AGU PUBLICATIONS

Geophysical Research Letters

RESEARCH LETTER
10.1002/2017GL074904

Key Points:
- Thermocline discharge driven by strong El Niño requires several years of La Niña conditions to return to a neutral state
- Forecast initialization at the peak of a strong El Niño event leads to skillful predictions of subsequent 2 year La Niña event
- Returning La Nina conditions in 2018 have 60% probability based on a dynamical model and 80% probability based on an empirical model

Supporting Information:
- Supporting Information S1

Correspondence to:
P. N. DiNezio,
pdn@ig.utexas.edu

Citation:
DiNezio, P. N., Deser, C., Karspeck, A., Yeager, S., Okumura, Y., Danabasoglu, G., … Meehl, G. A. (2017). A 2 year forecast for a 60–80% chance of La Niña in 2017–2018. *Geophysical Research Letters*, *44*, 11,624–11,635. https://doi.org/10.1002/2017GL074904

Received 17 JUL 2017
Accepted 1 NOV 2017
Accepted article online 6 NOV 2017
Published online 27 NOV 2017

A 2 Year Forecast for a 60–80% Chance of La Niña in 2017–2018

Pedro N. DiNezio[1], Clara Deser[2], Alicia Karspeck[2], Stephen Yeager[2], Yuko Okumura[1], Gokhan Danabasoglu[2], Nan Rosenbloom[2], Julie Caron[2], and Gerald A. Meehl[2]

[1]Institute for Geophysics, Jackson School of Geosciences, University of Texas at Austin, Austin, TX, USA, [2]Climate and Global Dynamics Division, National Center for Atmospheric Research, Boulder, CO, USA

Abstract Historical observations show that one in two La Niña events last for two consecutive years. Despite their outsized impacts on drought, these 2 year La Niña are not predicted on a routine basis. Here we assess their predictability using retrospective forecasts performed with a climate model that simulates realistic multiyear events, as well as with an empirical model based on observed predictors. The skill of the retrospective forecasts allows us to make predictions for the upcoming 2017–2018 boreal winter starting from conditions in November 2015. These 2 year forecasts indicate that the return of La Niña is more likely than not, with a 60% probability based on the climate model and an 80% probability based on the empirical model; the likelihood of El Niño is less than 8% in both cases. These results demonstrate the feasibility of predictions of the duration of La Niña.

Plain Language Summary Historical observations show that cold La Niña events in the tropical Pacific often reintensify for a second year. Despite their outsized climate impacts throughout the world, these 2 year La Niña events are not routinely predicted. Our study demonstrates that long-term forecasts of these events are feasible. Moreover, we show increased likelihood of returning La Niña for next boreal winter, a result that is directly relevant for assessing climate risks throughout the world, for instance, over the southern tier of the U.S., where La Niña events create extreme seasonal heat and drought during winter and spring, and the Maritime Continent and Northern Australia, where it causes excess rainfall and flooding.

1. Introduction

Historical observations show that most El Niño events last for several seasons after which they quickly transition into La Niña, the cold phase of the El Niño–Southern Oscillation (ENSO) phenomenon in the tropical Pacific. In contrast, La Niña events often reintensify for a second year (Kessler, 2002; Okumura & Deser, 2010). These 2 year La Niña events are not routinely predicted because operational forecasts are limited to lead times of less than 9 months (Barnston et al., 2010). Furthermore, these operational forecasts have proven to be unreliable at shorter lead times, for instance, during the most recent 2 year La Niña of 2010–2012. In particular, operational predictions initialized in June 2011 showed a range of positive and negative ENSO states in the following boreal winter 2011/2012, providing an unreliable picture of the evolution of La Niña that year (ENSO forecasts initialized in June 2011 available online at http://iri.columbia.edu/our-expertise/climate/forecasts/enso/archive/201106/SST_table.html). Indeed, La Niña ultimately reemerged during late 2011 to early 2012, favoring the occurrence of one of the most persistent droughts over the southeastern U.S. and Texas (Hoerling et al., 2013). Despite the weaker magnitude of the tropical cooling during the second year, the rainfall deficits over the southern U.S. are, on average, comparable in magnitude to the first year and extend further north, potentially exacerbating droughts associated with 2 year La Niña (Okumura, DiNezio, et al., 2017).

The *onset* of La Niña is highly predictable since it is governed by slowly propagating variations in the depth of the thermocline, the boundary separating the warm surface waters from deeper and colder waters. This process is initiated by wind patterns associated with the preceding El Niño, which drive a *shoaling* of the equatorial thermocline across the Pacific. These large-scale anomalies are typically characterized as a reduction in upper ocean heat content or "discharge" of equatorial heat content. The thermocline shoaling maximizes a few seasons after the peak of El Niño and is highly effective at cooling the ocean's surface over the central equatorial Pacific via increased entrainment and upwelling of cold subsurface waters. This lead-lag relationship is the basis of well-established theories explaining the onset and decay of ENSO events (Jin, 1997;

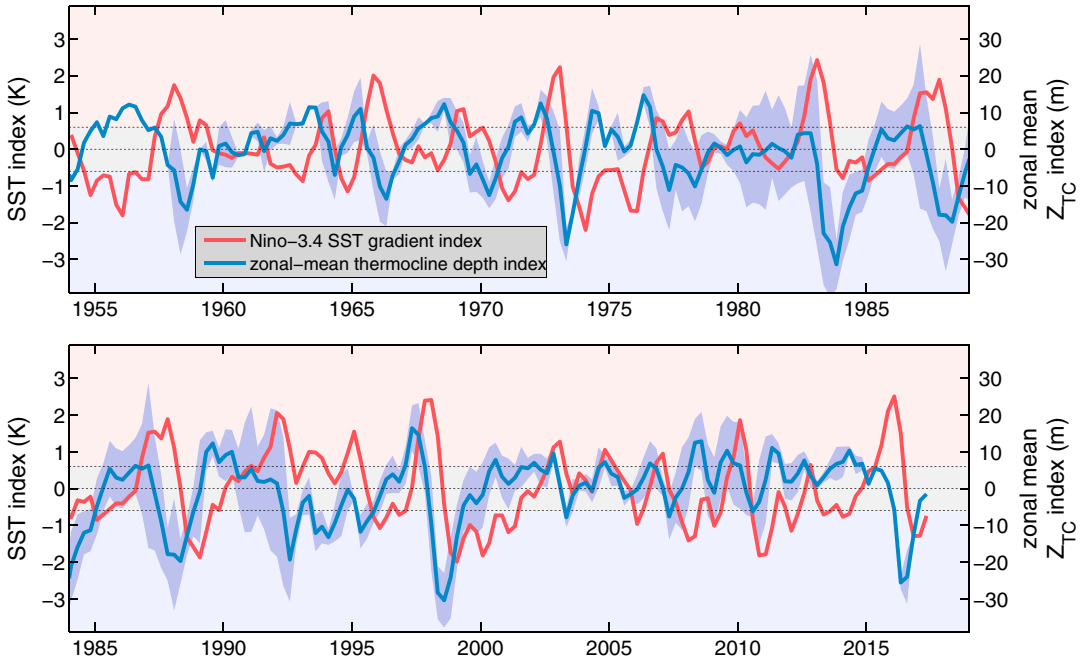

Figure 1. Observed ENSO events and associated thermocline depth anomalies. (a) Niño-3.4 sea surface temperature (SST) gradient index and (b) zonally averaged thermocline depth index from multiple observational data sets. SST data are from the ERSST3b reconstruction, and thermocline depth is computed from potential temperature data from the ORAS-4, GECCO2, and GODAS reanalyses and in situ observations from the TAO/TRITON moored array.

