Contrasting El Niño–La Niña Predictability and Prediction Skill in 2-Year Reforecasts of the Twentieth Century

S. SHARMILA,[a,b] H. HENDON,[a,c] O. ALVES,[a] A. WEISHEIMER,[d,e] AND M. BALMASEDA[d]

[a] *Bureau of Meteorology, Melbourne, Victoria, Australia*
[b] *Centre for Applied Climate Sciences, University of Southern Queensland, Toowoomba, Queensland, Australia*
[c] *Monash University, Melbourne, Victoria, Australia*
[d] *European Centre for Medium-Range Weather Forecasts, Reading, United Kingdom*
[e] *National Centre for Atmospheric Science, University of Oxford, Oxford, United Kingdom*

(Manuscript received 13 January 2022, in final form 11 October 2022)

ABSTRACT: Despite the growing demand for long-range ENSO predictions beyond 1 year, quantifying the skill at these lead times remains limited. This is partly due to inadequate long records of seasonal reforecasts that make skill estimates of irregular ENSO events quite challenging. Here, we investigate ENSO predictability and the dependency of prediction skill on the ENSO cycle using 110 years of 24-month-long 10-member ensemble reforecasts from ECMWF's coupled model (SEAS5-20C) initialized on 1 November and 1 May during 1901–2010. Results show that Niño-3.4 SST can be skillfully predicted up to ~18 lead months when initialized on 1 November, but skill drops at ~12 lead months for May starts that encounter the boreal spring predictability barrier in year 2. The skill beyond the first year is highly conditioned to the phase of ENSO: Forecasts initialized at peak El Niño are more skillful in year 2 than those initialized at peak La Niña, with the transition to La Niña being more predictable than to El Niño. This asymmetry is related to the subsurface initial conditions in the western equatorial Pacific: peak El Niño states evolving into La Niña are associated with strong upper-ocean heat discharge of the western Pacific, the memory of which stays beyond 1 year. In contrast, the western Pacific recharged state associated with La Niña is usually weaker and shorter-lived, being a weaker preconditioner for subsequent El Niño, the year after. High prediction skill of ENSO events beyond 1 year provides motivation for extending the lead time of operational seasonal forecasts up to 2 years.

KEYWORDS: ENSO; Pacific Ocean; Climate prediction; Forecast verification/skill; Tropics; Atmosphere-ocean interaction

1. Introduction

El Niño–Southern Oscillation (ENSO) is the most energetic naturally occurring coupled ocean–atmosphere phenomenon in the tropical Pacific (e.g., McPhaden et al. 2020) and a strong driver of weather/climate variability around the world via atmospheric teleconnections (Taschetto et al. 2020). ENSO is characterized by irregular episodes of anomalous sea surface temperature (SST) warming (El Niño) and cooling (La Niña) in the eastern and central equatorial Pacific from coupled dynamic and thermodynamic feedbacks with a periodicity of 2–7 years (e.g., Philander 1990; Timmermann et al. 2018). Owing to its profound global impacts extending to the environment, food security, economies, and social stability, the application of reliable ENSO prediction becomes a powerful way to provide meaningful and reliable guidance to decision-makers (Plummer et al. 2018). With the noticeable progress in coupled model developments, and improvement in initialization schemes in the recent decades, ENSO has become the most predictable climate mode at seasonal time scales (Tang et al. 2018), although untangling the complexities of ENSO dynamics and predicting ENSO precisely at longer lead time beyond a year remains a major challenge.

Over the recent decades, considerable research efforts have been dedicated to developing a hierarchy of dynamical forecast models, from idealized coupled models (e.g., Cane et al. 1986; Chen et al. 2004; Chen and Cane 2008) to fully coupled state-of-the-art general circulation models (GCMs) using sophisticated data assimilation techniques to generate initial conditions to predict ENSO evolution several seasons in advance (e.g., Alves et al. 2004; Jin et al. 2008; Luo et al. 2008, 2015; Balmaseda and Anderson 2009; Hendon et al. 2009; Barnston et al. 2012, 2019; Xue et al. 2013; Kirtman et al. 2014; Gonzalez and Goddard 2016; Hudson et al. 2017; Johnson et al. 2019; Wu et al. 2021a,b). Operational forecast centers around the world routinely provide real-time forecasts of ENSO events up to 9 months' lead time, although some models—such as from the North American Multi-Model Ensemble (NMME) and the European Centre for Medium-Range Weather Forecasts (ECMWF)—also provide ENSO outlooks up to 12 months ahead. Limiting lead time to less than a year derives, in part, from the inherent difficulty to predict across the boreal spring predictability barrier (BSPB)

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

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/JCLI-D-22-0028.s1.

