control simulation performed with CESM1 under constant pre-industrial forcings. *Orange* (*purple*) contours show positive (negative) thermocline depth anomalies on 5 m intervals. Equatorial SST and Z'_{TC} anomalies are averaged over the 5° S–5°N band. The observation-based composites are computed using 8 observed 2-year LN out of a total of 11 events from the 1958–2013 period. The model-based composite is computed using 156 simulated 2-year LN out of a total of 343 events from the CESM1 control. Refer to Sect. 2.6 for details on the methodology used to select 2-year LN events

wind anomalies during El Niño, and is typically characterized by a reduction in upper ocean heat content over the equatorial Pacific, hence the moniker “discharge” of equatorial heat content. La Niña onsets after the discharge (Fig. 2, blue shading around Jan⁻¹²), as the shallower thermocline enhances the cooling effect of equatorial upwelling. The termination of La Niña occurs 2 years later (Jul⁺⁶) during the “recharge” phase characterized by an anomalously deeper thermocline across the equatorial Pacific (Fig. 2). The simulated Z'_{TC} anomalies are strikingly realistic in terms of magnitude and timing suggesting that the simulated 2-year LN are governed by similar dynamics than in nature.

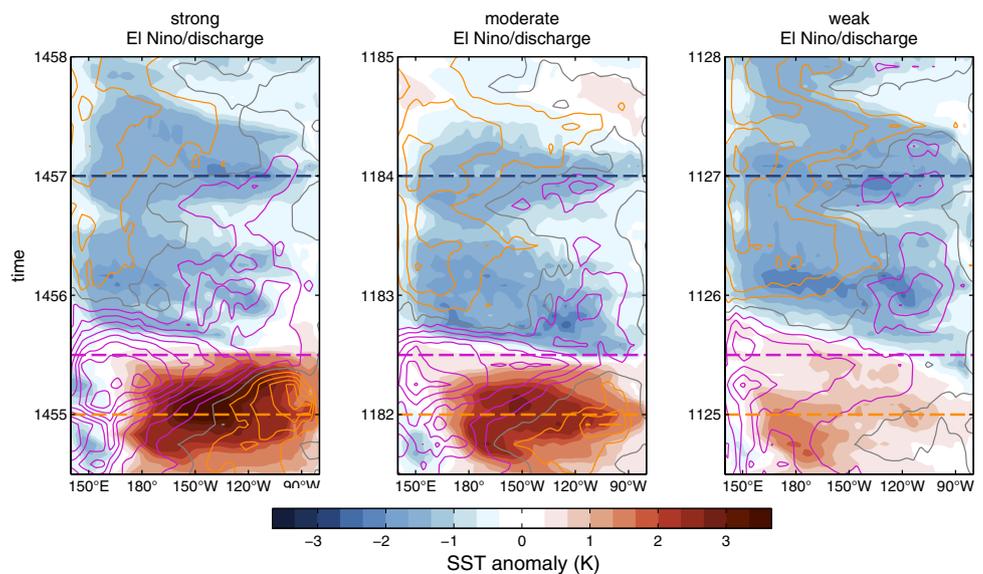
2.7 Perfect model forecasts

We selected three events from the control simulation and ran a series of forecasts for each case study. The forecasts are not affected by ocean initial condition uncertainty, initialization shocks, or model drift because the model is used to predict itself. Therefore the resulting “perfect model” forecasts can be used to assess the *potential* predictability of the duration of La Niña in CESM1.

All three selected events exhibit 2-year LN of similar magnitude, but are preceded by El Niño events of different amplitudes, both in terms of SST and thermocline depth anomalies (Fig. 3). The first case is characterized by a strong El Niño leading to a large shoaling of the thermocline (Fig. 3a). This event shows the largest SST anomalies, with peak values averaging more than 3 K over the Niño-3.4 region. This event also shows the largest thermocline shoaling (discharge) among the three cases. Averaged over the equatorial waveguide (5°S–5°N), the discharge is about –40 m at its peak in May of year 1455. The second case shows weaker positive SST anomalies at the peak of El Niño, as well as less pronounced thermocline shoaling, with peak Niño-3.4 SST anomalies of 2.5 K and discharge of about –30 m (Fig. 3b). The third event shows a much weaker El Niño and thermocline shoaling, with peak Niño-3.4 SST anomalies of less than 1 K and discharge of about –20 m (Fig. 3c). A lagged correlation analysis of subsurface ocean temperatures throughout the tropics shows that there are no significant off-equatorial signals correlated to the evolution of La Niña, other than the surface signature of meridional modes.

The three selected cases will be referred to as “strong”, “moderate” and “weak”, corresponding to the magnitude of the preceding El Niño and thermocline discharge anomalies. Ensemble forecasts were initialized for each case on January 1st of years 1455, 1182, 1125 respectively. We refer to these as “Peak El Niño” forecasts because the initial conditions coincide roughly with the time of maximum positive SST anomalies over the central Pacific. Additional

Fig. 3 As in Fig. 2 but for the events selected for the perfect model forecasts. *Dashed orange* and *magenta lines* show the initial January and July dates for the 24- and 18-month lead forecasts. *Dashed blue line* shows January peak of the second year La Niña, the target forecast time



ensemble forecasts were initialized on July 1st of the same years, which roughly coincides with the peak thermocline shoaling; we refer to these as “discharge” forecasts. We also refer to the “Peak El Niño” and “Discharge” ensemble forecasts as “24-month lead” and “18-month lead” forecasts respectively, corresponding the number of months prior to the second La Niña January peak. Note that the naming convention of both types of forecasts is based on the initialized state and not the predictand.

For each of the “strong”, “moderate” and “weak” events we generated an ensemble of 20 and 30 forecasts for the 18 and 24-month lead cases respectively. The 24-month forecasts consist of 30 members in order to estimate the spread and ensemble-mean with higher accuracy. The size of the ensembles allow us to minimally resolve the extremes of the forecast probability density function, such that, at least one member will fall outside the 95% confidence interval (e.g., for an ensemble of 20 members, 1 member will fall outside the 95% confidence interval). Each member of the forecast ensembles was initialized with the same July 1st or January 1st ocean, sea-ice, and land initial conditions. The forecasts were initialized with a perturbed atmosphere and perfect knowledge of the ocean, sea ice, and land. However, we expect that the ocean initial conditions will be the dominant source of ENSO predictability. To simulate irreducible atmospheric uncertainty in the initial conditions, a unique roundoff level perturbation was made to the full atmospheric temperature, velocity, and moisture fields for each member. Each ensemble member was then integrated forward in time for 4 years.

The spread of the Niño-3.4 SST index grows rapidly from the infinitesimal perturbation, approaching 0.2 K after the first two months. Note that both January- and July-initialized forecasts show similar spread over the first two

months. Thus for the purpose of assessing multi-year predictability in this model, there is apparently no long-term difference between perturbing with errors at the level of rounding, versus at the level of real-world observations. Any differences between these approaches would be effectively overcome after a few months of integration. The month of initialization influences the spread after the second month, showing evidence of seasonal modulation and state dependence. We discuss these features in great detail in Sect. 3.3.

2.8 Forecast plumes

For each forecast member, we computed SST anomalies by removing the monthly mean climatology of the control integration. We then computed each member’s Niño-3.4 SST index as described in Sect. 2.4. We follow the same procedure to compute the forecasted \bar{Z}'_{TC} index. We interpret the ensemble mean Niño-3.4 and \bar{Z}'_{TC} indices as the predictable component of each forecast. The standard deviation of these indices within a given ensemble was used to quantify the spread among the different members, a measure of error growth of the predicted signals.

2.9 Measuring predictability

A number of metrics have been used in the literature to quantify predictability. Most of them are related in some way to the magnitude of the ensemble spread relative to the variance of the system, in other words, the signal-to-noise ratio. However, other metrics can be useful for quantifying the utility of ensemble forecasts (e.g. Kleeman 2002). In this study we use the following two measures to examine predictability: (1) the forecast ensemble spread, a simple

indicator of our confidence in a given forecast, and (2) the “potential prediction utility” (Kleeman and Moore 1999), a measure of the usefulness of the forecast.

The potential prediction utility (PPU) can be expressed as:

$$PPU(t) = \frac{1}{1 + s^2(t)/\langle x \rangle^2(t)}, \quad (1)$$

where s and $\langle x \rangle$ are the standard deviation and ensemble mean Niño-3.4 SST indices, respectively. Like the commonly used correlation coefficient, the PPU also varies from zero to one, with a value of one indicating a perfect forecast. The PPU naturally tends toward unity as the ensemble spread tends to zero. According to this measure, the condition for a useful prediction is not simply low ensemble spread; predictions can be equally useful when the ensemble mean signal is large.

3 Results

3.1 Precursors of 2-year La Niña

We used the long control simulation to explore precursors of 2-year LN. In this section we solely focus on predictors based on equatorial dynamics and thus explore lead-lag relationships between the Niño-3.4 SST index and the zonal mean thermocline depth index, \bar{Z}'_{TC} , which tracks the thermocline discharge and recharge. We selected a total of 536 El Niño events simulated by CESM1 regardless of whether they transition into La Niña or not. In other words, we did not consider “a priori” information, such as, whether: (1) El Niño evolves into La Niña, or (2) La Niña lasts 2 years. The only criterion to select the events is that the Niño-3.4 SST index exceeds 0.5 standard deviations at the peak of El Niño during the December-January-February (DJF) season. The analysis thus includes a subset of El Niño events that start as a weak warming peaking to full amplitude 1 year later as discussed in Sect. 2.5. We discuss the implications of these events below.

The predictand in this analysis is the magnitude of the Niño-3.4 SST index two years after the peak of the selected El Niño events. Niño-3.4⁰ can be positive or negative depending on whether El Niño or La Niña conditions occur 2 years later. 2-year LN events are identified when Niño-3.4⁰ is negative and less than -0.5 standard deviations. We perform a lag-correlation analysis between Niño-3.4⁰ and the magnitude of the zonally averaged Z'_{TC} anomalies and the Niño-3.4 SST index through the evolution of the simulated ENSO events for lead times from 24 to 6 months. All correlation coefficients presented in this section are statistically significant with 99% confidence.

We begin the analysis with a “recharge” predictor, defined as \bar{Z}'_{TC} 6 months after the first peak of La Niña (\bar{Z}'_{TC}^{-6} , where the numeral superscript denotes the time lag in months between the predictor and the predictand). This predictor is based on the notion that La Niña events are terminated by a deepening of the thermocline which reduces the entrainment and upwelling of cold subsurface waters (Neelin et al. 1998). This deepening of the thermocline typically occurs 6 months after the peak of La Niña and is sometimes described as a “recharge” of heat content of the equatorial Pacific. We find that this precursor is significantly correlated ($r = 0.48$) with Niño-3.4⁰ (Fig. 4a). Note that Niño-3.4⁰ can either be positive (El Niño), neutral, or negative (La Niña) as shown by the red, gray, and blue shading in (Fig. 4a). This correlation indicates that La Niña is more likely to end (Niño-3.4⁰ > -0.45 K) when the recharge is high ($\bar{Z}'_{TC}^{-6} > 0$ m), and conversely, more likely to persist (Niño-3.4⁰ < -0.45 K) when the recharge is low or absent ($\bar{Z}'_{TC}^{-6} < 0$ m).

The “recharge” predictor provides information 6 months in advance. However, other predictors earlier in the evolution of ENSO events could provide longer lead time information on the duration of La Niña. For instance, the magnitude of the recharge is related to the preceding La Niña and therefore could provide a 12 month lead-time predictor. This “La Niña” predictor, the magnitude of La Niña during the first year (Niño-3.4⁻¹²), exhibits a relatively weak correlation ($r = 0.12$) with Niño-3.4⁰ (Fig. 4b). The fact that this correlation is positive is consistent with the idea that a stronger La Niña will tend to persist longer than a weaker one (Hu et al. 2013); however this correlation is lower than that for the “recharge” predictor: thus longer-lead time forecasts may not be more skillful. Note that in this and the \bar{Z}'_{TC}^{-6} predictor, we consider a subset of 320 events in which La Niña already peaked for a first year following El Niño.

Predictors could be found earlier in the evolution of ENSO events. For example, DD14 proposed that the magnitude of the thermocline anomalies 6 months *prior* to the onset of La Niña could be related to its magnitude on the second year (i.e. Niño-3.4⁰). This predictor would provide information on the return of La Niña with 18 month lead times. Our analysis of the CESM1 control shows a significant correlation ($r = 0.46$) between a \bar{Z}'_{TC}^{-18} “discharge” predictor and Niño-3.4⁰ (Fig. 4c). For this predictor we focus the analysis on a subset of El Niño events with neutral or negative Niño-3.4 SST index during July after the peak. This additional condition rules out El Niño events that return for a second year. Note that these events happen both in nature and in CESM1 as discussed in Sect. 2.5. Our finding of a significant correlation between \bar{Z}'_{TC}^{-18} and Niño-3.4⁰ suggests that the return of La Niña could be predicted 18 months in advance.