Meinen & McPhaden, 2000) and explains the consistent and predictable occurrence of La Niña after El Niño. Conversely, these theories propose that wind patterns associated with La Niña would drive a delayed *deepening* of the thermocline, or thermocline "recharge," that would warm the ocean's surface, terminating La Niña, and potentially triggering a subsequent El Niño. This response appears to be ineffective, since La Niña events rarely transition directly into an El Niño (Kessler, 2002). This breakdown of ENSO's oscillatory behavior limits our ability to predict La Niña's *termination* and hence its duration.

Recent studies suggest that the magnitude and persistence of thermal anomalies below the ocean's surface could aid in the prediction of La Niña's duration (DiNezio et al., 2017). According to that study, forecasts initialized at the peak of a strong El Niño reliably predict the occurrence of a subsequent 2 year La Niña in a coupled model. This occurs because strong El Niño events drive a thermocline shoaling that is disproportionally large, thus requiring several years of La Niña conditions to return the system to a neutral ENSO state. This idea is consistent with the observed evolution of the El Niño events of 1982 and 1997—the strongest on record at the time. Both events were followed by thermocline shoaling of more than 30 m averaged across the equatorial Pacific (Figure 1, blue curve). We posit that these thermocline depth anomalies—also the largest on record—played a dominant role on the multiyear persistence of the subsequent La Niña (Figure 1, red curves). A thermocline shoaling of more than 20 m was also observed during the spring and summer of 2016, after the peak of the recent record-breaking El Niño (Figure 1), suggesting that this year's La Niña could evolve according to similar dynamics.

Here we present retrospective forecasts of the multiyear evolution of past El Niño events that resulted in subsequent 2 year La Niña. Furthermore, we present predictions for the upcoming boreal winter of 2017–2018 that are initialized at the peak of the most recent El Niño, that is, 2 years in advance. Our forecasts are performed with an empirical statistical model and a comprehensive coupled climate model. The coupled climate model forecasts were performed with the Community Earth System Model Version 1 (CESM1)—a model capable of simulating very realistic El Niño–Southern Oscillation (ENSO) events, particularly the asymmetries in amplitude and duration between El Niño and La Niña events (DiNezio et al., 2017). The empirical model is based on a statistical analysis of multiple observational data sets combined with probabilities derived from a long control simulation performed with CESM1.

We first present the statistical analysis of precursors of 2 year La Niña from observations and the CESM1 control. We then assess retrospective skill of the CESM1 at predicting historical 2 year La Niña over the 1954–2014 period. To conclude, we use the forecasts initialized in November 2015 to estimate the probability of a 2 year La Niña following the most recent El Niño. The target season for these forecasts is November-December-January of 2017/2018, a full 2 years after initialization.

2. Data and Methods

2.1. Observational Data

We use the Extended Reconstructed Sea Surface Temperature version 3b (ERSST3b) (Smith et al., 2008) to compute sea surface temperature anomalies (SSTA) for the 1954 to present period. SSTA are computed relative to the monthly mean seasonal cycle over the 1986–2015 period, the same baseline used in NOAA's operational definition of ENSO events (http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_change.shtml). We use ocean subsurface temperature data to estimate the depth of the thermocline, Z_{TC}, as the location of the maximum vertical temperature gradient (DiNezio et al., 2017). The following ocean reanalysis data are used for different periods: ORA-S4 (Balmaseda et al., 2013) (1958–2014), GECCO2 (Kohl & Stammer, 2008) (1954–2010), GODAS (Huang et al., 2010) (1982–2016), and direct observations from the TAO/TRITON array (1986–2016). The monthly mean climatology over the common 1986–2010 period is removed from each data set to compute Z_{TC} anomalies, Z'_{TC}. Sea level pressure (SLP) data used to asses the realism of ENSO teleconnections are from the twentieth Century Reanalysis (Compo et al., 2011).

2.2. ENSO Indices

We track ENSO events using an index defined as the difference between the SSTA averaged over the Nino-3.4 region in the central equatorial Pacific (170°W–120°W, 5°S–5°N) minus SSTA averaged over the western equatorial Pacific (120°E–150°E, 5°S–5°N). This "gradient" Niño-3.4 SST index (hereinafter Niño-3.4g) captures the weaker or stronger zonal SST gradient associated with El Niño and La Niña and is highly correlated ($r = 0.98$) with the conventional Niño-3.4 SST index (Figure S1a in the supporting information). We use this index because it is less sensitive to long-term warming than the conventional Niño-3.4 SST index. The western equatorial Pacific region (120°W–170°W, 5°S–5°N) shows a consistent warming trend of 0.81 K (63 year)$^{-1}$ (Figure S1b). Referencing Niño-3.4 SSTA relative to this region isolates ENSO-related variability from the long-term warming trends, which can be of considerable magnitude and highly uncertain (Cane et al., 1997; Cravatte et al., 2009; Deser et al., 2010; Latif et al., 1997). As a result of this normalization, the observed Niño-3.4 SST gradient index shows a negligible trend over the 1954-to-present period; in contrast to a warming trend of 0.44 K (63 year)$^{-1}$ using the conventional definition (Figure S1a, red and blue curves). We use a threshold of 0.6 K to define El Niño and La Niña commensurate with the larger standard deviation of our Niño-3.4g SST index (Table S1). The thermocline discharge and recharge are estimated using an index defined as the area-averaged Z'_{TC} across the equatorial Pacific (140°E–80°W, 5°S–5°N). This zonal-mean Z'_{TC} index (hereinafter $\bar{Z}'_{TC}$) captures the timing and magnitude of the thermocline depth variations in all four data sets, particularly the large negative departures associated with the thermocline discharge following El Niño events (Figure S2c). We detrend and then average the $\bar{Z}'_{TC}$ index from each data set to obtain a multidata set mean (Figure 1, blue curve). Observational uncertainty is computed as the minimum-maximum range of the overlapping data sets (Figure 1, blue shading).

2.3. Empirical Model

We developed an empirical model for predicting 2 year La Niña combining observed predictors with probabilities from a 1,800 year-long control simulation performed with CESM1 (hereinafter CESM-CTL). CESM-CTL simulates realistic characteristics of ENSO events, including their spatial patterns and global teleconnections (Figure S2), amplitude (Table S1), and importantly the same frequency of 2 year La Niña events as observed. First, we selected 251 El Niño events from the CESM-CTL and computed a set of predictors following the methodology of DiNezio et al. (2017). The 2 year La Niña predictand (Niño $-$ 3.4g^{NDJ+2}) is defined as the magnitude of the Niño-3.4g index 2 years after the peak of an El Niño event. A "peak El Niño" predictor is defined as the Niño-3.4g SST index during the peak November-December-January (NDJ) season (Niño $-$ 3.4$g^{NDJ\,0}$) of each simulated El Niño event. A discharge predictor is defined as the $\bar{Z}'_{TC}$ index averaged from February to July following the peak of an El Niño event ($\bar{Z}'^{FMAMJJ+1}_{TC}$). We binned the 251 events according to the values of the "El Niño peak" and "thermocline discharge" predictors using overlapping bins of 1 K and 10 m width, respectively. The cumulative density function (CDF) of the Niño $-$ 3.4g^{NDJ+2} predictand was estimated