Corresponding author: S. Sharmila, sharmila.sur@bom.gov.au

DOI: 10.1175/JCLI-D-22-0028.1

when the signal-to-noise ratio (SNR) is small, and from the large uncertainty driven by high-frequency atmospheric noise and the inherent complexity of the climate system. Seasonal variations of the coupled atmosphere–ocean mean state and decadal variations in the tropical Pacific can also affect ENSO prediction skill, with some decades showing skill limited to much less than 9 months (Balmaseda et al. 1995; Kirtman and Schopf 1998; Zhao et al. 2016; Weisheimer et al. 2020). Nevertheless, forecasting El Niño and La Niña events at a longer lead time beyond 12 months has received limited attention (Wu et al. 2021a,b).

Theoretically, ENSO is suggested to be potentially predictable a few years in advance as a result of the slowly varying self-sustained nature of the tropical Pacific coupled ocean–atmosphere system (e.g., Latif et al. 1994). Nonetheless, only a handful of studies have demonstrated the potential of predicting recent ENSO events beyond 1 year using long-lead retrospective forecasts with coupled model prediction systems (e.g., Luo et al. 2008, 2017; DiNezio et al. 2017a; Park et al. 2018). Recently, Wu et al. (2021a) explored the predictability of ENSO duration using coupled CESM1 model initialized hindcasts for the period 1954–2014 and promoted the usefulness of extending the ENSO prediction beyond one additional year. According to those previous studies, the prospects of long-lead ENSO predictability primarily lie in the subsurface ocean memory in the equatorial Pacific (e.g., Meinen and McPhaden 2000; DiNezio and Deser 2014; Planton et al. 2018), although westerly wind bursts (e.g., Vecchi and Harrison 2000) and the interaction between ocean preconditioning in the tropical Pacific Ocean and subseasonal winds and SST variations in the subtropical Pacific and adjacent oceans, such as through the heat flux–driven seasonal footprint mechanism (e.g., Vimont et al. 2003) and ocean dynamics–driven trade wind charging (Chakravorty et al. 2020), can also affect the ENSO predictability.

There is still limited consensus or baseline about a potential dependence of long-lead (beyond 1 year) ENSO predictability on initial ENSO state in a real-time prediction scenario. Using a perfect model framework, Planton et al. (2021) found the transition from El Niño to La Niña to be more predictable than the transition from La Niña to El Niño. DiNezio et al. (2017b), also using a perfect model framework, proposed that it should be possible to predict the duration of La Niña events up to 2 years ahead. Larson and Kirtman (2017), on the other hand, found increased predictability for transitions toward El Niño events, also in a perfect model framework. The study by Larson and Pegion (2020), using a multimodel of operational seasonal ENSO forecasts, did not find any conclusive evidence of whether El Niño events are more predictable than La Niña, or vice versa. Their statistics were limited by the number of models providing long-lead predictions and by the number of cases in their hindcast period (1982–2010). The more recent work by Wu et al. (2021a), using a longer reforecast period (1954–2015), discusses the possibility of predicting the duration of 1- and 2-yr El Niño or La Niña events, and finds that the duration of La Niña can be predicted 2 years ahead if initialized from the peak of a strong El Niño event, consistent with DiNezio et al. (2017b).