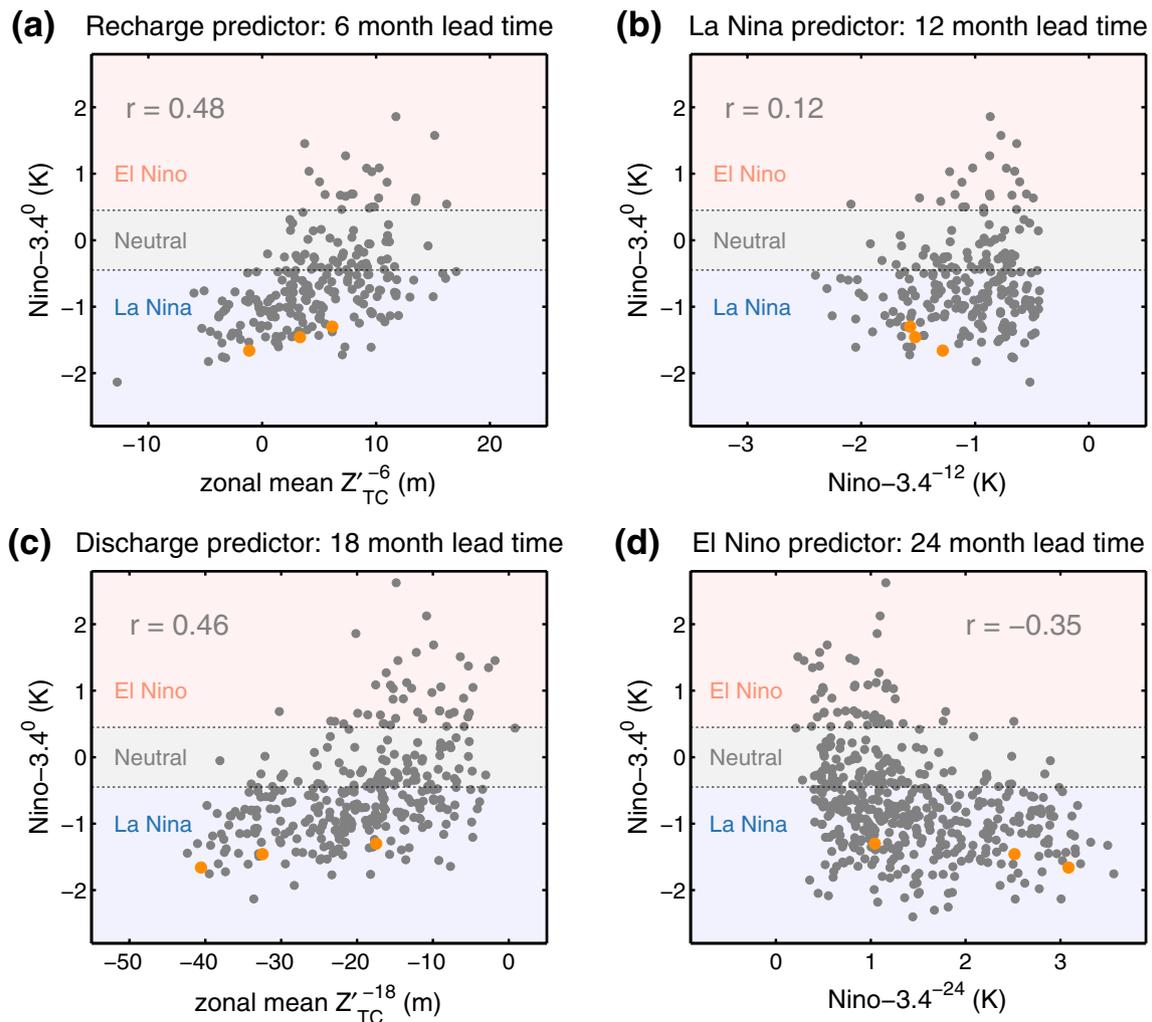


Fig. 4 Scatterplots showing correlations between the Niño-3.4 index on the year following the first peak of La Niña (Niño-3.4⁰, y-axis) and the following predictors (x-axis): **a** heat content recharge 6 months after the first year peak of La Niña ($\bar{Z}'_{TC}{}^{-6}$), **b** Niño-3.4 index during the first year peak of La Niña (Niño-3.4⁻¹²), **c** the heat content discharge during the transition between El Niño to La Niña ($\bar{Z}'_{TC}{}^{-18}$), and **d** the Niño-3.4 index at the peak of the El Niño event preceding La Niña (Niño-3.4⁻²⁴). The heat content discharge and recharge are

measured as the negative and positive zonally averaged thermocline depth anomalies (\bar{Z}'_{TC}) respectively. Niño-3.4⁰ values (y-axis) less than -0.45 K indicate return of La Niña conditions on the second year (*blue shading*), whereas values larger than 0.45 K indicate El Niño conditions during the second year (*red shading*). Correlation coefficient r between the predictor and the predictand are indicated in each panel. *Orange dots* correspond to the events selected for the “perfect model” forecast experiments

Thermocline anomalies associated with the discharge are driven by the preceding El Niño. Thus we also explored correlations between Niño-3.4⁰ and the magnitude of the preceding El Niño given by a Niño-3.4⁻²⁴ predictor (Fig. 4d). The correlation between these quantities is lower ($r = -0.35$) than for the discharge predictor. However, the scatter plot shows that the strongest El Niño events (Niño-3.4⁻²⁴ > 3 K) are virtually always followed by 2-year LN events (Niño-3.4⁰ < -0.45 K). These results suggest that strong El Niño events could lead to skillful 24-month lead time forecasts of 2-year LN. Note that these 18- and 24-month lead time predictions are performed before the initial onset of La Niña. Therefore in these calculations we

considered all El Niño events regardless of whether they are followed by La Niña or not.

The correlations shown by the “Peak El Niño” and “Discharge” predictors with the Niño-3.4 SST index during the 2nd year may not be sufficiently high to warrant skillful predictions of 2-year LN. However, the scatter plots show very few events with neutral or El Niño conditions on the second year when the predictors show large magnitude (Fig. 4c, d). This suggests that the likelihood of 2-year LN increases with the magnitude of the predictors. For instance, 94% of all simulated events with discharge values of $-40 \text{ m} \pm 5 \text{ m}$ become 2-year LN (Table 1). Similarly, 92% of all events with peak El Niño amplitude of 3 ± 0.5

Table 1 Percentage of events in the control simulation showing La Niña, neutral, and El Niño conditions on January of the second year for different values of the El Niño peak and discharge predictors

Control simulation					
Predictor name	Predictor amplitude	% La Niña	% neutral	% El Niño	% of total events
Discharge	−40 m	94	6	0	10
	−30 m	75	24	1	22
	−20 m	60	27	13	35
	−10 m	35	42	23	33
El Niño peak amplitude	3 K	92	6	2	16
	2 K	71	27	2	25
	1 K	44	36	18	59

ENSO events are binned according to their predictor values prior to computing the percentages. The central value of each bin is given in the second column

K become 2-year LN. Events initiated by a strong El Niño are not common (they happen less than 10% of the time in the long control), however, when they happen, their long term evolution will be highly predictable. The percentage of 2-year LN drops to 35 and 44% for the lowest values of the discharge (-10 ± 5 m) and peak El Niño amplitude ($1 \text{ K} \pm 0.5$). For these values of the predictors the frequency of occurrence of neutral, El Niño and La Niña conditions on the second year are about the same magnitude, therefore predictors of this magnitude would lack predictive skill for 2-year LN.

Lastly, we note that the peak amplitude of El Niño and the associated discharge are highly correlated ($r = -0.90$). The high correlation confirms that these predictors are not independent, since ENSO dynamics predicts that stronger El Niño events would drive a stronger discharge. This tight relationship, however, may allow skillful predictions as early in the evolution of an ENSO event as the peak of the El Niño phase. In the next subsections we present ensemble forecasts initialized at the peak of El Niño and at the discharge and discuss and compare their skill.

3.2 Role of initial conditions

3.2.1 Discharge

Next we turn our attention to the perfect model forecasts in order to more quantitatively explore the role of the “Discharge” and “Peak El Niño” predictors. We begin with the 18-month lead forecasts for the three (strong, moderate and weak) discharge predictor cases. All three forecasts start on July with the Niño-3.4 SST index near neutral conditions characteristic of the discharge phase. In all cases, the ensemble mean Niño-3.4 SST index shows La Niña’s first peak around January, 6 months after the forecast start date, as well as a second peak during the subsequent January, 18 months after the forecast start date (Fig. 5, left; blue lines).

The “strong” case predicts the strongest La Niña on the second year, with ensemble mean Niño-3.4 SST index of -1.37 K on January of the second year (model year 1457) (Fig. 5a), while the other cases show similar values of -1.05 and -0.94 K respectively, suggesting weaker 2-year LN relative to the “strong” case (Fig. 5c, e).

In all cases, the ensemble members track the control very closely during the first 6 months and diverge much faster during late summer and early fall of the second year (Fig. 5, left; green lines). We more quantitatively discuss the predictability and error growth of these forecasts by computing the spread (Fig. 7a) and PPU (Fig. 7c), which we discuss in more detail in Sect. 3.3. Despite the increasing forecasts spread suggested by the diverging trajectories of the individual members, the strong case shows only one member above the La Niña threshold during January of the second year (Fig. 5a, green line above the blue background). In contrast, the moderate and weak cases show 4 and 5 members above the La Niña threshold, respectively (Fig. 5c, green lines above the -0.45 K dashed line). We computed the probability of 2-year LN as the percentage of members predicting Niño-3.4 SST index under the -0.5 K threshold during January of the second year (Table 2). We find that the probability of predicting 2-year LN increases for forecasts with a larger initial discharge. This occurs because the value of the predictable signal, captured by the ensemble-mean, is sufficiently negative such as most members remain under the La Niña threshold. Conversely, the probability of predicting neutral or El Niño i.e. an unreliable forecast, is larger when the initial discharge is weaker.

These forecasts start from conditions characterized by an anomalously shallow thermocline, shown by \bar{Z}'_{TC} index values of -35 , -24 , and -12 m , respectively. Note that these values (given by the starting point of the red lines in Fig. 5b, d, f) are slightly less negative than the peak shoaling (given by the minima of \bar{Z}'_{TC} , gray lines in the same figures). The different events of the control

Fig. 5 Evolution of the Niño -3.4 index (left) and the zonal mean thermocline index (\bar{Z}'_{TC}) (right) in forecasts initialized from ocean conditions characterized by discharge conditions (i.e. negative \bar{Z}'_{TC}) ranging from strong (top), moderate (middle), and weak (bottom). The Niño -3.4 SST and \bar{Z}'_{TC} indices are computed as described in Fig. 4 and Sect. 22.4. Positive \bar{Z}'_{TC} indicate a deeper thermocline (recharge) and negative \bar{Z}'_{TC} indicate a shallower thermocline (discharge)

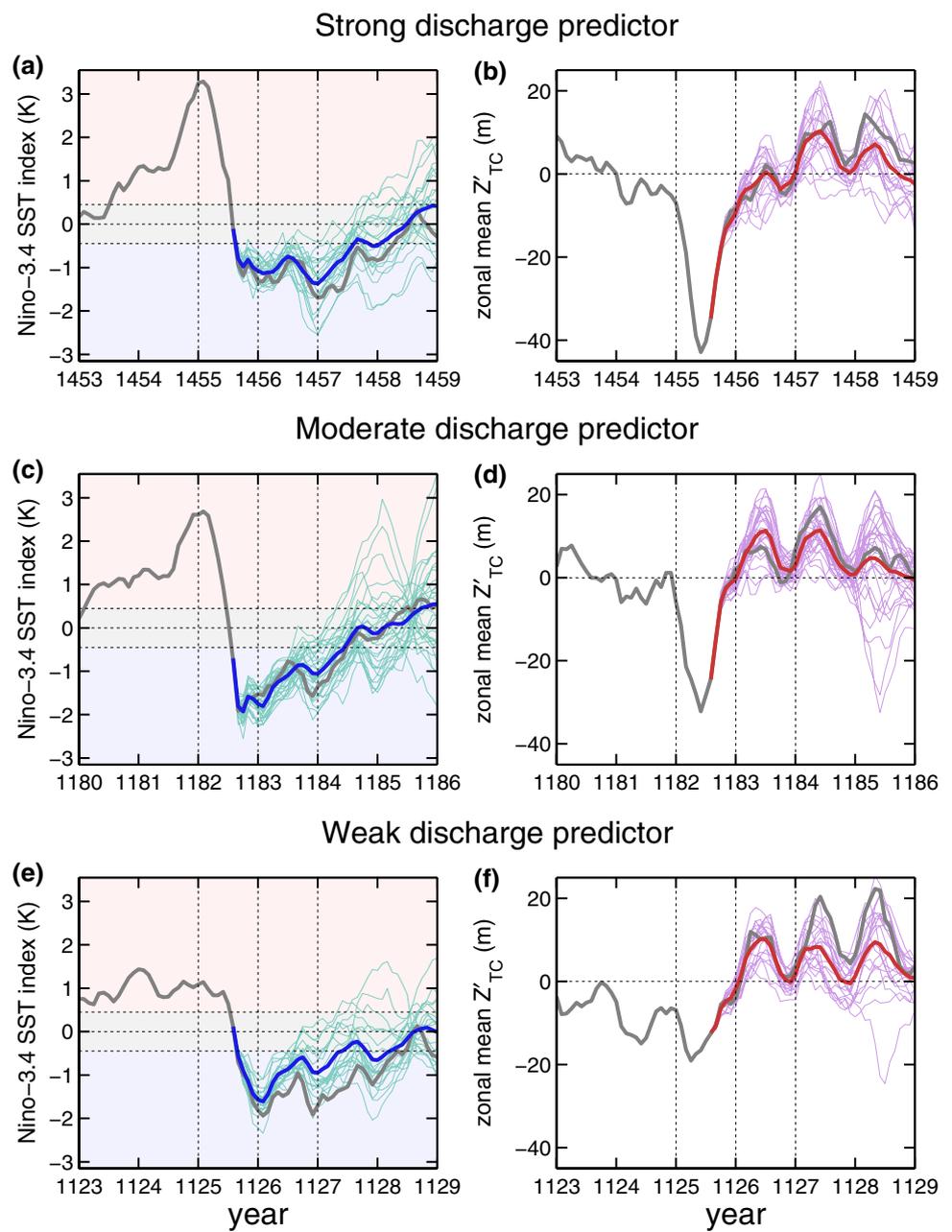


Table 2 Percentage of members in the perfect model forecasts showing La Niña, neutral, and El Niño conditions on January of the second year