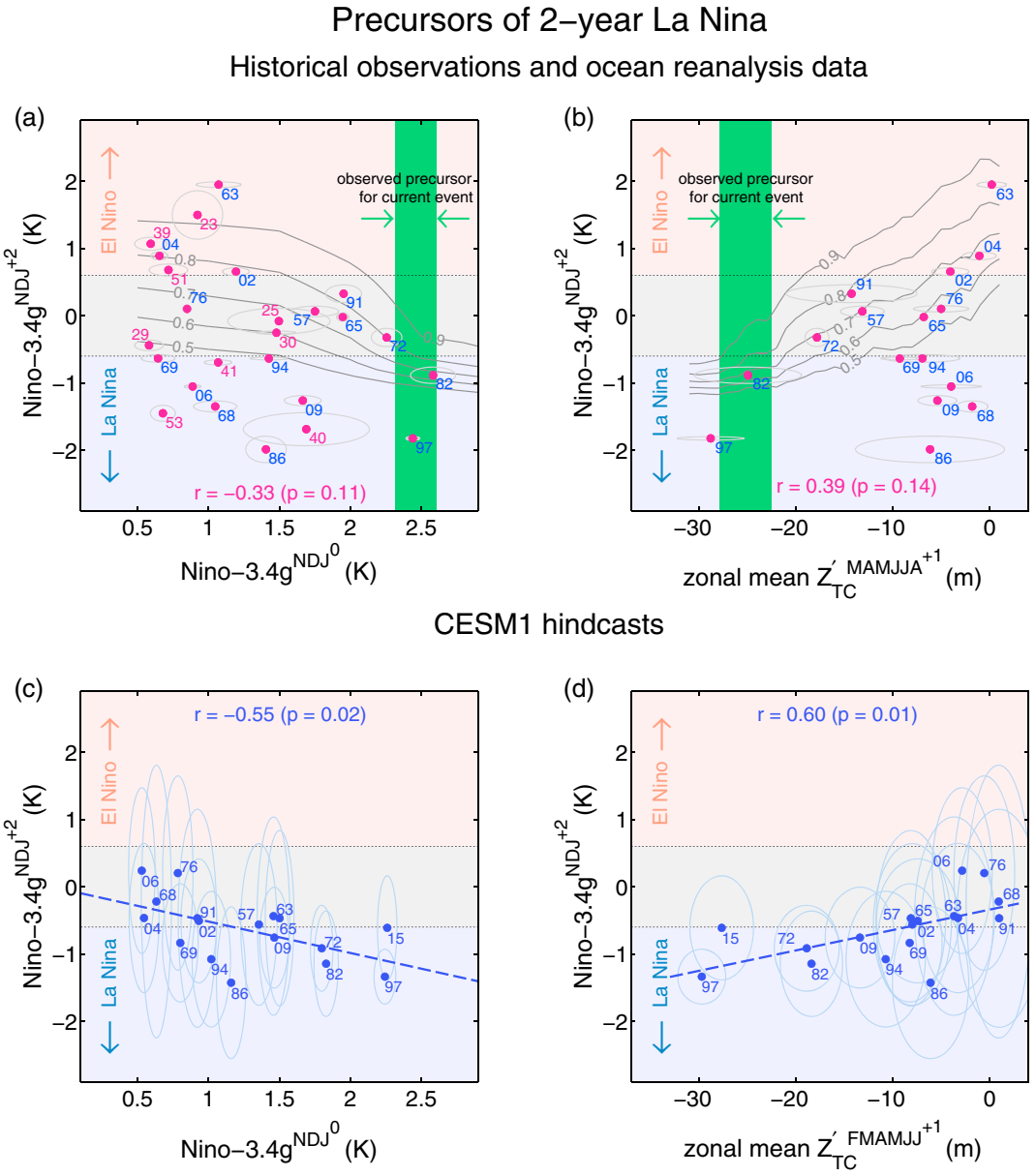

Figure 2. Statistical precursors of observed 2 year La Niña events in nature and in the CESM-DP-LE. Scatterplots showing correlations between the observed and predicted Niño-3.4g index in the year following the first peak of La Niña (Niño-3.4g^{NDJ+2}, *y* axis) and the following two predictors (*x* axis): (a and c) the Niño-3.4g index at the peak of the El Niño event preceding La Niña and (b and d) the thermocline discharge during the transition between El Niño to La Niña. Returning La Niña conditions for a second year, that is, a 2 year La Niña, are given by Niño-3.4g^{NDJ+2} values (*y* axis) less than −0.6 K (blue shading). El Niño conditions during the second year, that is, a 1 year La Niña are given by Niño-3.4g^{NDJ+2} values larger than 0.6 K (red shading). Each dot is labeled according to the year of the initial El Niño event (based on December). Figures 2a and 2b, and 2c and 2d show precursors derived from observational data and the CESM-DP-LE, respectively. Blue dots in Figures 2a and 2b identify events for which there is a CESM-DP-LE hindcast. Gray contours show the cumulative probability of Niño-3.4g^{NDJ+2} predictor as a function of the precursors computed from the uninitialized CESM1 control simulation. Ellipses indicate observational uncertainty for observed precursors, computed as the min-max range among the three data sets, and ensemble spread for hindcast precursors, computed as the ensemble standard deviation. For the CESM-DP-LE, the discharge predictor is computed from the ensemble-mean $\bar{Z}'_{TC}$ index averaged from February to July of the year after initialization, which we label $\bar{Z}'^{FMAMJJ+1}_{TC}$.

with MATLAB's ksdensity code using data from each bin. This resulted in a set of CDFs for different values of the Niño − 3.4$g^{\mathrm{NDJ}\,0}$ and $\bar{Z}'^{\mathrm{MAMJJA+1}}_{\mathrm{TC}}$ predictors. This dependence is plotted as contours in Figures 2a and 2b.

Last, we computed observed predictors using all available observational products following the same definitions as for the CESM-CTL. An averaging interval March to August is used for the observed thermocline discharge predictor ($\bar{Z}'^{\mathrm{MAMJJA+1}}_{\mathrm{TC}}$) based on timing of the peak of the observed discharge. Analysis of Niño-3.4$g^{\mathrm{NDJ}\,0}$ includes observed events since 1920, the period when all three data sets agree on the amplitude of El Niño events. Observational uncertainty is computed as the minimum-maximum range of the overlapping datasets for both the "El Niño peak" and "thermocline discharge" predictors, as well as for the 2 year La Niña predictand, shown in Figures 2a and 2b as ellipses. For each observed event, the probability of 2 year La Niña is given by value of the CDF for Niño-3.4$g^{\mathrm{NDJ+2}}$ = −0.6 K and the magnitude of the observed predictor.

2.4. Dynamical Model

The CESM1 forecasts belong to the Decadal Prediction Large Ensemble (hereinafter CESM-DP-LE), a suite of ensemble forecasts run on the Cheyenne platform at the National Center for Atmospheric Research (NCAR)-Wyoming Supercomputer Center (Computational and Information Systems Laboratory, 2017). The suite consist of 61 ensemble hindcasts (retrospective forecasts) of global climate initialized in every 1 November from years 1954 to 2014, and an additional ensemble forecast initialized in 1 November 2015. We used the former for assessing the retrospective predictability of 2 year La Niña events since 1954, and the latter for predicting the evolution of the current La Niña. Ocean and sea ice initial conditions were obtained from a simulation performed with POP2 and CICE—the CESM ocean and sea ice models—forced at the surface with atmospheric reanalysis data defined by the Coordinated Ocean-ice Reference Experiments (CORE) phase II as in Yeager et al. (2015). This CESM-CORE simulation captures key ocean dynamical processes involved in the onset and decay of 2 year La Niña in agreement with the ORAS-S4 reanalysis (Figure S3). The land and atmosphere were initialized from an uninitialized historical simulation.