Motivated by the growing demand for knowledge of the likelihood of El Niño or La Niña in the following year and for multiyear regional climate information, here we aim to assess the potential of long-lead ENSO prediction up to 2 years using the latest 110-years of 24-month-long 10-member ensemble retrospective forecasts for the period 1901–2010 from the ECMWF coupled model (SEAS5-20C; Weisheimer et al. 2021). We explore the conditional long-range prediction skill and predictability of ENSO, with focus on the dependence on the phase of ENSO in the initial conditions. Unlike Wu et al. (2021a), this study does not deal with the specific aspect of predicting the duration of the events. With these 110 years of hindcast, we overcome some of the limitations of previous assessment with idealized modeling techniques that used substantially shorter historical analysis periods or hindcast periods. Compared to previous idealized modeling studies of long-lead ENSO prediction using hindcasts initialized over the past ~100 years (e.g., Chen et al. 2004), the atmosphere and ocean initial conditions are created with a state-of-the-art assimilation system. Although this assimilation cannot make up for the lack of in situ and satellite data during earlier periods, it does provide a consistent analysis using a fixed coupled model system that optimally uses the available ocean and atmosphere observations. Hindcasts from a state-of-the-art coupled model assimilation and prediction system spanning the period 1901–2010 will provide an improved estimate of ENSO predictability accounting for its large multidecadal variation [as suggested by Weisheimer et al. (2022)] and differences in predictability of El Niño and La Niña based on a much longer sample size.

The rest of the paper is organized as follows: section 2 briefly presents the model and hindcast experiments, data used for verifications and methods used for this study. In section 3, we assess ENSO predictability, focusing on the skill dependency of ENSO on initial ENSO states, the predictability of El Niño versus La Niña events and underlying climate processes affecting the predictability. The summary of key findings and the implications for future operational use are provided in section 4.

2. Model, data, and methods

a. SEAS5-20C hindcast experiments

In this study, we use the SEAS5-20C dataset (ECMWF 2021; Weisheimer et al. 2021), a suite of 2-yr-long ensemble hindcast experiments spanning the period 1901–2010. The forecast model used is a low-resolution configuration of ECMWF's fifth-generation operational coupled model seasonal forecasting system, SEAS5, documented in Johnson et al. (2019). The performance of the model, as well as the sensitivity to resolution, has been further assessed in the ocean–atmosphere coupled simulations in Roberts et al. (2018). They found that the model could reasonably reproduce various aspects of observed ENSO behavior, including magnitude, periodicity, asymmetry of ENSO variability, coupled feedback, and teleconnections in the historical simulations.

TABLE 1. Lists of identified El Niño and La Niña events and the subset of ENSO years transitioning to opposite phases for the period 1901–2010, denoted by the years when detected in November.

	El Niño years (EN_All)	La Niña years (LN_All)
All cases	1902, 1904, 1905, 1911, 1913, 1914, 1918, 1923, 1925, 1930, 1940, 1941, 1951, 1957, 1963, 1965, 1968, 1969, 1972, 1976, 1977, 1979, 1982, 1986, 1987, 1991, 1994, 1997, 2002, 2004, 2006, 2009	1903, 1906, 1909, 1910, 1912, 1915, 1916, 1917, 1922, 1924, 1933, 1938, 1942, 1944, 1947, 1949, 1954, 1955, 1956, 1964, 1970, 1971, 1973, 1974, 1975, 1983, 1984, 1988, 1995, 1998, 1999, 2007, 2010
	El Niño to La Niña (EN → LN)	La Niña to El Niño (LN → EN)
Transition subset	1902/03, 1905/06, 1911/12, 1914/15, 1923/24, 1941/42, 1963/64, 1969/70, 1972/73, 1982/83, 1987/88, 1994/95, 1997/98, 2006/07, 2009/10	1903/04, 1910/11, 1917/18, 1924/25, 1956/57, 1964/65, 1971/72, 1975/76

The initial conditions for SEAS-20C are provided by ECMWF's first Coupled Reanalysis of the 20th century (CERA-20C; Laloyaux et al. 2018) which has an ensemble of 10 members. CERA-20C aims at reconstructing the weather and climate of the coupled atmosphere, ocean, land, ocean waves, and sea ice system for the past period 1901–2010. CERA-20C assimilates observed subsurface temperature and salinity profiles in the ocean and conventional surface observations (surface pressure and marine winds) in the atmosphere, while no satellite-derived data went into the reanalysis. In CERA-20C, SSTs are constrained via relaxation toward the monthly HadISST2 analysis (Titchner and Rayner 2014). Weisheimer et al. (2022) have shown that the skill of SEAS5-20C for ENSO prediction during the recent 1981–2010 period is comparable to that of the operational system SEAS5 (see Fig. S1 in the online supplemental material).