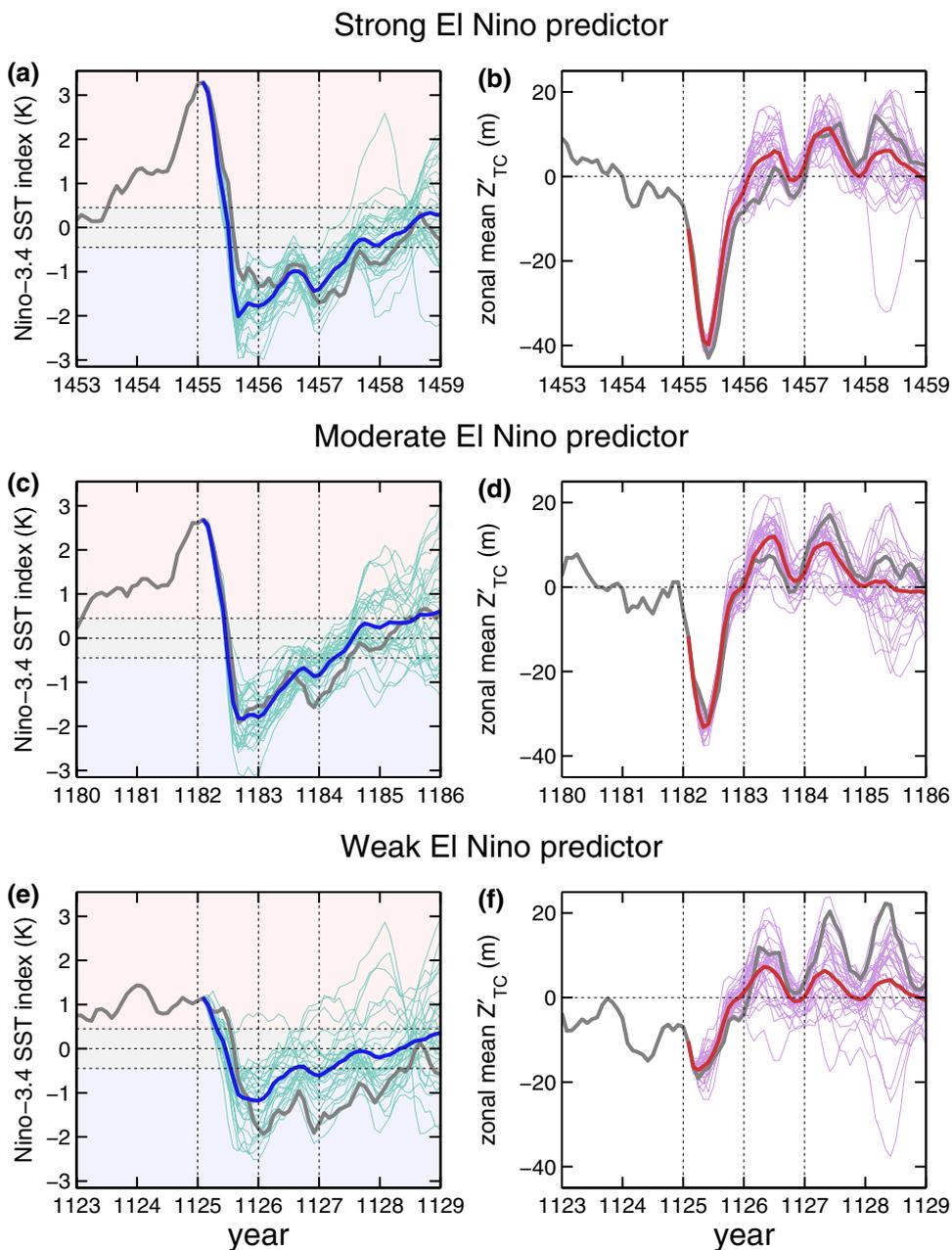
Perfect model forecasts					
Initial state	Forecast name	Predictor amplitude	% La Niña	% neutral	% El Niño
Discharge	Strong	-41 m	90	10	0
	Moderate	-33 m	77	23	0
	Weak	-17 m	57	33	10
El Nino peak	Strong	3.1 K	97	3	0
	Moderate	2.6 K	63	37	0
	Weak	1 K	57	33	10

simulation show discharge peaking from April to July (not shown). We selected initial conditions in July to ensure that all cases are initialized after the peak discharge. The ensemble-mean \bar{Z}'_{TC} shows the thermocline slowly returning to neutral depth throughout the duration of La Niña in all three cases (Fig. 5, right; red lines). The index shows that the thermocline hovers around neutral depths during the first peak of La Niña deepening during the following summer, returning to neutral conditions during the second year peak of La Niña. Note, that the \bar{Z}'_{TC} plumes track the control very closely until the first peak of La Niña (Fig. 5, right; lilac lines) starting to spread out afterwards.

3.2.2 El Niño peak

We continue with the analysis of the 24-month lead forecasts for the three (strong, moderate and weak) “Peak El Niño” cases. All three forecasts start in January when the Niño-3.4 SST index reaches its maximum value at the peak of El Niño. In all three forecasts, the ensemble mean Niño-3.4 SST index shows negative values associated with the peak of La Niña the following January, 12 months after the forecast start date, along with a second peak the following January, 24 months after the forecast start date (Fig. 6, left; blue lines). The “strong” case predicts the strongest second year La Niña, with ensemble mean Niño-3.4 SST index of

Fig. 6 As in Fig. 5, but for “Peak El Niño” forecasts



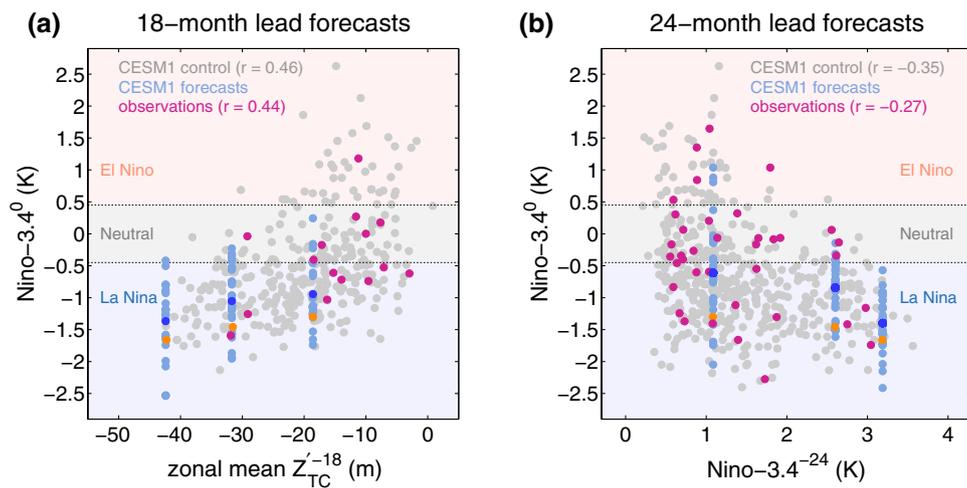


Fig. 7 Scatterplots showing correlations between the Niño-3.4 index on the year following the first peak of La Niña (Niño-3.4⁰, y-axis) and: **a** the heat content discharge and **b** the magnitude of the preceding El Niño (y-axis). Data are from the CESM1 control simulation (gray dots), perfect model forecasts (light blue dots), and ocean reanalysis data (magenta). The heat content discharge is measured as the zonally averaged thermocline depth anomalies (\bar{Z}'_{TC}) during the transition between El Niño to La Niña (\bar{Z}'_{TC}^{-18}). Niño-3.4⁰ values (y-axis)

−1.44 K in January of the second year (model year 1457) (Fig. 6a), while the other cases show values of −0.87 and −0.63 K respectively, suggesting weaker 2-year LN relative to the “strong” case (Fig. 6c, e). During the first year of the forecasts, the Niño-3.4 SST plumes do not track the control as closely as in the “Discharge” forecasts (Fig. 6, left; green lines). Despite this larger initial error growth, 29 out of 30 members of the “strong” case and 26 out of 30 of the “moderate” case predict the return of La Niña with a 24 month lead time.

We find that the probability of predicting 2-year LN increases for forecasts initialized from strong El Niño, which is explained by the larger predictable signals given by the ensemble mean. However, we also find that the likelihood of a 2-year LN forecasts are comparable to the corresponding “Discharge” cases despite the longer lead time (Table 2). The temporal evolution of the spread explains why these forecasts remain as skillful. We find that the spread of these forecasts increases rapidly during the first 6 months, but then decreases reaching values similar to the “Discharge” forecasts by January of the second year (Fig. 7b) and PPU (Fig. 7d). This, explains why these forecasts appear to be as reliable as the corresponding “Discharge” cases.

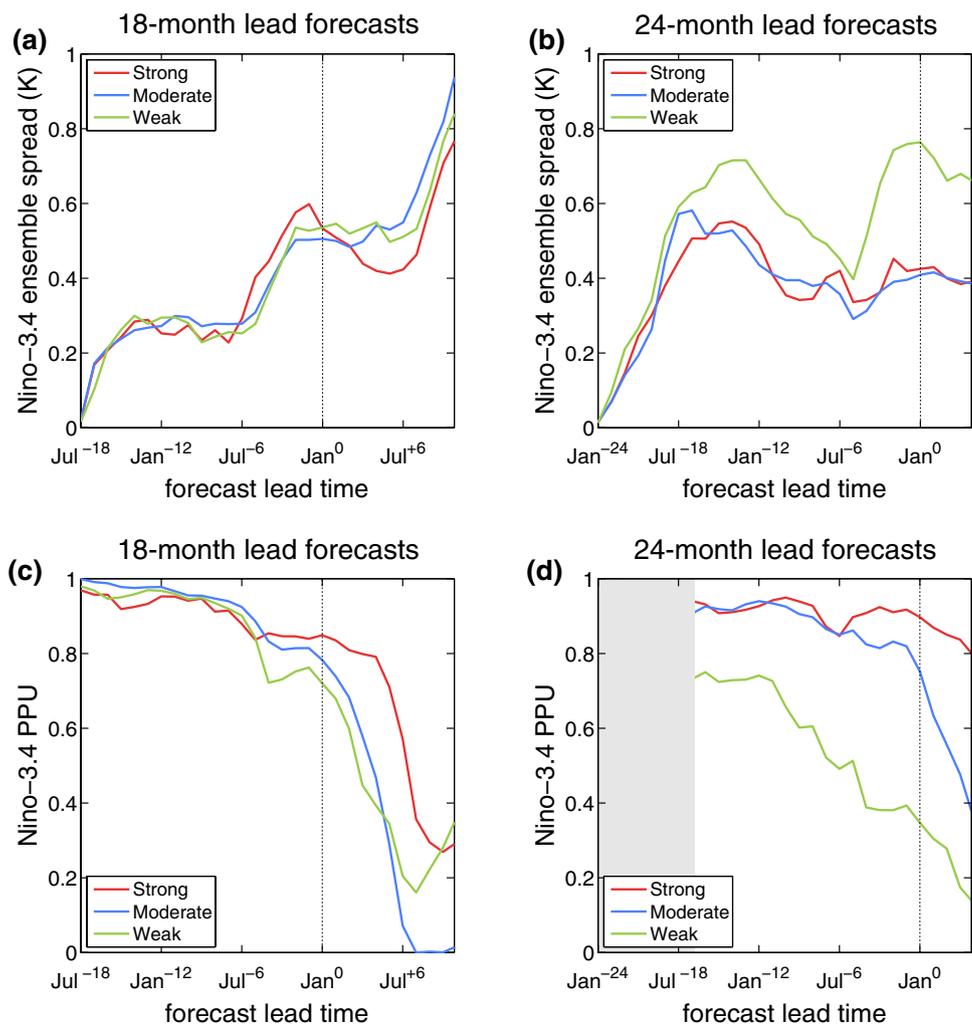
All three cases of the 24-month lead forecasts exhibit a similar evolution for the ensemble-mean \bar{Z}'_{TC} (Fig. 6, right; red lines), with the initial thermocline discharge rebounding to slight positive values after the onset of La Niña. After then, the “moderate” and “weak” cases

less than −0.45 K indicate return of La Niña conditions on the second year (blue shading), whereas values larger than 0.45 K indicate El Niño conditions during the second year (red shading). Orange dots indicate the forecasted events selected from the control simulation. Dark blue dots indicate the ensemble-mean forecasts. Correlation coefficient r between the predictor and the predict and are indicated for both the CESM1 control simulation (gray) and the ocean reanalysis data (magenta)

show \bar{Z}'_{TC} plumes with positive values consistent with the recharge phase (Fig. 6, right; lilac lines). The “strong” case shows \bar{Z}'_{TC} plumes close to zero for the second year. Lastly, the \bar{Z}'_{TC} plumes show little error growth during the first year for all cases. This suggest that the initial error growth of SST anomalies may be decoupled from the error growth of the thermocline anomalies, thus allowing for long lead predictability.

We conclude the analysis of the role of initial conditions by comparing the 18- and 24-month lead forecasts with the statistical predictors from the control simulation and from observational data. We find that the three “Discharge” cases exhibit ensemble-mean Niño-3.4⁰ values that are linearly related to the magnitude of the initial discharge (\bar{Z}'_{TC}^{-18} ; Fig. 8a, dark blue dots), consistent with the correlation between these predictors seen in the CESM1 control (Fig. 8a, gray dots). Observational data also show significant correlations between Niño-3.4⁰ and \bar{Z}'_{TC}^{-18} ($r = 0.44$ with 95% confidence; Fig. 8a, magenta dots). The three “Peak El Niño” cases are also consistent with the CESM1 control, both regarding the relationship of the ensemble-mean Niño-3.4⁰ with the magnitude of the preceding El Niño, Niño-3.4^{−24}, and the scatter (Fig. 8, blue and light blue dots respectively). The observed data, shows a similar correlation ($r = -0.27$) and scatter (Fig. 8, magenta dots). The scatter of the CESM control, the perfect model forecasts, and observations, agree in suggesting that strong El Niño events

Fig. 8 Temporal evolution of the spread (*top*) and potential prediction utility (PPU) (*bottom*) of the Niño-3.4 index in forecasts initialized **a** during the discharge phase and **b** at the peak of the preceding El Niño event. The *dashed black line* indicates January of the second year (Jan⁰). The initial 6 months of the El Niño forecasts are not shown (*shaded gray*) because the PPU becomes zero when the ensemble-mean Niño-3.4 index goes to zero during the transition from El Niño to La Niña



are likely to be followed by 2-year LN. In contrast, these lines of evidence suggest that weak El Niño events lead to less predictable 2-year LN.