Ensembles were generated for each initial date perturbing each member's atmospheric initial conditions with a unique round-off level difference (order of 10^{14} K) to the air temperature field. Each ensemble consists of 40 members run for 10 years providing forecasts with sufficient length to explore the evolution of 2 year La Niña. We used the 1954- to 2014-initialized hindcasts also to compute and remove forecast drift following the methodology of International CLIVAR Project Office (2011) extensively used in seasonal and decadal climate forecasts (Hazeleger et al., 2013; Meehl & Teng, 2012; Yeager et al., 2012). Further details are available in section 2 of the supporting information. This approach removes the drift that is common to all ensembles, isolating internally generated and externally forced climate anomalies. The latter can be quite pronounced over a 62 year period such as 1954–2015 and thus introduce biases in the amplitude of ENSO events. For instance, the ensemble-mean SSTA between earlier and recent events reveals a pronounced warming trend over the tropical Pacific in the CESM-DP-LE (Figures S4 and S5, left). A similar warming trend is evident in the observed SSTA computed relative to the same 1964–2014 baseline (Figures S4 and S5, right). Previous studies have addressed this issue by removing the forced component from the hindcasts (e.g., van Oldenborgh et al., 2012). We circumvent this issue by using the Niño-3.4g SST index defined in section 2.2.

Lastly, retrospective verification is required to demonstrate the skill of a forecast system prior to its use for actual predictions (Barnston et al., 2010; Ham et al., 2014; Kirtman et al., 2001; Wang et al., 2010; Tang et al., 2005; Tippett et al., 2011). Our analysis shows that that the CESM-DP-LE is particularly skillful at predicting the duration of historical La Niña with 2 and 1 year lead times when initialized from the peak of the preceding El Niño and from the first peak of La Niña, respectively (Figures S7 and S8). Further details available in section 3 of the SI. To conclude, we use the 2015-initialized forecasts to predict the evolution of this year's event and estimate the likelihood of returning La Niña during the target season of November-December-January (NDJ) of 2017/2018.

2.5. ENSO Prediction Plumes

We computed the mean Niño-3.4g and $\bar{Z}'_{\mathrm{TC}}$ indices among all members of each ensemble to quantify the predictable component of each forecast. We define a second year La Niña when the Niño-3.4g SST index is less than −0.60 K, during the November-December-January (NDJ) season, 2 years after the forecast start date (NDJ+2). Probability of 2 year La Niña are computed as the fraction of ensemble members showing Niño-3.4g SST index under the −0.60 K threshold during NDJ+2. A composite of events starting from strong and moderate El Niño shows that on average the CESM-DP-LE hindcasts capture the evolution of the subsequent La Niña events in terms of timing and amplitude (Figures S6a and S6b).

3. Results

3.1. Observed Precursors

Analysis of precursors of historical ENSO events supports a link between the occurrence of 2 year La Niña and the strength of the preceding El Niño as proposed by previous theoretical and modeling studies (DiNezio & Deser, 2014; DiNezio et al., 2017; Okumura, Sun, et al., 2017). The peak amplitude of El Niño events, measured by the gradient Niño-3.4g SST index during the NDJ season (Niño-3.4g$^{NDJ\,0}$), is anticorrelated ($r = -0.33$, $p = 0.11$) with the magnitude of the Niño-3.4g index 2 years later (Niño-3.4g^{NDJ+2}). A correlation of this magnitude would not warrant a skillful prediction; however, the scatterplot between these quantities suggests that the likelihood of 2 year La Niña increases with the strength of the preceding El Niño (Figure 2a). For instance, the observed SST data show that a returning El Niño after La Niña would be highly unlikely when the preceding El Niño has a peak amplitude larger than 1.5 K (Figure 2a, Niño-3.4g$^{NDJ+2} > 1.5$ K). Conversely, 2 year La Niña are more likely to occur following strong El Niño events since two out of the three strongest El Niño events were followed by La Niña conditions 2 years later and none by a returning El Niño. The remaining event, the El Niño of 1972/1973, showed neutral ENSO conditions during the second year.

Our empirical model based on the CESM-CTL shows increasing likelihood of 2 year La Niña as a function of the strength of the preceding El Niño (Figure 2a, gray contours). According to this model, La Niña events preceded by El Niño amplitudes within the range observed during NDJ 2015 (Figure 2a, green shading) would have a 65% to 85% chance of lasting 2 years (An 80% chance for the averaged Niño-3.4g$^{NDJ\,0}$ predictor among the three SST data sets). The magnitude of the thermocline discharge can also be used as a predictor for 2 year La Niña. The observed $\bar{Z}'^{MAMJJA+1}_{TC}$ predictor is also correlated ($r = 0.39$, $p = 0.14$) with the amplitude of the Niño-3.4g index in NDJ of the following year. Furthermore, the probabilities from the CESM-CTL (Figure 2b, gray contours) show that La Niña events preceded by a discharge predictor within the range observed during 2016 (Figure 2b, green shading) would have a 80% to 90% chance of becoming a 2 year event.

3.2. Retrospective Predictions

The CESM-DP-LE shows a similar link between predictions of past 2 year La Niña events and the peak amplitude and thermocline discharge of the preceding El Niño as in the CESM-CTL. Hindcasts initialized during El Niño Novembers show ensemble-mean Niño-3.4g^{NDJ+2} values 2 years later that are anticorrelated with the magnitude of the ensemble-mean Niño-3.4g$^{NDJ\,0}$ predictor ($r = -0.58$, $p = 0.01$). The ensemble-mean Niño-3.4g^{NDJ+2} predictand is also highly correlated ($r = 0.60$, $p = 0.01$) with the discharge predictor computed from the ensemble-mean $\bar{Z}'_{TC}$ index averaged from February to July of the year after initialization. The hindcasts show higher correlation than the observational data because the ensemble-mean predictors average out random variability. This suggests that in the CESM-DP-LE, the magnitude of the thermocline discharge determines a large fraction of the predictable component of 2 year La Niña. This result also suggest that many observed 2 year La Niña preceded by weak El Niño/discharge were influenced by unpredictable factors (i.e., atmospheric noise). Moreover, the CESM-DP-LE supports our conclusion that the likelihood of 2 year La Niña increases after strong El Niño events. The hindcasts initialized in November of 1972, 1982, and 1997, at the peak of the strongest El Niño events in our period of study, show highly predictable negative SST anomalies during the NDJ season 2 years later, as seen by the ensemble mean (Figures 2c and 2d, circles) and ensemble spread (Figures 2c and 2d, ellipses) of the Niño-3.4g SST index. Forecasts initialized from states characterized by weaker El Niño and associated discharge lead to less predictable 2 year La Niña due to a smaller ensemble-mean signal, and also due to larger ensemble spread.

3.3. Predictions of the Current Event

The CESM-DP-LE initialized in November 2015 predicts an ensemble-mean evolution similar to previous 2 year La Niña events preceded by strong or moderate El Niño events (Figures 3a and 3b). In particular, our model predicted the observed onset of La Niña during boreal fall of 2016 and its subsequent peak near the end of 2016, but more importantly, it also predicts continued persistence of La Niña conditions through the upcoming boreal winter (Figure 3a). The ensemble-mean Niño-3.4g SST index remains under the −0.6 K threshold used to define La Niña conditions throughout our target forecast season NDJ of 2017/2018 (Figure 3a, solid red curve) as well as the preceding August-September-October season, thus satisfying operational criteria used to classify ENSO events (NOAA's Climate Prediction Center defines ENSO events when Niño-3.4 SST anomalies exceed a 0.5 K threshold for five consecutive 3 month running mean seasons). The 2015-initialized forecast also shows zonally averaged thermocline depth anomalies evolving with the same magnitude and timing as observed (Figure 3b, blue versus green curves), with thermocline discharge peaking during February to July of 2016, shown by $\bar{Z}'_{TC}$ values of about −27 m, slightly shallower than observed.