The experiments for generating hindcasts were run at the atmospheric resolution Tco199 (~50 km) horizontally with 91 vertical levels and with a 1° horizontal ocean resolution using 42 vertical levels. "Tco" stands for triangular–cubic–octahedral grid (the shortest wave resolution is represented by four grid points), commonly used in lower-resolution atmospheric experiments at ECMWF. Historical constructions of greenhouse gases (CMIP5) and volcanic stratospheric sulfate aerosols were used as radiative forcings similar to SEAS5. All experiments have an ensemble size of 10 members, sampling the 10 realizations of each set of ocean ICs. The reforecasts were started on each 1 November and 1 May from 1901 to 2010 and have a lead time of 24 months. Recent studies show that an ensemble size of 10 is sufficient to estimate the ensemble mean and spread for 2-yr lead forecasts for predictions of El Niño and La Niña (Wu et al. 2021a), although likely more ensemble members will be needed for conditions where the SNR is lower. We have used the words *hindcast*, *forecast*, *reforecast*, and *prediction* synonymously throughout the remainder of this paper.

b. Data for forecast verification

We have utilized certain observational and reanalyses datasets to validate the model forecasts for the period of 1901–2010. For the assessment of SST forecasts, we used the ERSST.v5 monthly averaged gridded analyses with a horizontal resolution of 2.0° × 2.0° for the period 1901–2010 (Huang et al. 2017). The CERA-20C ocean heat content analyses for the upper 300 m (OHC300; Laloyaux et al. 2018) and ERA-20C monthly 10-m wind gridded reanalyses with a horizontal resolution of 1° × 1° for 1901–2010 (Poli et al. 2016) have also been used to verify reforecasts at each lead month.

c. Diagnostics for verification

Before assessing the 2-yr lead reforecasts, we remove the model climatology of each variable at each lead month by averaging the ensemble mean (averaged over 10 members) forecasts across all start years as a function of lead month for the period of 1901–2010. We computed the climate drift-adjusted forecast anomalies of each variable by subtracting the calculated model climatology of that variable for each lead month from each ensemble member forecast lead month and then we assessed their performance compared to observed anomalies that are similarly formed by subtracting the observed climatology of the same study period.

We have utilized various statistical approaches and skill scores/metrics to assess the skill and predictability of ENSO events. ENSO events are selected based on the observed SST anomalies averaged over the Niño-3.4 region (5°S–5°N, 170°–120°W, defined as the Niño-3.4 index or N34) for 1901–2010. The year when an ENSO event first develops is denoted as year 0; herein Nov^0 is used to refer to the month of November in the same year when the mature phase of ENSO starts and May^0 denotes the month of May of the same calendar year. For instance, for the 1997/98 El Niño event, Nov^0 and May^0 refer to November and May of 1997. To assess the state-dependency and long-lead prediction skill of ENSO from November-initialized (Nov^0) forecasts, individual near mature or peak El Niño (La Niña) events are selected when the normalized Niño-3.4 index in Nov^0 is $\geq +0.5$ (≤ -0.5). The identified El Niño (EN_All) and La Niña (LN_All) years and their transition subsets are listed in Table 1.

To quantify overall long-lead prediction skill and predictability of ENSO in SEAS5-20C, we have applied both deterministic and probabilistic approaches. The ensemble-mean

anomaly correlation coefficient (ACC) is used to quantify forecast accuracy, which is corroborated using the root-mean-square error (RMSE). Forecast uncertainty is quantified using ensemble spread and signal-to-noise ratio (SNR; the ratio of the ensemble mean divided by the ensemble spread). We use composite analysis to measure the capability of SEAS5-20C to predict the temporal evolution of Niño-3.4 SST anomalies and other oceanic/atmospheric variables of El Niño and La Niña events and their transitions starting from Nov^0-initialized forecasts compared to the observations.