The “moderate” and “weak” cases show a range of outcomes (Fig. 8, light blue dots) that is distributed within the scatter of the control run (Fig. 8, light blue dots). The “strong” cases, in contrast, do not map onto the distribution of the control, possibly because strong El Niño and their associated discharge are not sampled with sufficient frequency in the control. In other words, the “strong” forecasts show that the control run does not provide complete information of the outcomes associated with these extreme events. In these cases the perfect model forecasts provide more complete information on the amplitude of 2-year LN events. For instance, both strong cases have members predicting extreme Niño-3.4⁰ values < -2.5 K, suggesting that very strong 2-year LN events are possible, but extremely unlikely.

3.3 Predictability and error growth

The 18-month lead time forecasts show fast error growth, measured by the ensemble spread, from July to September, remaining constant afterward until July of the following year (Jul⁻⁶), when the error resumes the growth at about the same rate as the previous year (Fig. 7a). The seasonality of the spread suggests that error growth is governed by a mechanism active during late summer and early fall. The spread of the “Peak El Niño” forecasts evolves differently. All three El Niño forecasts show fast error growth during the first six months, followed by a decline in error growth (Fig. 7b). The “strong” and “moderate” cases show constant spread for an additional year that remains lower than the corresponding discharge forecasts (Fig. 7a, red and blue lines). The “weak” El Niño case, in contrast, resumes the error growth, which exceeds the spread of the corresponding discharge case by January of the second year (Jan⁰; Fig. 7a, green). This

is consistent with the results from the weak El Niño forecast, which show largest scatter for the amplitude of the 2nd year La Niña (Fig. 8b). It remains unclear why all El Niño forecasts show decreasing spread from Jul⁻¹⁸ to Jul⁻⁶. The reduction in spread could be indicative of the following two processes. First, initial error growth may not be driven by coupled process, therefore the coupled system may not be affected by the initial spread. Second, that the thermocline anomalies are not affected by this uncoupled variability, and once the system goes into La Niña, the spread decreases because the uncoupled variability decreases.

We computed the PPU of the Niño-3.4 SST indices as a measure of the usefulness of each ensemble forecast. PPU values near one are indicative of useful predictions, while values near zero indicate that the ensemble predictions lack skill. All forecasts start with PPU close to 1, which decays slowly until about Jan⁰, when all cases show values higher than 0.7 (Fig. 7c) indicative of useful forecasts. PPU quickly decays after the target forecast date of Jan⁰, indicating that the forecasts lose their usefulness beyond 18-month lead times. Each case, however, decays at a different rate. The strong discharge case shows the largest PPU values, which stays at about 0.85 by Jan⁰, whereas the weak discharge case shows lower values of about 0.75, consistent with a link between PPU and the magnitude of the initial discharge.

In contrast, not all “Peak El Niño” forecasts remain useful at 24 month lead times. The “strong” and “moderate” cases show values of PPU larger than 0.7 by Jan⁰ (Fig. 7d, red and blue lines), indicative of useful forecasts. The “weak” case, in contrast, shows much smaller PPU values throughout the forecasts reaching values under 0.4 by Jan⁰ (Fig. 7d, green line). The strong El Niño case shows the largest PPU values, whereas the weak case shows the lowest values, suggesting a link between PPU and the magnitude of the preceding El Niño. Note that by normalizing the spread by the square of the ensemble-mean, as is done in the PPU measure (Eq. 1), there is no longer a seasonal modulation of the error growth. This simply suggests that both the spread and the ensemble-mean have a similar seasonality, possibly because they are both controlled by the seasonality of the equatorial Pacific climate.

We further explored the processes causing error growth by analyzing the spatio-temporal evolution of predicted SST anomalies over the equatorial Pacific and Indian basins. Here we discuss results only from the *moderate* discharge and El Niño cases because they are representative of all other cases, particularly since the spread of the Niño-3.4 SST index evolves similarly in each type of forecast (Fig. 7, top). The ensemble-mean SST anomalies evolve similarly, with slightly larger amplitudes in the “Discharge” case compared to the “Peak El Niño” case (Fig. 9a, c). In particular, negative SST anomalies begin in the eastern Pacific

around Jun⁻¹⁹ and subsequently propagate westward over the following year, peaking around Jan⁻¹² (Fig. 9, left). The negative SST anomalies persist in the central and western Pacific over the following year, and then re-emerge in the eastern Pacific around December (Dec⁻¹).

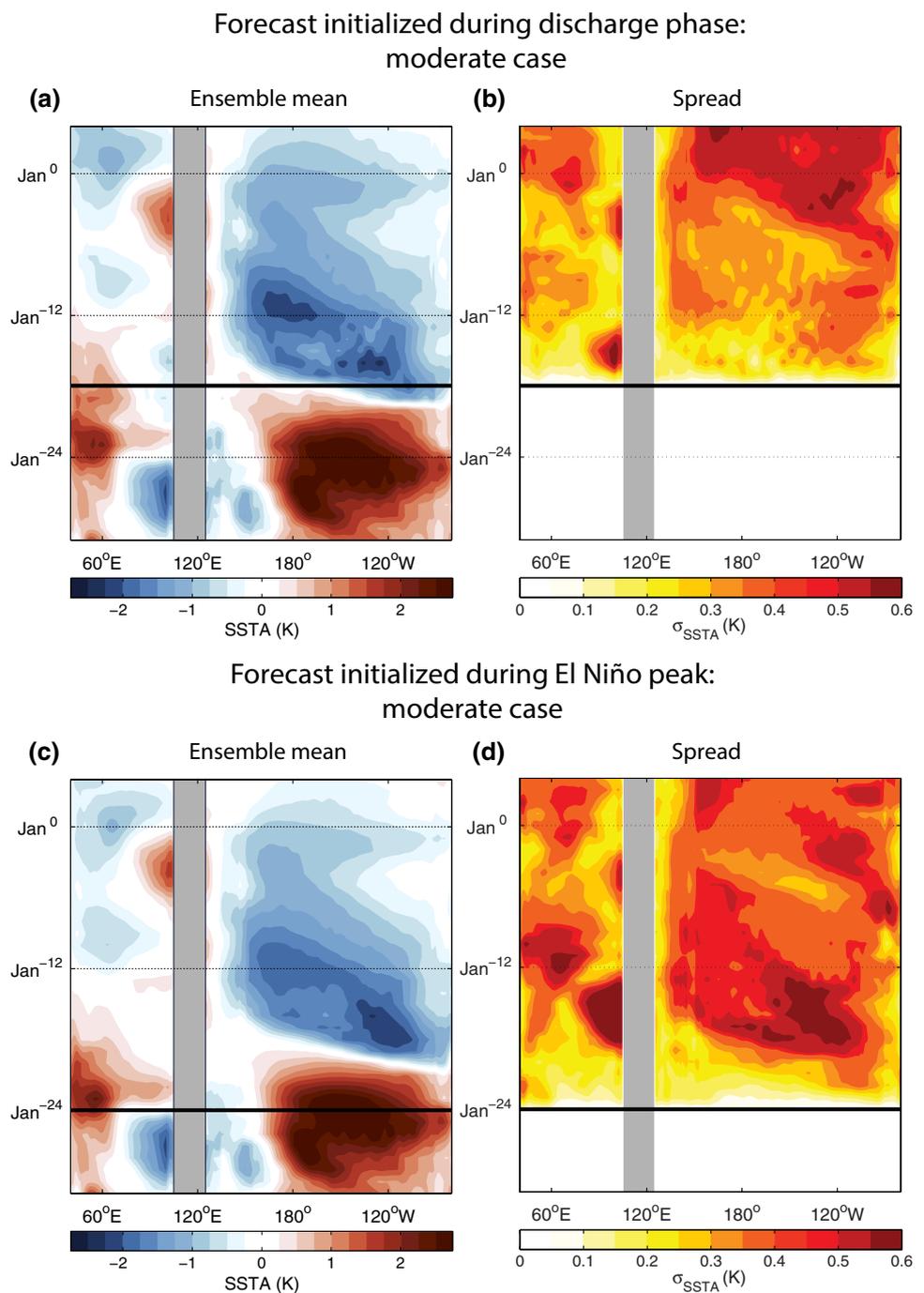
Both cases also show a similar evolution of the ensemble-mean SST anomalies in the Indian Ocean (IO), with weak negative (positive) SST anomalies over the eastern (western) basin around September of the first year (Sep⁻¹⁶). This pattern reverses during the second year, when the eastern (western) IO shows stronger positive (negative) SST anomalies. These SST anomalies show the same pattern and September–October–November (SON) seasonality as Indian Ocean Dipole (IOD) events observed in nature (Webster et al. 1999; Saji et al. 1999). However, the simulated evolution of IO SST anomalies, with a negligible warming over the eastern IO during the first La Niña year and a strong warming on the second year, is not supported by observations, which show strong negative IOD events during both years (not shown).

Despite the similarity of their ensemble-mean SST anomalies, the two cases exhibit pronounced differences in the spatial evolution of the SST spread (Fig. 9, right). Both cases show the largest initial error growth in the *eastern* equatorial Pacific with values up to 0.4 K during January of the first year for the discharge case (Fig. 9b, Jan⁻¹²) and over 0.6 K during July for the El Niño case (Fig. 9d, Jul⁻¹⁸). Additional differences between the two cases arise as the forecasts evolve. The spread of the discharge forecasts increases during the second year reaching values over 0.6 K at the peak of the second year La Niña (Fig. 9b, Jan⁰). In contrast, the spread of the El Niño forecasts decreases after the second year with values remaining below 0.5 K at the peak of the second year La Niña (Fig. 9d, Jan⁰). An additional similarity between the cases is the large forecast spread over the eastern Indian Ocean during September of the first and second years. The seasonality and spatial pattern of these maxima in the SST spread suggest a role for climate variability associated with the IOD. The eastern IO exhibits the largest spread at the beginning of the forecasts; however, we explore whether this error growth has an influence on the ENSO forecasts in the next subsection.

3.4 Role of Indian Ocean variability

Several studies suggest that IO variability associated with dipole events could influence the evolution of ENSO events (Saji and Yamagata 2003; Kug and Kang 2006; Kug et al. 2006; Ohba and Ueda 2007; Yamanaka et al. 2009; Annamalai et al. 2010; Izumo et al. 2010). Therefore the large forecasts spread seen over the eastern IO could be playing a role in the spread over the Pacific

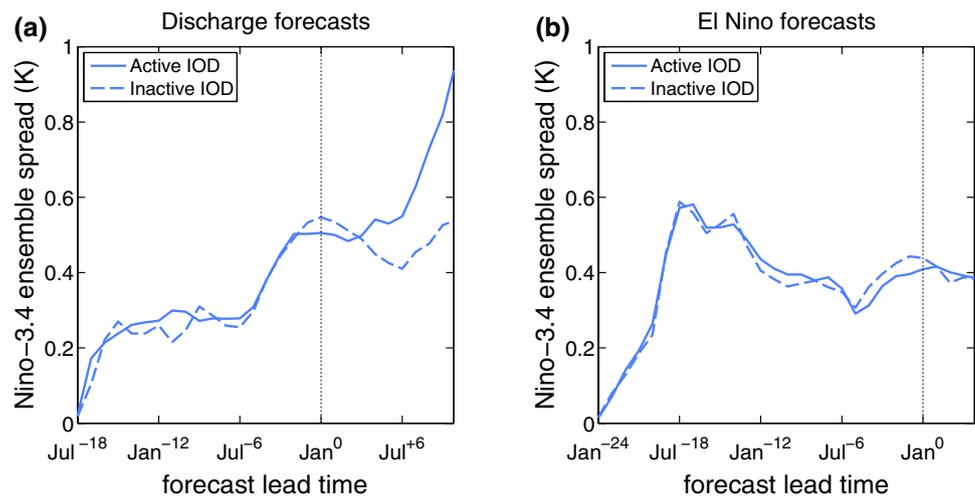
Fig. 9 Spatiotemporal evolution of equatorial sea-surface temperature (SST) anomalies predicted by CESM1 in the moderate discharge (*top*) and El Niño (*bottom*) cases. Ensemble mean (*left*) and spread (*right*) of the predicted SST anomalies. The forecast start date is marked by the *solid horizontal lines*. SST anomalies before these dates are from the control simulation. In this and subsequent figures, Jan⁻¹² and Jan⁰ indicate the peaks of the first and second years of La Niña. Jan⁻²⁴ indicates the peak of the preceding El Niño. SST anomalies are averaged over the 5°S–5°N equatorial band over the Pacific Ocean and over 10°S–0° over the Indian Ocean



(Fig. 9). Moreover, this SST spread over the IO shows a maximum around September, which could be related to the seasonality of the spread of the Niño-3.4 SST index (Fig. 7a, b). We performed an additional set of forecasts in order to test this link. In these forecasts we prescribed climatological SSTs in the eastern Indian Ocean (75°E–105°E 9°S–1°N). This effectively disables the variability associated with the eastern lobe of the IOD and therefore the forecasts show negligible ensemble-mean

and spread over the eastern equatorial IO. Beyond this difference, disabling the IOD does not lead to substantial differences in the evolution of the forecasts in the equatorial Pacific (not shown). Disabling the IOD also does not lead to obvious changes in the spread of the Niño-3.4 SST index (Fig. 10). These results suggest that the evolution of the La Niña forecasts is largely uncoupled from the strong IOD variability as suggested by the study of Ohba and Watanabe (2011).

Fig. 10 Temporal evolution of the spread of the Niño-3.4 index in forecasts initialized **a** during the discharge phase and **b** at the peak of the preceding El Niño event. *Dashed lines* indicate the spread of forecasts where the IOD is inactive



3.5 Role of other modes of variability

We explore additional mechanisms responsible for error growth using output from the perfect model forecasts. We identify additional predictors via lagged-correlation analysis of various anomaly fields [sea level pressure (SLP), wind stress, thermocline depth and SST] with the departures of the Niño-3.4 SST index from the ensemble-mean at the Jan⁻¹² and Jan⁰ peaks. Data from 60 “Discharge” forecasts and 60 moderate and strong “Peak El Niño” forecasts were combined in each calculation. These cases show similar error growth (Fig. 7, top), thus we expect that will be governed by similar processes, allowing us to combine their data and thus increase the statistical significance of the analysis. We removed the ensemble mean anomaly from each member’s predicted SLP, wind stress, thermocline depth, and SST anomalies in order to focus the analysis on the departures from the ensemble mean (i.e. ensemble-spread).