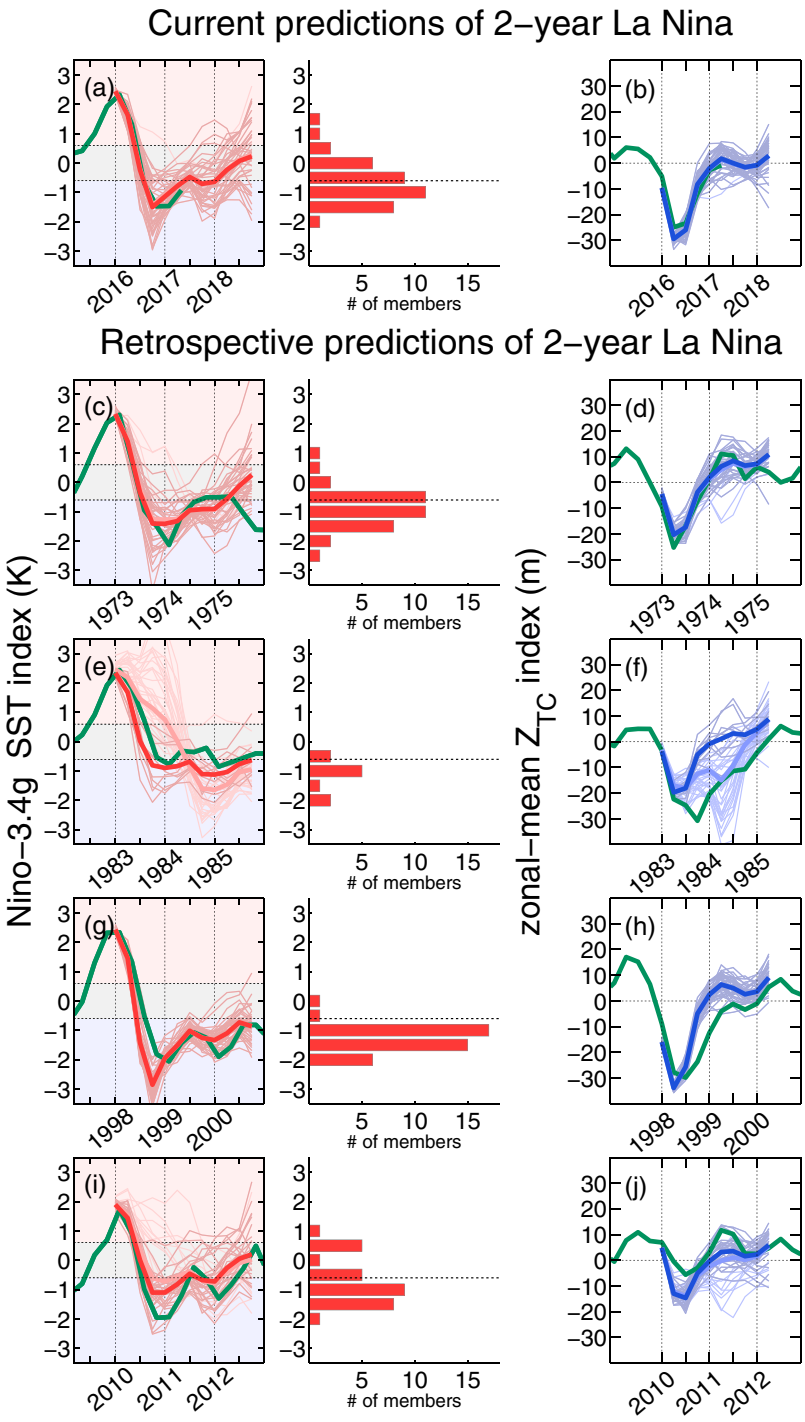

Figure 3. Current and retrospective predictions of La Niña events performed with the CESM1. Predicted and observed (a, c, e, g, and i) Nino-3.4 SST index and (b, d, f, h, and j) zonally averaged thermocline depth, Z_{TC}, index, of ENSO events initiated in 1972, 1982, 1997, 2009, and 2015. Histograms show the distribution of predicted Niño-3.4g SST values for the target season of November-December-January 2 years after initializations (NDJ+2). Figures 3a and 3b show the prediction plumes from the forecast initialized in November 2015, at the peak of the El Niño preceding the current La Niña. Figures 3c–3j show retrospective forecasts of previous events initiated by strong and moderate El Niño. Solid green curves are the observed indices. Refer to section 2 for details on how these indices are computed from multiple available observational data sets. Red and blue thick curves identify the predicted ensemble-mean indices. Thin curves identify the indices predicted by each individual ensemble member. Pink and lilac curves correspond to ensemble members that show returning El Niño conditions and are not considered in the forecasts (see supporting text section 3.4 for more details on how these members are identified).