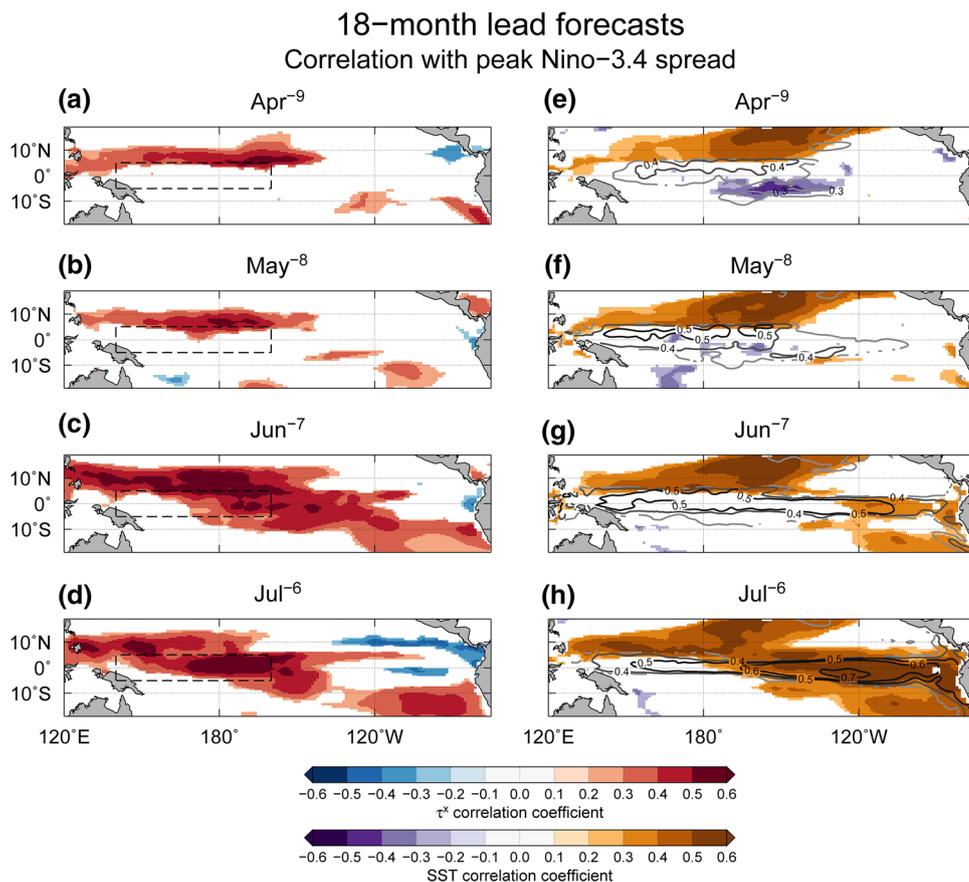
We first analyze processes within the equatorial Pacific related to the spread of the 2-year LN forecasts in both the “Discharge” and “Peak El Niño” cases. We focus our analysis on the correlations during months preceding the strong error growth prior to the second year peak, as shown by the spread of the Niño-3.4 SST index (Fig. 7, top). The correlation analysis shows that the spread of the Niño-3.4 SST index is positively correlated with the spread of zonal wind stress, τ'_x , over the western equatorial Pacific during the previous April (Apr⁻⁹, Figs. 11a, 12a, shading). Both cases show concurrent correlations with the Z'_{TC} spread characterized by a pattern of positive correlations extending across the equatorial Pacific (Figs. 11e, 12e, contours). The Z'_{TC} correlation patterns evolve with time and by July their magnitude becomes largest over the eastern equatorial Pacific (Figs. 11f–h, 12f–h, shading). Positive SST correlations appear in the eastern equatorial Pacific around June

and continue to increase and propagate westward by July (Figs. 11f–h, 12f–h, contours).

Overall, the evolution of these correlation patterns is consistent with the dynamics of the equatorial Kelvin waves triggered by wind fluctuations over the western equatorial Pacific. The positive Z'_{TC} and SST correlations in June and July also suggest that these Kelvin waves influence SSTs over the eastern equatorial Pacific. Moreover, as the westward propagation of the SST anomalies appear to reinforce the initial zonal wind anomalies over the western side of the basin, as shown by the positive τ'_x correlations (Figs. 11d, 12d, shading). We extended this analysis to months before Apr⁻⁹, however no significant correlations were found. This suggests that variability prior to first La Niña peak has little influence on the spread of the second year forecasts.

The relationship between τ'_x , Z'_{TC} , and SSTA revealed by the correlation analysis strongly suggests that coupled instabilities within the equatorial waveguide contribute to forecast spread. However, the analysis is unclear regarding the drivers of the initial wind stress variability. We identify subtle differences in the location of the initial zonal wind stress anomalies. The “Discharge” forecasts show initial τ'_x correlations centered around 10°N (Fig. 11a, shading). This τ'_x variability coincides with SST variability over the subtropical north Pacific (Fig. 11e, shading) reminiscent of the “seasonal footprinting” mechanism (Chang et al. 2007; Vimont et al. 2001, 2003, 2009). We note that the τ'_x correlations do not fully project on the equatorial waveguide, thus may be unable to explain the Z'_{TC} variability revealed by correlations (Fig. 11e, contours). However, off-equatorial wind variability could trigger equatorial Kelvin waves via the so-called “trade wind-induced charging” mechanism proposed by Anderson et al. (2013), and thus generate SST spread in the Niño-3.4 region (Fig. 11h, shading).

Fig. 11 Correlation coefficient of zonal wind stress (left shading), thermocline depth (contours) and sea-surface temperature (right shading) with the peak Niño-3.4 SST spread in the “Discharge” forecasts. Correlation coefficients with 95% statistical confidence are shown. The rectangle indicates the region used to compute the $\tau_{\text{WeqPac-9}}^x$ index used in the statistical prediction model (Eq. 2)



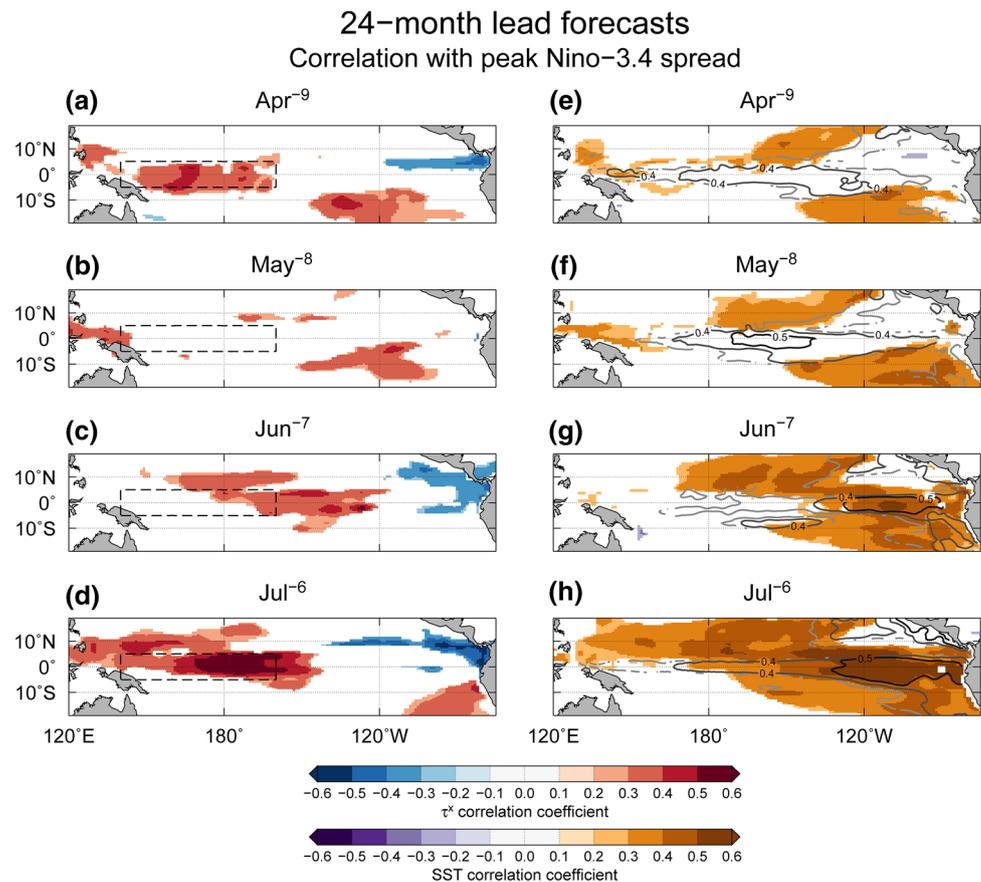
In contrast, the “Peak El Niño” forecasts show correlated wind variability centered on the western equatorial Pacific (Fig. 12a, shading), which could effectively trigger Kelvin waves, as suggested by the Z'_{TC} correlations (Fig. 12e, contours), generating SST variability over the Niño-3.4 region (Fig. 12h, shading).

Additional predictors outside the equatorial waveguide were explored focusing on the summer season prior to both the first and second peaks. Note that in this analysis we also focus on the mechanisms leading to spread of the first year forecasts, as they could show similarities with the mechanisms generating spread at the second year peak. Prior to performing the lag-correlation analysis, we regressed each member’s anomaly fields on the corresponding May⁻⁸ value of the zonally averaged Z'_{TC} . This procedure allows us remove the effect of the equatorial processes discussed above and focus on additional processes. The resulting correlation maps for July prior to the first peak (Jul⁻¹⁸) show few areas over the tropics correlated with the Niño-3.4 SST spread 6 months later (Fig. 13a). The sole exception is the southeastern (SE) tropical Pacific, where SLP anomalies show significant negative correlations. The negative correlation between SLP anomalies persists during August and September when SSTs also start to be positively correlated, showing the largest values over the SE Pacific (Fig. 13b, c).

The negative correlations indicate that negative SLP anomalies in the SE Pacific are related to positive departures of the Niño-3.4 SST index relative to the ensemble mean. This suggests that low SLP anomalies are associated with Niño-3.4 SST forecasts that are warmer than the ensemble mean, and conversely high SLP anomalies with cooler La Niña forecasts. The correlations between SST and Niño-3.4 index are initially stronger off the equator (Fig. 13b) and then become stronger in the eastern equatorial Pacific (Fig. 13c). A similar evolution of SLP and SST anomalies is found for the second year (Fig. 13, bottom), suggesting that a common mechanism influences the error growth of the La Niña forecasts during the first and second years. The second year exhibits negative SLP correlations associated with positive SST correlations over the central north Pacific from May to July (Fig 13d–f). In addition, positive SLP correlations are found over the western equatorial Pacific during June and July (Fig 13e, f).

The El Niño forecasts show similar evolution of SLP and SST correlations, with the only difference that the anomalies in the north Pacific are also present during the first year (Fig. 14, left). The spatial pattern and temporal evolution of these negative and positive SLP and SST correlations suggests that circulation anomalies over the SE and NE Pacific contribute to the spread of Niño-3.4 SSTs. These SST

Fig. 12 As in Fig. 11, but for “Peak El Niño” forecasts



anomalies first emerge in the SE Pacific and then propagate equatorward. The negative SLP correlation suggests that a cyclonic anomaly in the SE Pacific is associated with warming over the eastern equatorial Pacific, resulting in a Niño-3.4 forecast that is warmer than the ensemble mean. Conversely, an anticyclonic anomaly in the SE Pacific is associated with cooling over the eastern equatorial Pacific, resulting in a Niño-3.4 forecast that is colder than the ensemble mean. In all cases the magnitude of the negative correlation increases with time as the positive SST correlations emerge, indicating that SST anomalies are driven by circulation anomalies. Moreover, the correlation maps do not show evidence for large-scale wave train patterns prior to the occurrence of the SLP correlations, suggesting that local internal atmospheric dynamics are more likely to contribute to variability in the southeast trade winds. A similar mechanism could be invoked to explain the SLP and SST correlation results over the NE Pacific.

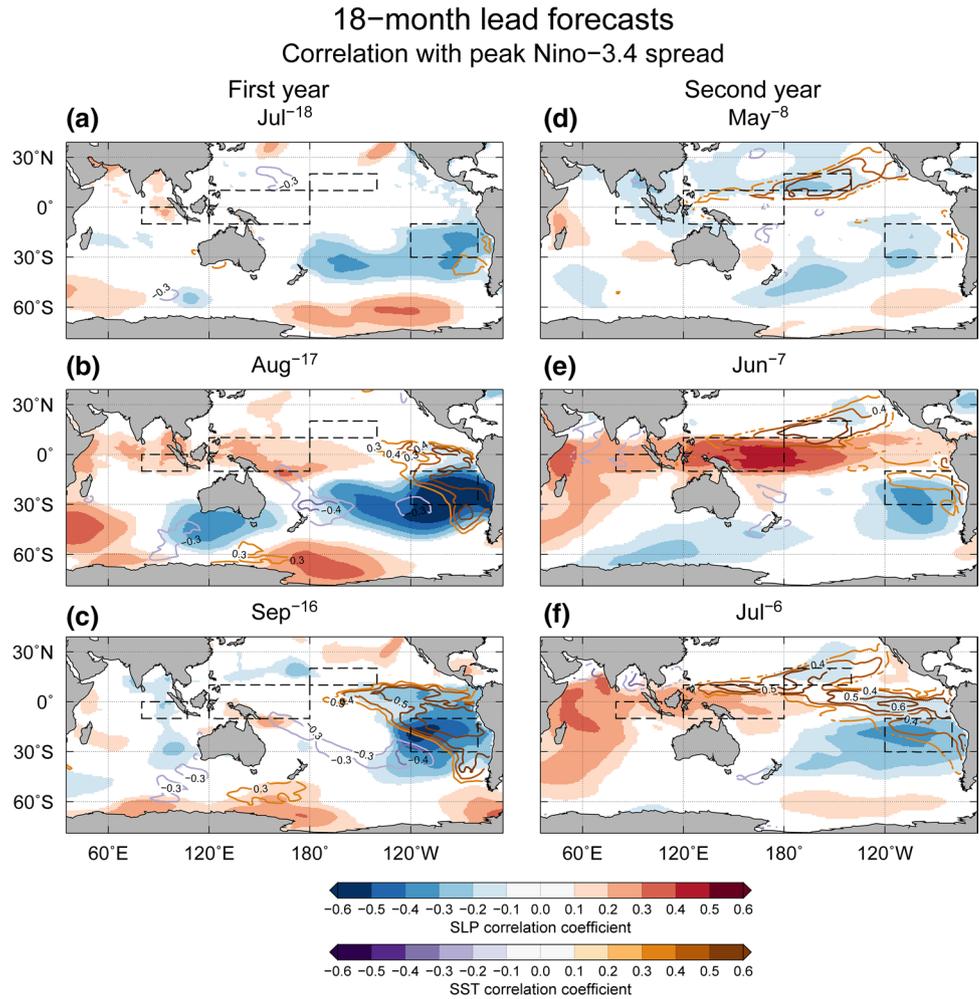
The spatial pattern and temporal evolution of these correlations are consistent with the South Pacific Meridional Mode (Zhang et al. 2014), according to which, fluctuations in the strength of the SE trades associated with regional-scale cyclonic anomalies can have an influence on the equatorial Pacific. According to this mechanism, trade wind variability associated with monthly and sub-monthly

fluctuations of the atmospheric circulation in the mid-latitude South Pacific can initiate SST anomalies in the SE Pacific which can then propagate westward and equatorward through the wind–evaporation–SST (WES) feedback mechanism (Xie and Philander 1994). Moreover, the initial SLP pattern, not only shows negative correlations over the tropical SE Pacific, but also a lobe of positive correlations farther south. This SLP dipole resembles the Pacific–South American (PSA) pattern, the leading mode of internal atmospheric variability over the South Pacific. This mode plays an important role in modulating the trade winds and sea surface temperature (SST) in the southeast tropical Pacific (Okumura 2013). These fluctuations in the trade winds, in the presence of ocean dynamics, can be amplified by the positive Bjerknes feedback, altering the evolution of the La Niña forecasts.

3.6 Multivariate statistical prediction models

We conclude the study with a systematic evaluation of predictors in the long control simulation. Here we combine our understanding on the subsurface ocean predictors developed in Sect. 3.2 with the atmospheric predictors suggested from the results in Sect. 3.5. We formulated a series of multi-variate regression models of Niño-3.4⁰ by linearly

Fig. 13 Correlation coefficient of SLP (shading) and SST (contours) anomalies with the peak Niño-3.4 SST spread in the “Discharge” forecasts. Months prior to strong error growth are shown. Correlations with the spread of the Niño-3.4 SST index during the first year peak (Niño-3.4⁻¹², left) and second year peak (Niño-3.4⁰, right). Correlation coefficients with 95% statistical confidence are shown. The rectangles indicate the regions used to compute the DMI⁻³, SLP⁻⁷_{WeqPac}, SLP⁻⁶_{SEpac} and SLP⁻⁶_{NPac} indices used in the statistical prediction model (Eq. 2)



combining these predictors with the thermocline predictors discussed in the first part of our study. The model showing the highest correlations ($r = 0.85 \pm 0.04$; error given at 99% confidence level) with Niño-3.4⁰ is:

$$\text{Niño} - 3.4^0 = a_1 \bar{Z}'_{TC}{}^{-18} + a_2 \tau^x_{\text{WeqPac}^{-9}} + a_3 \bar{Z}'_{TC}{}^{-6} + a_4 \text{SLP}^{-6}_{\text{SEpac}} + a_5, \tag{2}$$

where $\bar{Z}'_{TC}{}^{-18}$ and $\bar{Z}'_{TC}{}^{-6}$ are the “discharge” and “recharge” predictors discussed in Sect. 3.1, $\tau^x_{\text{WeqPac}^{-9}}$ is an index of zonal wind variability in the western equatorial Pacific, and $\text{SLP}^{-6}_{\text{SEpac}}$ is an index of SLP variability in the SE Pacific. The $\tau^x_{\text{WeqPac}^{-9}}$ index is defined as the zonal wind stress anomalies averaged over the western equatorial Pacific (120°E–160°W 30°N–10°S) and shows peak correlation ($r = 0.63$) with Niño-3.4⁰ at a lead time of 9 months. The $\text{SLP}^{-6}_{\text{SEpac}}$ index is defined as the SLP anomalies averaged over the SE Pacific (120°W–80°W 30°S–10°S) and shows peak correlation ($r = -0.70$) with Niño-3.4⁰ at a lead time of 6 months. The a_i constants in (Eq. 2) are computed using least-squares linear regression. Table 3 summarizes the

18-month lead forecasts
Correlation with peak Niño-3.4 spread

results from this analysis and provides more details on the definitions and lags of the different predictors.

Additional indices were considered, however these predictors did not increase the correlation significantly (Table 3). An index of SLP anomalies over the central north Pacific (180°–140°W 10°N–20°N) is highly correlated with Niño-3.4⁰ ($r = -0.40$) at a lead time of 6 months. However, including this $\text{SLP}^{-6}_{\text{NPac}}$ predictor in Eq. 2 does not lead to a statistically significant increase in the correlation with Niño-3.4⁰. We also considered the Dipole Mode Index (DMI) (Saji et al. 1999) during September of the first year (DMI⁻³), which is highly correlated ($r = -0.61$) with Niño-3.4⁰; however, including this predictor in Eq. 2 does not improve the correlation.

We also explored the role of SSTs over the tropical Atlantic. We found that SST anomalies averaged over the equatorial Atlantic (0°–20°E 5°S–5°N), the so-called ATL3 SST index (Zebiak 1993), is correlated ($r = -0.36$) with Niño-3.4⁰ with a lag of 12 months, suggesting that tropical Atlantic SSTs could have an influence on the evolution of La Niña. Including this ATL3⁻¹² predictor in

Fig. 14 As in Fig. 13, but for forecasts initialized at El Niño peak

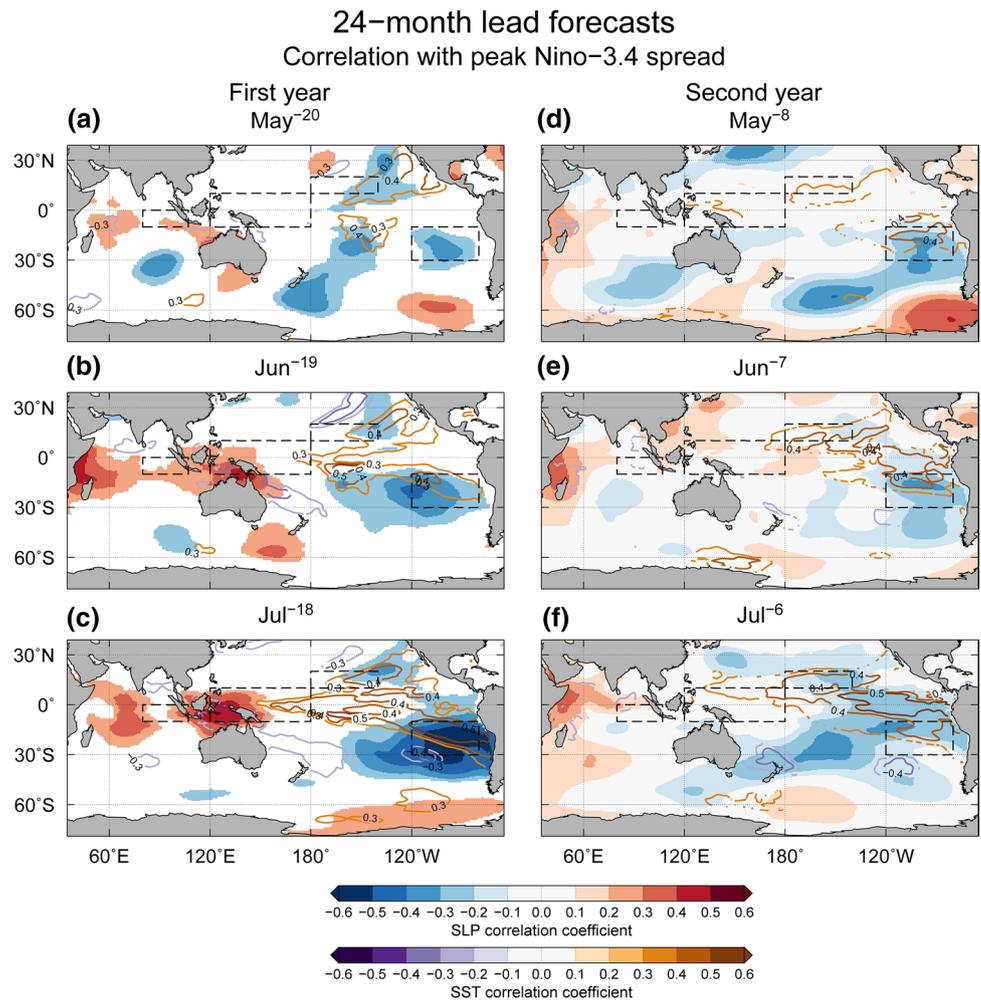


Table 3 Indices considered as predictors of the magnitude of the Niño-3.4 SST index on the second year (Niño-3.4⁰)

Predictor						r_{model}
ID	Name	Variable	Area	Lead (months)	$r_{\text{Niño-3.4}^0}$	
$\bar{Z}'_{TC}{}^{-18}$	Discharge	Thermocline depth	140°W–80°W 5°S–5°N	18	0.46	0.46
$\tau^x_{\text{WeqPac}}{}^{-9}$	Coupled instabilities	Zonal winds	140°E–160°W 5°S–5°N	9	0.63	0.65
$\bar{Z}'_{TC}{}^{-6}$	Recharge	Thermocline depth	140°W–80°W 5°S–5°N	6	0.46	0.73
$\text{SLP}_{\text{SE Pac}}{}^{-6}$	S. Pacific meridional mode	Sea level pressure	120°W–80°W 30°S–10°S	6	–0.70	0.85
DMI^{-3}	Indian Ocean Dipole	Sea-surface temperature	80°E–107°E 10°S–0°	3	–0.58	0.85
Niño-3.4^{-12}	1 st year La Niña	Sea-surface temperature	120°W–170°W 5°S–5°N	12	0.12	0.85
$\text{SLP}_{\text{Weq Pac}}{}^{-7}$	Western equatorial Pacific	Sea level pressure	120°W–180° 10°S–10°N	6	0.20	0.85
$\text{SLP}_{\text{NPac}}{}^{-6}$	N. Pacific meridional mode	Sea level pressure	180°–140°W 10°N–20°N	6	–0.40	0.83
ATL3^{-12}	Atlantic Niño	Sea-surface temperature	0°–20°E 5°S–5°N	12	–0.36	0.83

The first five columns provide information on each predictor. The sixth column shows the correlation coefficient (r) between each predictor with Niño-3.4⁰. The seventh column shows r values of the linear multi-variate model combining all the predictors listed in the table from top ($\bar{Z}'_{TC}{}^{-18}$) to the given predictor. Bold r values indicate when the increase in r relative to the previous model is statistically significant with 99% confidence

Eq. 2 does not lead to a substantial increase in correlation ($r = 0.83 \pm 0.08$) mainly because $ATL3^{-12}$ is also correlated ($r = -0.28$) with $\bar{Z}'_{TC}{}^{-18}$. In general, these predictors are highly correlated with $\bar{Z}'_{TC}{}^{-18}$, which is related to the magnitude of preceding El Niño: thus the bulk of the predictable information on the duration of La Niña is already included in $\bar{Z}'_{TC}{}^{-18}$.

4 Discussion

The purpose of this study was to assess the predictability of the duration of La Niña out to 2 year lead times in a perfect-model setting. In particular, we explored predictability as a function of ocean initial state during different phases of ENSO. We performed ensembles of forecasts starting from two different types of initial states. The first state occurs during the transition from El Niño to La Niña, during which the equatorial Pacific is characterized by near-zero SST anomalies and a shallower than normal thermocline. In this state the thermocline has not fully adjusted to prior wind anomalies and therefore provides a source of predictability for the evolution into La Niña state. The second state occurs at the peak of El Niño, when the source of predictability is the associated westerly wind anomaly, the driver of the delayed thermocline shoaling responsible for the onset of La Niña. The main objective of our study was to explore how different magnitudes of these “discharge” and “Peak El Niño” predictors control the reliability and skill of forecasts of 2-year La Niña.

Our results show that the forecasts initialized with the largest discharge and peak El Niño amplitudes exhibit the largest skill at predicting a 2-year La Niña event (according to the PPU metric). The higher skill in these cases stems from the larger magnitude of the initial conditions, which leads to a larger predictable signal throughout the first 2 years of the forecast. The forecast spread, in contrast, appears to be independent of the magnitude of the initial conditions. Our results show that CESM1 can produce consistent forecasts of La Niña for the first and second years (up to 24 month lead time), mainly because the ensemble-mean signal has sufficiently large amplitude to overcome the forecast spread. The forecasts become significantly less skillful and reliable beyond the second year peak, despite robust ensemble-mean signals showing La Niña conditions for an additional (third) year. For instance, the El Niño event selected for the “strong” cases was followed by a 3-year La Niña (Figs. 5, 6, gray lines). However, only 13 in 30 members predicted a 3-year La Niña, rendering the forecast considerably less reliable. This occurs because the spread grows over time, dominating the decreasing ensemble-mean signal. Our conclusion that the ensemble-mean component has a larger contribution to prediction utility

than the ensemble spread is consistent with previous studies showing that ensemble spread is not a good indicator of skill (Kirtman et al. 1997). This is a well known difference with weather prediction, where ensemble spread is the main determinant of potential forecast skill. Because the ENSO system could be viewed as a stochastically forced damped linear system, the forecasts would also be controlled by the magnitude of the initial conditions. In other words, the ensemble-mean component of the PPU, in contrast to ensemble-spread, is related to the initial conditions, as shown by Kleeman (2002) using a simplified ENSO model.