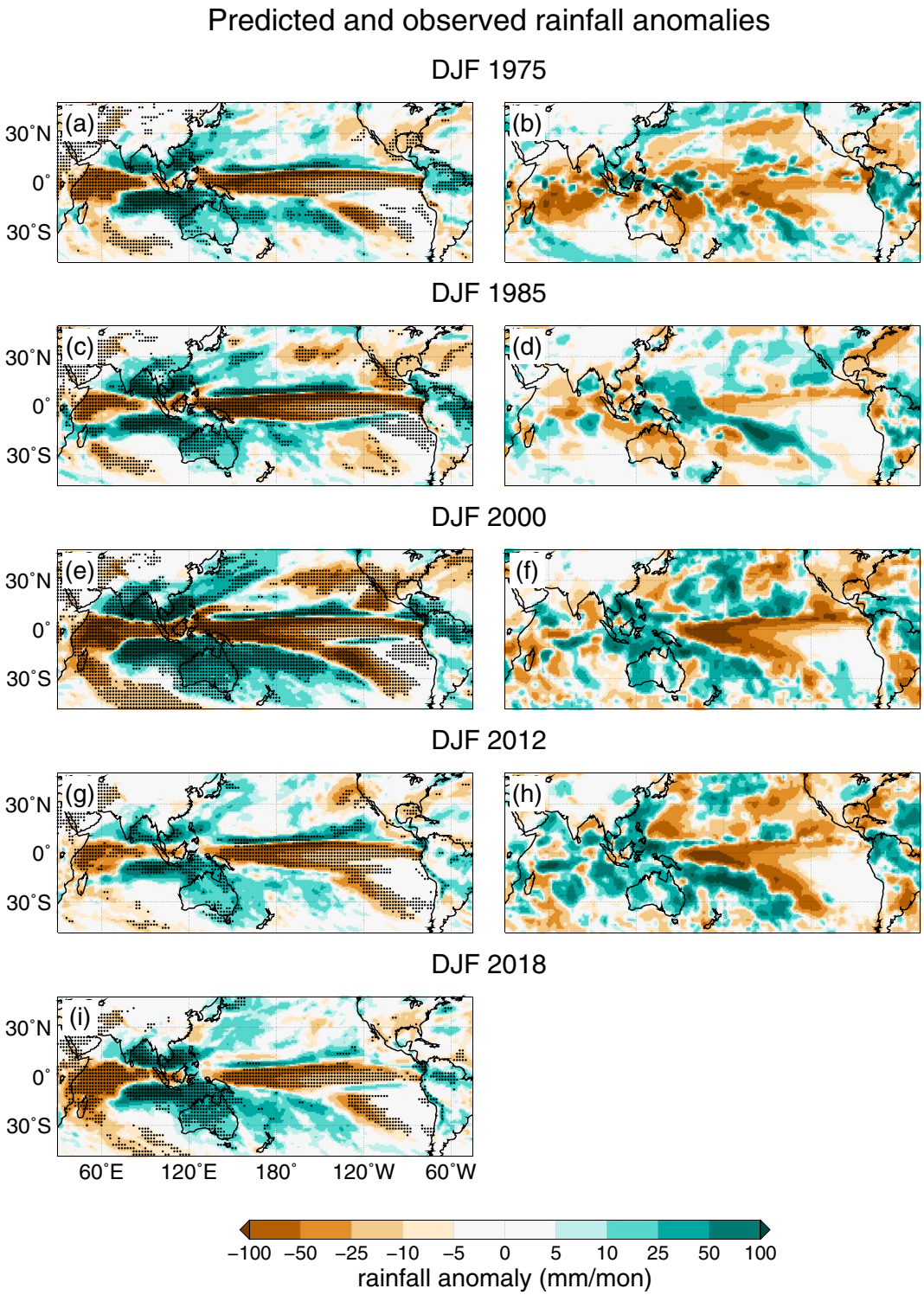

Figure 4. Predicted and observed rainfall anomalies during the second year peak of La Niña events. (a, c, e, g, and i) Ensemble-mean and (b, d, f, and h) observed rainfall anomalies corresponding to the target December-January-February forecast season of selected 2 year La Niña that were preceded by strong or moderate El Niño events. Stippling indicates locations where the ensemble signal to noise ratio is larger than 0.5. Observed rainfall anomalies prior to and after 1979 are from ERA40 reanalysis data (Uppala et al., 2005) and GPCPv2 satellite observations (Adler et al., 2003), respectively.

Is the evolution of the current event similar to that of the 1997 record-breaking El Niño? Both events show the largest predicted thermocline discharge among our hindcasts (Figures 3b, 3d, 3f, 3h, and 3j and also 2d) consistent with the record-breaking strength of the preceding El Niño. And both cases show the onset of La Niña one year after the peak of El Niño, along with its return for a second year (Figure 3a). However, these two events show striking differences in the consistency of the predictions. The *onset* of La Niña was predicted by all 40 members of the 1997-initialized ensemble (Figure 3h, pink curves), but only by 33 members of the 2015-initialized ensemble (Figure 3a, thin red curves). Twenty-four members predict the *return* of La Niña for next winter (i.e., a 60% chance), in contrast to 38 members predicting La Niña conditions for NDJ of 2000 (a 95% chance of 2 year La Niña following the El Niño of 1997). Despite the strength of the preceding El Niño, the forecasts of the current event show one of the weakest ensemble-mean La Niña predictions for the second year among the events analyzed (Figures 3a, 3c, 3e, 3g, and 3 and also 2d). However, both sets of forecasts provide a low probability of a returning El Niño during the second year (7.5% for the current event and 0% for the 1997 event).

Several processes could explain the distinctions between the likelihood of a 2 year La Niña obtained from the 1997- and 2015-initialized forecasts. First, CESM1 predicts an ensemble-mean Niño-3.4g SST index of −0.61 K for the target season of NDJ 2017, compared to the much larger value of −1.33 K for NDJ of 1999. This difference cannot be explained by differences in the thermocline discharge because both cases show the same peak amplitude. CESM1 predicted not only a much stronger first year La Niña relative to observed for 1998 but also a much lower ensemble spread (Figure 3g). The differences with the 2015-initialized forecasts become conspicuous by NDJ of 2017/2018 when the ensemble spread, measured by the standard deviation of the Niño-3.4g SST index, is 0.77 K compared with 0.41 K for the 1997-initialized forecasts. The different spread between these cases could also be related to the strength of La Niña's first year peak, which would suppress uncoupled atmospheric variability capable of causing forecast spread.

The CESM-DP-LE also predicted a 2 year La Niña following the El Niño events of 1972, 1983, and 2009 (Figures 3a, 3c, 3e, 3g, and 3i). Among these events, the 1972- and 1982-initialized hindcasts could be closer analogs to the current predictions since they show similar ensemble mean and spread as the 2015-initialized forecast (Figure 2d). However, these events show a weaker peak discharge in the CESM-DP-LE (Figure 2d). It is possible that our definition of discharge, based on the depth of the maximum vertical temperature gradient, could be overestimating the ENSO-driven thermocline depth variations. For instance, the thermocline could be experiencing long-term shoaling in response to global warming (DiNezio et al., 2009), thus biasing our definition of the $\bar{Z}'_{TC}$ index. This effect would be more pronounced in the recent hindcasts where is difficult to isolate externally forced trends from model drift. Note that one of these events, the El Niño of 1972, led to a 1 year La Niña in nature, followed by weak cold conditions during the target NDJ season of 1974/1975 that did not meet the criterion for La Niña per se (Figure 3c, green curve). This is consistent with the predictions for the current event, which has a 40% chance of being limited to 1 year.

3.4. Predicted Impacts

The CESM-DP-LE predicted typical La Niña rainfall anomalies for the 2 year events preceded by strong or moderate El Niño events (Figure 4). In particular, the predicted rainfall anomalies for the second year peak in boreal winter show drying over the central equatorial Pacific and wetter conditions over the Maritime Continent (Figures 4a, 4c, 4e, 4g, and 4i). These anomalies show a high ensemble signal to noise (Figure 4, stippling) suggesting reliable rainfall predictions over the tropics. Observations show similar patterns, albeit with more spatial variability consistent with the more unpredictable nature of rainfall. Remote rainfall impacts associated with ENSO teleconnections, such as drying over the southeastern U.S., show signal to noise ratios smaller than 0.5. The one exception is boreal winter of 1999/2000, for which the CESM-DP-LE predictions show signal to noise larger than 0.5 over the southern tier of the United States (Figure 4e, stippling). The rainfall predictions for the upcoming boreal winter (Figure 4i) show large signal over the tropics, but not over North America, highlighting the more unpredictable nature of extratropical ENSO teleconnections (Deser et al., 2017).

4. Conclusions

The results presented here demonstrate the feasibility of skillful forecasts of 2 year La Niña events at 2 year lead times. These forecasts are skillful when initialized at the peak of strong El Niño events because more than 1 year is required for the thermocline to return to neutral conditions after the large initial discharge.

Our statistical analysis shows that the magnitude of the peak thermocline discharge controls the predictable signal for up to 2 years. The predictability of La Niña duration is consistent with processes governed by persistence, in which events driven by large initial perturbations will take longer to decay. The forecasts remain skillful at 2 year lead times because the ensemble-mean signal has sufficiently large amplitude to overcome the growing forecast spread. This conclusion is consistent with previous studies showing that ensemble spread is not a good indicator of skill (Kirtman et al., 1997), a well-known difference with weather prediction, where low ensemble spread is critical for skillful forecasts. The striking similarity in the evolution of the thermocline depth anomalies in the hindcasts and observations suggest that the skillful prediction of past 2 year La Niña arises from CESM1's realistic simulation of the dynamics of these events, which arises from nonlinearities in dynamical feedbacks (Choi et al., 2013; DiNezio & Deser, 2014; DiNezio et al., 2017; Okumura et al., 2011).