We used the perfect model forecasts to investigate how early in the evolution of an ENSO event can the duration of La Niña be skillfully predicted. The fact that the discharge and the peak El Niño amplitude are highly correlated allows the forecasts initialized at the peak of an El Niño event to remain as skillful as those initialized at the discharge, despite being initialized six months earlier. Particularly, we find that the forecasts initialized at the peak of *strong* El Niño events are more skillful, measured by the PPU value at target date of Jan^0 , as forecasts initialized during their discharge phase. Because of the tight El Niño–discharge relationship, particularly for strong El Niño events, the predictability (in a PPU sense) of the “Peak El Niño” and the “Discharge” forecasts will have the same contribution from the ensemble-mean signal and the differences in PPU will arise from differences in the spread. Note that the strong “Peak El Niño” ensemble forecast has a PPU of 0.90 by the target date of Jan^0 , while the strong “Discharge” forecast has a lower PPU of 0.85 (Fig. 8, bottom, red curves), indicating that the former forecast is more skillful despite the longer lead time. The higher PPU of the “Peak El Niño” forecasts relative to the “Discharge” occurs because these forecasts have lower spread by Jan^0 (Fig. 8, top, red curves). The lower Jan^0 spread in the “Peak El Niño” forecasts (relative to the “Discharge” forecasts) maybe related to the seasonality and state of ENSO. Further investigation of the seasonality of the spread and dependence on ENSO state is beyond the scope of this study and will be left for future work.

In a linear ENSO system, strong events would drive a commensurate response of the delayed thermocline feedback, thus leading to events with the same duration. Our results, in contrast, suggest that the predictability of La Niña duration (in a PPU sense) is consistent with processes governed by persistence, such that, events driven by large initial perturbations will take longer to decay. However, the enhanced persistence of La Niña should not be interpreted as the effect of thermodynamic damping; rather it is caused by nonlinearities in dynamical feedbacks as proposed by DD14. The nonlinearity of the delayed thermocline feedback is caused by two processes. First, strong El

Niño events have much larger wind anomalies than La Niña events and therefore drive much larger discharge (Frauen and Dommenges 2010; Choi et al. 2013). The recharge, in contrast, is driven by the magnitude of the La Niña wind anomalies, which are proportionally weaker. The second process involves coupling between the thermocline and the mixed layer. DD14 propose that shoaling of the thermocline during the discharge has a much larger influence on SSTs because the thermocline can continuously shoal, effectively entraining cold thermocline waters into the mixed layer. This effect is not limited—in fact becomes more pronounced—as the thermocline becomes shallower (e.g. in response to strong El Niño events). In contrast, deepening of the thermocline *decouples* the thermocline waters from the mixed layer. In this case, strong recharge and associated positive thermocline depth anomalies become increasingly ineffective at detrainning cold waters from the mixed layer. This asymmetry appears to be more pronounced in the central Pacific where the climatological mixed layer is shallower than the thermocline.

We speculate that this nonlinearity could also lead to threshold behavior in the evolution of 2-year LN. This idea is motivated by the fact that the perfect model forecasts do not completely map onto the $\bar{Z}'_{TC}{}^{-18}$ -Niño-3.4⁰ space defined by the events of the control. A sizable fraction of La Niña events simulated in the control are preceded by weak discharge, last 1 year, and are followed by moderate and strong El Niño in the second year (Fig. 8a, gray dots over the region approximately defined by $\bar{Z}'_{TC}{}^{-18} > -20$ m and Niño-3.4⁰ > 0.45 K). In contrast, none of the ensemble members initialized from weak discharge conditions lead to El Niño on the second year as seen in many events from the control. Instead, most of the perfect model forecasts predict La Niña conditions for a second year (17 out of 20) and a lesser fraction (3 out of 20) show neutral conditions in the second year (Fig. 8a, right blue dots). We cannot rule out that our ensemble size is insufficient to fully describe the range of events simulated in the control simulation. However, it is also possible that the forecasts undergo a bifurcation depending on the magnitude of the discharge, resulting in forecasts with a more pronounced tendency to evolve into 2-year LN, as in our case, or forecasts with a tendency to evolve into neutral or El Niño. It is also possible that this threshold behavior is controlled by decadal variability, and our selected events coincided with periods more conducive to 2-year LN. Both hypotheses deserve further exploration since they could lead to additional sources of predictability.

The discharge-recharge asymmetry is reflected in the correlation between the magnitude of the discharge ($\bar{Z}'_{TC}{}^{-18}$) and the value of the Niño-3.4 SST index (Niño-3.4⁰) 18 months later. La Niña events preceded by strong discharge must last 2 years in order for the recharge to drive SST anomalies back to neutral. Our analyses of the CESM1

control and perfect model forecasts show that the initial conditions characterized by large-amplitude $\bar{Z}'_{TC}{}^{-18}$ virtually always lead to a 2-year LN. The large PPU of these cases suggests that operational forecasts, which are generally less skillful due to initial condition uncertainty and model drift, could still provide reliable forecasts of 2-year La Niña. Conversely, our analysis of observational data indicates that the $\bar{Z}'_{TC}{}^{-18}$ and Niño-3.4⁻²⁴ predictors could be used to determine the return of La Niña. For instance, the La Niña following the 1997/98 El Niño is one of three observed events falling in the area of the Niño-3.4⁻²⁴-Niño-3.4⁰ space where CESM1 would have predicted a 2 year duration. The first 2 years of this La Niña event could have been predicted using the correlation between the peak El Niño amplitude and the magnitude of the 2-year La Niña. Note, however, that other events, such as the 2010–2012 La Niña, lie in a statistical regime where both the CESM1 control and the perfect model forecasts are less reliable. Last, the limited number of events with subsurface reanalysis data precludes a more robust assessment of this relationship in observations.

The evolution of the spread of the Niño-3.4 SST index suggest that the processes causing error growth are modulated by both the state of ENSO and the seasonal cycle. The forecasts initialized during the “discharge” phase show error growth during late boreal summer and fall, and no error growth during the rest of the calendar year. Similar behaviour has been reported in July-initialized forecasts of El Niño onset performed with CCSM4, a previous version of our model (Larson and Kirtman 2015). Our correlation analysis of the forecast spread shows that most of the forecast error is associated with (1) coupled instabilities in the equatorial Pacific and (2) trade wind variability associated with the South Pacific meridional mode. Our analysis shows that coupled instabilities are triggered by stochastic wind variability in boreal spring (March/April), influencing Niño-3.4 SSTs via Kelvin waves one or two seasons later (i.e. during boreal and fall summer). Variability in the trade winds is caused with winter-time atmospheric variability in the Southern Hemisphere, which is stronger during austral winter (boreal summer). This extratropical variability influences Niño-3.4 SSTs via mixed-layer processes taking about a season to have an impact on the central equatorial Pacific SSTs and thus on forecasts error. These two processes together could explain the large Niño-3.4 spread during boreal summer and fall. The subsequent plateau of forecast spread occurs because coupled instabilities subside during boreal fall when the equatorial Pacific is seasonally colder, and also during austral summer because atmospheric variability in the SE Pacific is seasonally weaker.

The forecasts initialized at the peak of El Niño show the largest spread in the first 3 months, which then stabilizes at a lower level throughout the subsequent 2 years. This initial forecast error does not appear to impact the skill of

the 2-year LN predictions, which overall exhibit less spread than the “discharge” forecasts, particularly during the target forecast season (DJF of the second year). The decrease in spread after spring may be caused by the emergence of La Niña and the suppression of convective instability by negative SST anomalies. This suggests that the more predictable thermocline anomalies are able to “pull through” the predictability barrier, resulting in skillful forecasts. While exploring these seasonal and state dependencies of the error growth would require a much larger number of forecasts, our results suggest that the spring predictability barrier may be less influential when initializing at specific ENSO phases. This results agree with a recent study suggesting that thermocline precursors during the onset of El Niño can also overcome predictability loss during boreal spring (Larson and Kirtman 2016).

We find two key processes causing error growth in the predictions of 2-year LN. The first one involves coupled instabilities within the equatorial Pacific. Our correlation analysis of the perfect model forecasts shows that these instabilities are initiated around April when stochastic wind variability over the western Pacific triggers ocean Kelvin waves. These waves reach the eastern equatorial Pacific two months later initiating the spread of SST anomalies. This result agrees with recent studies showing the importance of high frequency wind variability over the western equatorial Pacific in the evolution of El Niño events (McPhaden and Yu 1999; Thompson and Battisti 2001; Fedorov et al. 2003; Chiodi and Harrison 2015; Larson and Kirtman 2015). According to these studies, Kelvin waves triggered by westerly wind bursts over the western Pacific can influence SSTs over the eastern as seen in our La Niña forecasts. We find that April prior to peak 2-year La Niña is an important window for coupled instabilities to impact the final forecasts spread. This result agrees with the study of Larson and Kirtman (2016), which using a previous version of the model, showed that March is when stochastic zonal wind stress is more effective at initiating coupled instabilities that have a lasting influence forecast error.

The second mechanism involves atmospheric variability over the SE Pacific in agreement with previous observational studies suggesting that the Southern Oscillation could also be a trigger of ENSO events (van Loon and Shea 1985; Trenberth and Shea 1987). The spatial pattern and temporal evolution of these correlations are consistent with the South Pacific Meridional Mode (Zhang et al. 2014). According to this mechanism, trade wind variability associated with the monthly and sub monthly cyclonic variability in the extra tropical SE Pacific can initiate SST anomalies in the SE Pacific which can then propagate westward and equatorward through the wind–evaporation–SST (WES) feedback (Xie and Philander 1994). In our forecasts, this mechanism is initiated around July by mid-latitude

atmospheric variability, which influences the spread of equatorial SSTs by August.

Both mechanisms appear to be active during the evolution of 2-year LN, however with a less disruptive effect than during the onset of El Niño, because of the more predictable influence of the thermocline dynamics discuss in this study. Our statistical model derived from the control simulation suggest that both processes could play a comparable role in altering the evolution of 2-year LN. The extent to which forecasts of 2-year LN events might be improved by considering these additional predictors is left for future work.

Our correlation analysis also reveals SLP and SST patterns in the North Pacific reminiscent of the meridional mode and seasonal “foot printing” mechanisms (Chang et al. 2007; Vimont et al. 2001, 2003, 2009). However, it remains unclear whether these processes influence the evolution of the forecasts. Off-equatorial wind variability associated with these mechanisms could be initiating coupled instabilities leading to spread in the “discharge” forecasts. The “Peak El Niño” forecasts do not show evidence for this link. Moreover, our predictor analysis of the long control did not reveal statistically significant correlation between indices of north Pacific SST or SLP variability and 2-year LN. Further research is needed to elucidate whether these modes of variability play a key role on the evolution of La Niña forecasts or not. Similarly, SST variability in the Indian and tropical Atlantic show correlations with 2-year LN; however these predictors are also correlated with the peak magnitude of the preceding El Niño and thus do not provide additional independent information. Our experiments with disabled IOD confirm that this mode does not play a role in the error growth, in agreement with a previous study showing that IO variability accelerates the transition from El Niño to La Niña, but does not play a role in the termination of La Niña (Ohba and Watanabe 2011).

5 Conclusions

We assessed the predictability of the duration of La Niña using CESM1, a climate model that simulates realistic ENSO characteristics, including the multi-year duration of La Niña. CESM1 skillfully predicts the return of La Niña for a second year in forecasts initialized from ocean states characterized by a strong discharge of the equatorial thermocline or by a strong El Niño. The high model skill, as measured by PPU, stems from larger ensemble-mean signals, particularly since the spread evolves independently of the magnitude of the initial conditions. Thus, initial conditions with weaker anomalies lead to less skillful forecasts, mainly because the ensemble-mean signal becomes weaker and is more influenced by error growth. Our results

show that the predictability of La Niña duration is consistent with a process governed by persistence of thermocline depth anomalies, wherein large thermocline depth anomalies will decay over multiple years. The enhanced persistence of La Niña should not be interpreted as the effect of thermodynamic damping; rather it is caused by nonlinearities in dynamical feedbacks as proposed by DD14.

The high skill demonstrated by the CESM1 perfect-model forecasts suggests that useful operational forecasts of the duration of La Niña in nature may be feasible. The magnitude of the thermocline discharge in the equatorial Pacific could be used as an observational diagnostic of the likelihood of a 2-year LN event. For instance, the La Niña following the 1997/98 El Niño is among very few observed events falling in the range of $\bar{Z}'_{TC}-18$ values where CESM1 would have consistently predicted a 2 year duration. The first 2 years of this 3-year La Niña event could therefore have been predicted using the statistical relationship put forward in this study. Our results also suggest that the realism of ENSO simulations could be critical for achieving these long-range predictions. Therefore improvements in the simulated characteristics of ENSO may lead to further improvements in predictability. Due to the disproportionate impact of La Niña on drought throughout the world, such further research could result in concrete and direct societal benefits.

Acknowledgements This study was supported by NOAA's Climate Program Office's Modeling, Analysis, Predictions, and Projections program (Grant NA14OAR4310229). PDN received additional support from the Climate and Global Dynamics division during a long term visit to NCAR. Computing resources were provided by the Climate Simulation Laboratory at NCAR's Computational and Information Systems Laboratory (CISL), sponsored by the National Science Foundation (NSF) and other agencies. NCAR is sponsored by the NSF.

References

- Anderson BT, Perez RC, Karspeck A (2013) Triggering of El Niño onset through trade wind-induced charging of the equatorial Pacific. *Geophys Res Lett* 40(6):1212–1216. doi:[10.1002/grl.50200](https://doi.org/10.1002/grl.50200)
- Annamalai H, Kida S, Hafner J (2010) Potential Impact of the tropical Indian Ocean-Indonesian Seas on El Niño characteristics. *J Clim* 23(14):3933–3952. doi:[10.1175/2010JCLI3396.1](https://doi.org/10.1175/2010JCLI3396.1)
- Balmaseda MA, Mogensen K, Weaver AT (2013) Evaluation of the ECMWF ocean reanalysis system ORAS4. *Q J R Meteorol Soc* 139(674):1132–1161. doi:[10.1002/qj.2063](https://doi.org/10.1002/qj.2063)
- Cane MA, Zebiak SE (1985) A theory for El Niño and the Southern Oscillation. *Science* 228(4703):1085–1087, doi: [10.1126/science.228.4703.1085](https://doi.org/10.1126/science.228.4703.1085). <http://www.sciencemag.org/content/228/4703/1085.abstract>. <http://www.sciencemag.org/content/228/4703/1085.full.pdf>
- Chang P, Zhang L, Saravanan R, Vimont DJ, Chiang JCH, Ji L, Seidel H, Tippett MK (2007) Pacific meridional mode and El Niño-Southern oscillation. *Geophys Res Lett* 34(16): doi:[10.1029/2007GL030302](https://doi.org/10.1029/2007GL030302)
- Chiodi AM, Harrison DE (2015) Equatorial Pacific easterly wind surges and the Onset of La Niña events. *J Clim* 28(2):776–792. doi:[10.1175/JCLI-D-14-00227.1](https://doi.org/10.1175/JCLI-D-14-00227.1)
- Choi KY, Vecchi GA, Wittenberg AT (2013) ENSO transition, duration and amplitude asymmetries: role of the nonlinear wind stress coupling in a conceptual model. *J Clim* doi:[10.1175/JCLI-D-13-00045.1](https://doi.org/10.1175/JCLI-D-13-00045.1)
- Cook B, Seager R, Miller R (2011) Atmospheric circulation anomalies during two persistent north american droughts: 1932–1939 and 1948–1957. *Clim Dyn* 36(11–12):2339–2355
- Deser C, Phillips AS, Tomas RA, Okumura YM, Alexander MA, Capotondi A, Scott JD, Kwon YO, Ohba M (2012) ENSO and Pacific decadal variability in the community climate system model version 4. *J Clim* 25(8):2622–2651. doi:[10.1175/JCLI-D-11-00301.1](https://doi.org/10.1175/JCLI-D-11-00301.1)
- DiNezio PN, Deser C (2014) Nonlinear controls on the persistence of La Niña. *J Clim* 27(19):7335–7355
- Fedorov AV, Harper SL, Philander SG, Winter B, Wittenberg A (2003) How predictable is El Niño? *Bull Am Meteorol Soc* 84(7):911–919
- Frauen C, Dommengot D (2010) El Niño and La Niña amplitude asymmetry caused by atmospheric feedbacks. *Geophys Res Lett* 37(18): doi:[10.1029/2010GL044444](https://doi.org/10.1029/2010GL044444)
- Hoerling M, Kumar A (2003) The Perfect Ocean for drought. *Science* 299(5607):691–694
- Hoerling M, Quan X-W, Eischeid J (2009) Distinct causes for two principal U.S. droughts of the 20th century. *Geophys Res Lett* 36:L19708. doi:[10.1029/2009GL039860](https://doi.org/10.1029/2009GL039860)
- Hoerling M, Kumar A, Dole R, Nielsen-Gammon JW, Eischeid J, Perlwitz J, Quan XW, Zhang T, Pegion P, Chen M (2013) Anatomy of an extreme event. *J Clim* 26(9):2811–2832
- Hu ZZ, Kumar A, Xue Y, Jha B (2013) Why were some La Niñas followed by another La Niña? *Clim Dyn* 42:1–14
- Hurrell JW, Holland MM, Gent PR et al. (2013) The community earth system model: a framework for collaborative research. *Bull Am Meteorol Soc* 94(9):1339–1360
- Izumo T, Vialard J, Lengaigne M, de Boyer Montegut C, Behera SK, Luo JJ, Cravatte S, Masson S, Yamagata T (2010) Influence of the state of the Indian Ocean Dipole on the following year's El Niño. *Nat Geosci* 3(3):168–172. doi:[10.1038/ngeo760](https://doi.org/10.1038/ngeo760)
- Kay JE, Deser C, Phillips A, Mai A, Hannay C, Strand G, Arblaster JM, Bates SC, Danabasoglu G, Edwards J, Holland M, Kushner P, Lamarque JF, Lawrence D, Lindsay K, Middleton A, Munoz E, Neale R, Oleson K, Polvani L, Vertenstein M (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 Am Meteorol Soc* 96(8):1333–1349
- Kessler WS (2002) Is ENSO a cycle or a series of events? *Geophys Res Lett* 29(23):40-1–40-4. doi:[10.1029/2002GL015924](https://doi.org/10.1029/2002GL015924)
- Kirtman B, Shukla J, Balmaseda M, Graham N, Penland C, Xue Y, Zebiak S (2001) Current status of ENSO forecast skill. *Tech. Rep. 56*, CLIVAR Working Group on Seasonal to Interannual Prediction
- Kirtman BP, Shukla J, Huang B, Zhu Z, Schneider EK (1997) Multiseasonal predictions with a coupled tropical ocean-global atmosphere system. *Mon Weather Rev* 125(5):789–808. doi:[10.1175/1520-0493\(1997\)125<0789:MPWACT>2.0.CO;2](https://doi.org/10.1175/1520-0493(1997)125<0789:MPWACT>2.0.CO;2)
- Kleeman R (2002) Measuring dynamical prediction utility using relative entropy. *J Atmos Sci* 59(13):2057–2072. doi:[10.1175/1520-0469\(2002\)059<2057:MDPUUR>2.0.CO;2](https://doi.org/10.1175/1520-0469(2002)059<2057:MDPUUR>2.0.CO;2)
- Kleeman R, Moore AM (1999) A new method for determining the reliability of dynamical ENSO predictions. *Mon Weather Rev*

- 127(5):694–705. doi:[10.1175/1520-0493\(1999\)127<0694:ANMFDT>2.0.CO;2](https://doi.org/10.1175/1520-0493(1999)127<0694:ANMFDT>2.0.CO;2)
- Kug JS, Kang IS (2006) Interactive feedback between ENSO and the Indian Ocean. *J Clim* 19(9):1784–1801. doi:[10.1175/JCLI3660.1](https://doi.org/10.1175/JCLI3660.1)
- Kug JS, Kirtman BP, Kang IS (2006) Interactive feedback between ENSO and the Indian Ocean in an interactive ensemble coupled model. *J Clim* 19(24):6371–6381. doi:[10.1175/JCLI3980.1](https://doi.org/10.1175/JCLI3980.1)
- Larson SM, Kirtman BP (2015) Revisiting ENSO coupled instability theory and SST error growth in a fully coupled model. *J Clim* 28(12):4724–4742. doi:[10.1175/JCLI-D-14-00731.1](https://doi.org/10.1175/JCLI-D-14-00731.1)
- Larson SM, Kirtman BP (2016) Drivers of coupled model ENSO error dynamics and the spring predictability barrier. *Climate Dyn* 1–14. doi:[10.1007/s00382-016-3290-5](https://doi.org/10.1007/s00382-016-3290-5)
- Latif M, Anderson D, Barnett T, Cane M, Kleeman R, Leetmaa A, O'Brien J, Rosati A, Schneider E (1998) A review of the predictability and prediction of ENSO. *J Geophys Res Oceans* 103(C7):14375–14393, doi:[10.1029/97JC03413](https://doi.org/10.1029/97JC03413)
- van Loon H, Shea DJ (1985) The Southern Oscillation. Part IV: The precursors South of 15° to the extremes of the oscillation. *Mon Weather Rev* 113(12):2063–2074. doi:[10.1175/1520-0493\(1985\)113<2063:TSOPIT>2.0.CO;2](https://doi.org/10.1175/1520-0493(1985)113<2063:TSOPIT>2.0.CO;2)
- McPhaden MJ, Yu X (1999) Equatorial waves and the 1997–98 El Niño. *Geophys Res Lett* 26(19):2961–2964
- McPhaden MJ, Zhang X (2009) Asymmetry in zonal phase propagation of ENSO sea surface temperature anomalies. *Geophys Res Lett* 36(13). doi:[10.1029/2009GL038774](https://doi.org/10.1029/2009GL038774)
- Neelin JD, Battisti DS, Hirst AC, Jin FF, Wakata Y, Yamagata T, Zebiak SE (1998) ENSO theory. *J Geophys Res Oceans* 103(C7):14261–14290
- Ohba M, Ueda H (2007) An Impact of SST Anomalies in the Indian Ocean in acceleration of the El Niño to La Niña transition. *J Meteorol Soc Jpn Ser II* 85(3):335–348. doi:[10.2151/jmsj.85.335](https://doi.org/10.2151/jmsj.85.335), <http://ci.nii.ac.jp/naid/110006318050/en/>
- Ohba M, Ueda H (2009) Role of nonlinear atmospheric response to SST on the asymmetric transition process of ENSO. *J Clim* 22(1):177–192. doi:[10.1175/2008JCLI2334.1](https://doi.org/10.1175/2008JCLI2334.1)
- Ohba M, Watanabe M (2011) Role of the Indo-Pacific interbasin coupling in predicting asymmetric ENSO transition and duration. *J Clim* 25(9):3321–3335. doi:[10.1175/JCLI-D-11-00409.1](https://doi.org/10.1175/JCLI-D-11-00409.1)
- Okumura YM (2013) Origins of tropical Pacific decadal variability: role of stochastic atmospheric forcing from the South Pacific. *J Clim* 26(24):9791–9796. doi:[10.1175/JCLI-D-13-00448.1](https://doi.org/10.1175/JCLI-D-13-00448.1)
- Okumura YM, Deser C (2010) Asymmetry in the duration of El Niño and La Niña. *J Clim* 23(21):5826–5843. doi:[10.1175/2010JCLI3592.1](https://doi.org/10.1175/2010JCLI3592.1)
- Okumura YM, Ohba M, Deser C, Ueda H (2011) A proposed mechanism for the asymmetric duration of El Niño and La Niña. *J Clim* 24(15):3822–3829. doi:[10.1175/2011JCLI3999.1](https://doi.org/10.1175/2011JCLI3999.1)
- Penland C, Sardeshmukh PD (1995) The optimal growth of tropical sea surface temperature anomalies. *J Clim* 8(8):1999–2024. doi:[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)
- Rayner NA, Parker DE, Horton EB, Folland CK, Alexander LV, Rowell DP, Kent EC, Kaplan A (2003) Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J Geophys Res Atmos* 108(D14). doi:[10.1029/2002JD002670](https://doi.org/10.1029/2002JD002670)
- Saji NH, Yamagata T (2003) Possible impacts of Indian Ocean Dipole mode events on global climate. *Clim Res* 25(2):151–169. doi:[10.3354/cr025151](https://doi.org/10.3354/cr025151). <http://www.int-res.com/abstracts/cr/v25/n2/p151-169/>
- Saji NH, Goswami BN, Vinayachandran PN, Yamagata T (1999) A dipole mode in the tropical Indian Ocean. *Nature* 401(6751):360–363
- Schubert SD, Suarez MJ, Pegion PJ, Koster RD, Bacmeister JT (2004) Causes of long-term drought in the U.S. Great Plains. *J Clim* 17(3):485–503
- Seager R (2007) The turn of the century North American drought: global context, dynamics, and past analogs. *J Clim* 20(22):5527–5552
- Seager R, Kushnir Y, Ting M, Cane M, Naik N, Miller J (2008) Would advance knowledge of 1930s SSTs have allowed prediction of the Dust Bowl drought? *J Clim* 21(13):3261–3281
- Thompson CJ, Battisti DS (2001) A linear stochastic dynamical model of ENSO part II: analysis. *J Clim* 14(4):445–466
- Trenberth KE, Shea DJ (1987) On the evolution of the Southern Oscillation. *Mon Weather Rev* 115(12):3078–3096. doi:[10.1175/1520-0493\(1987\)115<3078:OTEOTS>2.0.CO;2](https://doi.org/10.1175/1520-0493(1987)115<3078:OTEOTS>2.0.CO;2)
- Trenberth KE, Branstator GW, Arkin PA (1988) Origins of the 1988 North American drought. *Science* 242(4886):1640–1645
- Vimont DJ, Battisti DS, Hirst AC (2001) Footprinting: a seasonal connection between the tropics and mid-latitudes. *Geophys Res Lett* 28(20):3923–3926. doi:[10.1029/2001GL013435](https://doi.org/10.1029/2001GL013435)
- Vimont DJ, Wallace JM, Battisti DS (2003) The seasonal footprinting mechanism in the Pacific: implications for ENSO. *J Clim* 16(16):2668–2675
- Vimont DJ, Alexander M, Fontaine A (2009) Midlatitude excitation of tropical variability in the Pacific: the role of thermodynamic coupling and seasonality. *J Clim* 22(3):518–534. doi:[10.1175/2008JCLI2220.1](https://doi.org/10.1175/2008JCLI2220.1)
- Webster PJ, Moore AM, Loschnigg JP, Leben RR (1999) Coupled ocean-atmosphere dynamics in the Indian Ocean during 1997–98. *Nature* 401(6751):356–360
- Xie XP, Philander SG (1994) A coupled ocean-atmosphere model of relevance to the ITCZ in the eastern Pacific. *Tellus* 46(4):340–350
- Yamanaka G, Yasuda T, Fujii Y, Matsumoto S (2009) Rapid termination of the 2006 El Niño and its relation to the Indian Ocean. *Geophys Res Lett* 36(7). doi:[10.1029/2009GL037298](https://doi.org/10.1029/2009GL037298) (107702)
- Zebiak SE (1993) Air-sea interaction in the equatorial Atlantic region. *J Clim* 6(8):1567–1586. doi:[10.1175/1520-0442\(1993\)006<1567:AIITEA>2.0.CO;2](https://doi.org/10.1175/1520-0442(1993)006<1567:AIITEA>2.0.CO;2)
- Zhang H, Clement A, Nezio PD (2014) The South Pacific meridional mode: a mechanism for ENSO-like variability. *J Clim* 27(2):769–783. doi:[10.1175/JCLI-D-13-00082.1](https://doi.org/10.1175/JCLI-D-13-00082.1)