We estimate the probability of La Niña returning next boreal winter at between 60%, based on the CESM-DP-LE forecast initialized in November 2015, and 80%, based on a statistical analysis of observed and simulated predictors. We predict a rather weak La Niña based on the magnitude of the ensemble-mean Niño-3.4g SST index (Figure 3a, red curve); however, its probability distribution is skewed toward more moderate La Niña conditions (Figure 3a, histogram). While the La Niña of 2016–2017 was considered weak by the conventional definition (Figure S1a, red curve), it was not in terms of the Niño-3.4g SST index (Figure S1a, blue curve). Furthermore, the predicted rainfall patterns for the upcoming boreal winter are not much different from those of previous La Niña (Figures 4a, 4c, 4e, 4g, and 4i). This occurs because rainfall patterns are sensitive to SSTA gradients, and less to absolute SSTA, which increase with global warming. Therefore, the rainfall anomalies more clearly show that the predicted SSTA for next winter will have La Niña-like impacts of considerable amplitude over the tropical Indian and Pacific oceans. To date, our ensemble-mean predictions have tracked the observed evolution of the Niño-3.4g SST index (Figure 3a, blue and red curves). However, this does not necessarily indicate that the event will continue to evolve according to the ensemble-mean predictions, as models and observations show that the magnitude of the first and second year peaks are not correlated (DiNezio & Deser, 2014; DiNezio et al., 2017; Okumura, Sun, et al., 2017).

In addition, our forecasts show large spread, suggesting that the future evolution until our target season of NDJ 2017/2018 could still be affected by unpredictable processes in the Pacific. Idealized predictions performed with CESM1 show that random wind variability in the western Pacific and the state of the Southern Oscillation have the largest short-term impact on forecast spread (DiNezio et al., 2017). Observations from the TAO/TRITON array captured easterly wind anomalies during February to April of 2017, which according to these idealized forecasts will be more conducive to a 2 year La Niña. The Southern Oscillation index exhibited values closer to neutral throughout the past austral winter and has turned positive since August of 2017 a state that would also be conducive for returning La Niña conditions. Future studies should assess the role of these stochastic processes in past 2 year La Niña.

The empirical model complements the dynamical forecasts because it shows that despite the large magnitude of the El Niño peak and thermocline discharge predictors, there is a nonnegligible chance (20%) that La Niña will not return next boreal winter. Conversely, the dynamical model provides a lower bound to the likelihood of 2 year La Niña, which is comparable to the historical odds of a 50% chance, based on the observed frequency of occurrence of 2 year La Niña events (13 out of 27 since 1880). The 60% prediction by CESM-DP-LE is larger than the historical odds from the CESM-CTL, in which 2 year La Niña events occur with a 30% frequency. This supports our conclusion that that the initial conditions associated with the El Niño of 2015 have increased the odds of 2 year La Niña. Lastly, our forecasts show very low probability of El Niño conditions for next winter. Additional work is needed to compare the skill of our predictions relative to that of operational systems, which provide forecasts with shorter lead times. This is an important issue since operational forecasts initialized in April of 2017 consistently predicted El Niño conditions for next boreal winter (Figure S9a). Since then, these forecasts have shifted to predicting neutral conditions, when initialized in July 2017 (Figure S9b), and to predicting returning La Niña conditions when initialized in September 2017 (Figure S9c). This drastic shift in predicted outcomes could have arisen from limited predictive skill during spring, the so-called "spring predictability barrier" (McPhaden, 2003; Webster & Yang, 1992), as well as from unrealistic dynamics of 2 year La Niña in the operational models. For instance, models in which ENSO is excessively oscillatory would tend to predict a subsequent El Niño when initialized right after the first La Niña peak (e.g., April 2017). Such model dependencies on dynamics of 2 year La Niña would need to be addressed in a multimodel framework. In sum, all lines of evidence presented here point to a return of La Niña that is more likely than not, along with an

unlikely return of El Niño. Despite the uncertainty inherent to any initialized climate forecast, the predictions presented here could be useful for assessing risk of flooding and drought during the upcoming boreal winter in areas throughout the world impacted by La Niña.

Acknowledgments
The authors would like to thank Michelle L'Heureux for her comments on this work and two anonymous reviewers for their insightful comments and constructive criticism. P. D. N., C. D., Y. O., A. K., and S. Y. are supported by the National Oceanic and Atmospheric Administration (NOAA) Climate Program Office's Modeling, Analysis, Predictions, and Projections program (P. D. N. NA14OAR4310229 and C. D. NA14OAR4310228 and NA17OAR4310145, and Y. O. NA14OAR4310227 and NA17OAR4310149) and Climate Variability and Predictability Program (A. K. and S. Y. NA13OAR4310138). A. K. and S. Y. are also supported by the National Science Foundation (NSF) Collaborative Research EaSM2 grant OCE-1243015. NCAR is sponsored by the NSF. Computing resources were provided by the Climate Simulation Laboratory at NCAR's Computational and Information Systems Laboratory (CISL), sponsored by the National Science Foundation and other agencies. The model output is available via the Earth System Grid. We acknowledge the Global Tropical Moored Buoy Array Project of NOAA/PMEL for providing subsurface ocean temperature data.

References

Adler, R. F., Huffman, G. J., Chang, A., Ferraro, R., Xie, P.-P., Janowiak, J., … Nelkin, E. (2003). The Version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation analysis (1979–present). *Journal of Hydrometeorology*, *4*, 1147–1167. https://doi.org/10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2

Balmaseda, M. A., Mogensen, K., & Weaver, A. T. (2013). Evaluation of the ECMWF ocean reanalysis system ORAS4. *Quarterly Journal of the Royal Meteorological Society*, *139*(674), 1132–1161. https://doi.org/10.1002/qj.2063

Barnston, A. G., Li, S., Mason, S. J., DeWitt, D. G., Goddard, L., & Gong, X. (2010). Verification of the First 11 Years of IRI's Seasonal Climate Forecasts. *Journal of Applied Meteorology and Climatology*, *49*(3), 493–520. https://doi.org/10.1175/2009JAMC2325.1

Cane, M. A., Clement, A. C., Kaplan, A., Kushnir, Y., Pozdnyakov, D., Seager, R., … Murtugudde, R. (1997). Twentieth-century sea surface temperature trends. *Science*, *275*(5302), 957–960. https://doi.org/10.1126/science.275.5302.957

Choi, K.-Y., Vecchi, G. A., & Wittenberg, A. T. (2013). ENSO transition, duration and amplitude asymmetries: Role of the nonlinear wind stress coupling in a conceptual model. *Journal of Climate*, *26*, 9462–9476. https://doi.org/10.1175/JCLI-D-13-00045.1

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., … Worley, S. J. (2011). The twentieth century reanalysis project. *Quarterly Journal of the Royal Meteorological Society*, *137*(654), 1–28. https://doi.org/10.1002/qj.776

Computational and Information Systems Laboratory (2017). Cheyenne: SGI ICE XA system (accelerated scientific discovery). Boulder, CO: National Center for Atmospheric Research. https://doi.org/10.5065/D6RX99HX

Cravatte, S., Delcroix, T., Zhang, D., McPhaden, M., & Leloup, J. (2009). Observed freshening and warming of the western Pacific Warm Pool. *Climate Dynamics*, *33*(4), 565–589. https://doi.org/10.1007/s00382-009-0526-7

Deser, C., Phillips, A. S., & Alexander, M. A. (2010). Twentieth century tropical sea surface temperature trends revisited. *Geophysical Research Letters*, *37*, L10701. https://doi.org/10.1029/2010GL043321

Deser, C., Simpson, I. R., McKinnon, K. A., & Phillips, A. S. (2017). The Northern Hemisphere extra-tropical atmospheric circulation response to ENSO: How well do we know it and how do we evaluate models accordingly? *Journal of Climate*, *30*, 5059–5082. https://doi.org/10.1175/JCLI-D-16-0844.1

DiNezio, P. N., & Deser, C. (2014). Nonlinear controls on the persistence of La Niña. *Journal of Climate*, *27*(19), 7335–7355.

DiNezio, P. N., Clement, A. C., Vecchi, G. A., Soden, B. J., Kirtman, B. P., & Lee, S.-K. (2009). Climate response of the equatorial pacific to global warming. *Journal of Climate*, *22*(18), 4873–4892. https://doi.org/10.1175/2009JCLI2982.1

DiNezio, P. N., Deser, C., Okumura, Y., & Karspeck, A. (2017). Predictability of 2-year La Niña events in a coupled general circulation model. *Climate Dynamics*, 1–25. https://doi.org/10.1007/s00382-017-3575-3

Ham, Y.-G., Schubert, S., Vikhliaev, Y., & Suarez, M. J. (2014). An assessment of the ENSO forecast skill of GEOS-5 system. *Climate Dynamics*, *43*(9), 2415–2430. https://doi.org/10.1007/s00382-014-2063-2

Hazeleger, W., Guemas, V., Wouters, B., Corti, S., Andreu-Burillo, I., Doblas-Reyes, F. J., … Caian, M. (2013). Multiyear climate predictions using two initialization strategies. *Geophysical Research Letters*, *40*, 1794–1798. https://doi.org/10.1002/grl.50355

Hoerling, M., Kumar, A., Dole, R., Nielsen-Gammon, J. W., Eischeid, J., Perlwitz, J., … Chen, M. (2013). Anatomy of an extreme event. *Journal of Climate*, *26*(9), 2811–2832.

Huang, B., Xue, Y., Zhang, D., Kumar, A., & McPhaden, M. J. (2010). The NCEP GODAS ocean analysis of the tropical pacific mixed layer heat budget on seasonal to interannual time scales. *Journal of Climate*, *23*(18), 4901–4925. https://doi.org/10.1175/2010JCLI3373.1

International CLIVAR Project Office (2011). *Data and bias correction for decadal climate predictions* (4 p.). Southampton, UK: International CLIVAR Project Office, CLIVAR Publication Series 150.

Jin, F.-F. (1997). An equatorial ocean recharge paradigm for ENSO. Part I: Conceptual model. *Journal of the Atmospheric Sciences*, *54*(7), 811–829. https://doi.org/10.1175/1520-0469(1997)054<0811:AEORPF>2.0.CO;2

Kessler, W. S. (2002). Is ENSO a cycle or a series of events? *Geophysical Research Letters*, *29*(23), 2125. 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 (*Technical Report 56*). Southampton, UK: CLIVAR Working Group on Seasonal to Interannual Prediction.

Kirtman, B. P., Shukla, J., Huang, B., Zhu, Z., & Schneider, E. K. (1997). Multiseasonal predictions with a coupled tropical ocean–global atmosphere system. *Monthly Weather Review*, *125*(5), 789–808. https://doi.org/10.1175/1520-0493(1997)125<0789:MPWACT>2.0.CO;2

Kohl, A., & Stammer, D. (2008). Variability of the meridional overturning in the North Atlantic from the 50-year GECCO state estimation. *Journal of Physical Oceanography*, *38*(9), 1913–1930. https://doi.org/10.1175/2008JPO3775.1

Latif, M., Kleeman, R., & Eckert, C. (1997). Greenhouse warming, decadal variability, or El Niño? An attempt to understand the anomalous 1990s. *Journal of Climate*, *10*(9), 2221–2239. https://doi.org/10.1175/1520-0442(1997)010<2221:GWDVOE>2.0.CO;2

McPhaden, M. J. (2003). Tropical Pacific Ocean heat content variations and ENSO persistence barriers. *Geophysical Research Letters*, *30*, 1480. https://doi.org/10.1029/2003GL016872

Meehl, G. A., & Teng, H. (2012). Case studies for initialized decadal hindcasts and predictions for the Pacific region. *Geophysical Research Letters*, *39*, L22705. https://doi.org/10.1029/2012GL053423

Meinen, C. S., & McPhaden, M. J. (2000). Observations of warm water volume changes in the equatorial pacific and their relationship to El Niño and La Niña. *Journal of Climate*, *13*(20), 3551–3559.

Okumura, Y. M., & Deser, C. (2010). Asymmetry in the duration of El Niño and La Niña. *Journal of Climate*, *23*(21), 5826–5843. https://doi.org/10.1175/2010JCLI3592.1

Okumura, Y. M., Ohba, M., Deser, C., & Ueda, H. (2011). A proposed mechanism for the asymmetric duration of El Niño and La Niña. *Journal of Climate*, *24*(15), 3822–3829. https://doi.org/10.1175/2011JCLI3999.1

Okumura, Y. M., DiNezio, P., & Deser, C. (2017). Evolving impacts of multi-year La Niña events on atmospheric circulation and US drought. *Geophysical Research Letters*, *44*. https://doi.org/10.1002/2017GL075034

Okumura, Y. M., Sun, T., & Wu, X. (2017). Asymmetric modulation of El Niño and La Niña and the linkage to tropical Pacific decadal variability. *Journal of Climate*, *30*(12), 4705–4733. https://doi.org/10.1175/JCLI-D-16-0680.1

Smith, T. M., Reynolds, R. W., Peterson, T. C., & Lawrimore, J. (2008). Improvements to NOAA's historical merged land-ocean surface temperature analysis (1880–2006). *Journal of Climate*, *21*(10), 2283–2296. https://doi.org/10.1175/2007JCLI2100.1

Stockdale, T. N. (1997). Coupled ocean–atmosphere forecasts in the presence of climate drift. *Monthly Weather Review*, *125*(5), 809–818. https://doi.org/10.1175/1520-0493(1997)125<0809:COAFIT>2.0.CO;2

Tang, Y., Kleeman, R., & Moore, A. M. (2005). Reliability of ENSO dynamical predictions. *Journal of the Atmospheric Sciences*, *62*(6), 1770–1791. https://doi.org/10.1175/JAS3445.1

Tippett, M. K., Barnston, A. G., & Li, S. (2011). Performance of recent multimodel ENSO forecasts. *Journal of Applied Meteorology and Climatology*, *51*(3), 637–654. https://doi.org/10.1175/JAMC-D-11-093.1

Uppala, S. M., Kllberg, P. W., Simmons, A. J., Andrae, U., Bechtold, V. D. C., Fiorino, M., … Woollen, J. (2005). The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, *131*(612), 2961–3012. https://doi.org/10.1256/qj.04.176

van Oldenborgh, G. J., Doblas-Reyes, F. J., Wouters, B., & Hazeleger, W. (2012). Decadal prediction skill in a multi-model ensemble. *Climate Dynamics*, *38*(7), 1263–1280. https://doi.org/10.1007/s00382-012-1313-4

Wang, W., Chen, M., & Kumar, A. (2010). An assessment of the CFS real-time seasonal forecasts. *Weather and Forecasting*, *25*(3), 950–969. https://doi.org/10.1175/2010WAF2222345.1

Webster, P. J., & Yang, S. (1992). Monsoon and ENSO: Selectively interactive systems. *Quarterly Journal of the Royal Meteorological Society*, *118*(507), 877–926. https://doi.org/10.1002/qj.49711850705

Yeager, S., Karspeck, A., Danabasoglu, G., Tribbia, J., & Teng, H. (2012). A decadal prediction case study: Late twentieth-century North Atlantic Ocean heat content. *Journal of Climate*, *25*(15), 5173–5189. https://doi.org/10.1175/JCLI-D-11-00595.1

Yeager, S. G., Karspeck, A. R., & Danabasoglu, G. (2015). Predicted slowdown in the rate of Atlantic sea ice loss. *Geophysical Research Letters*, *42*, 10,704–10,713. https://doi.org/10.1002/2015GL065364