3. Results and discussion

a. *State dependency of ENSO prediction skill*

We start with assessing the overall ENSO prediction skill of SEAS5-and how the prediction skill varies with selected hindcast periods. Figure 1 shows the ACC skill of the Niño-3.4 index as a function of forecast lead month initialized on 1 November (Fig. 1a) and 1 May (Fig. 1b) respectively for three different hindcast periods, the entire 1901–2010 period (110 years; green curve), the most recent 50 years (1961–2010; blue curve), and the most recent 30 years (1981–2010; red curve). We refer to the forecast accuracy as being skillful as long as the ACC is significantly different from zero. Here, ACC $\geq$ 0.4 is significant at 95% confidence level for a sample size $\geq$ 30 using a standard t test. For Nov^0 initialized forecasts (Fig. 1a), the ENSO skill varies from 6 to 18 months lead for selected hindcast periods but performs best in the most recent 30-yr period as might be expected due to increased availability of observations used to initialize the forecasts. All hindcast periods exhibit a skill drop associated with BSPB and subsequent skill recovery, but with different magnitudes. Forecasts initialized in May, which is around the peak of BSPB, show comparatively lower ACC than those initialized in November. The forecasts from May remain skillful only out to 10–13 lead months, which coincides with the BSPB in the second year. The possible role of any externally forced trends for causing the skill variation across the three epochs considered here was explored by recomputing the ACC using detrended data and no appreciable difference in skill was found (See in supplemental Fig. S2).

We additionally complement the analysis using ACC with corresponding RMSE and ensemble spread and standard deviation over the three hindcast periods (supplemental Fig. S3). The model is seen to consistently simulate too weak ENSO variability in year 2 and beyond, which presumably limits the seasonal recovery of skill after the BSPB. For the most skillful recent period (1981–2010), the ensemble spread follows the RMSE during the first year for November-initialized forecasts, indicating a reliable ensemble, but then remains underdispersed in year 2. There is also an indication of recovery of skill from Aug^{+1} onward, which matches well the seasonal increase in observed standard deviation. The ACC is seen to reach <0.4 when the RMSE is nearly equal to the observed standard deviation around the BSPB in $May–Jun^{+2}$ in year 2 beyond lead month 18. However, RMSE grows rapidly after lead month 7 during other historical periods. Further analysis reveals that the model consistently

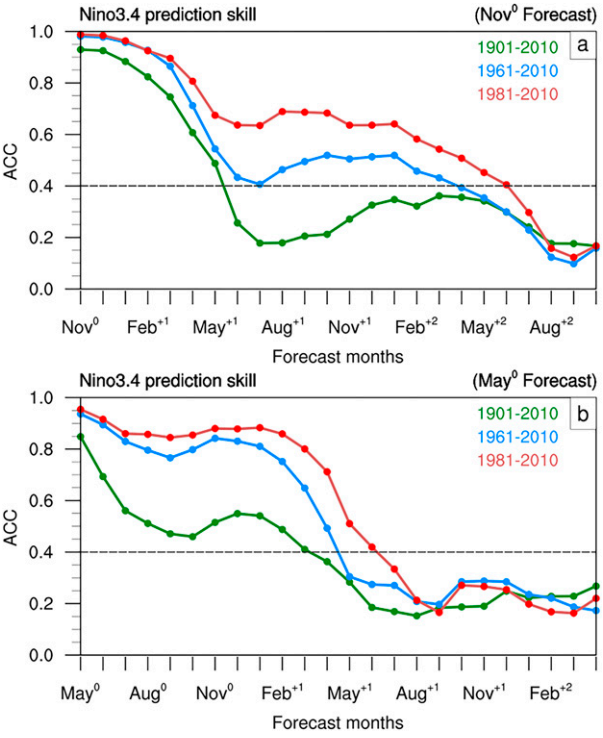

FIG. 1. Anomaly correlation coefficient (ACC) skill scores of ensemble-mean monthly Niño-3.4 anomaly as a function of forecast lead months initialized at (a) 1 Nov and (b) 1 May for selected three hindcast periods: 1901–2010, 1961–2010, and 1981–2010.

simulates too weak ENSO variability in year 2 and beyond, and weaker wind variability at the early stage of the seasonal forecasts (supplemental Figs. S3 and S4), which presumably limits the seasonal recovery of skill after the BSPB.

Although the results in Fig. 1 seem to indicate that the best skill is achieved in the most recent epoch, when presumably the initial conditions are best constrained, the story is not that simple. Recently, Weisheimer et al. (2022) highlighted distinct periods of enhanced long-range ENSO skill at the beginning and at the end of the twentieth century, and an extended multidecadal epoch of reduced skill during the 1930s to 1950s in SEAS5-20C. Their study suggests that the nonmonotonic skill modulation (e.g., Zhao et al. 2016) is largely driven by multidecadal variations in the dynamical characteristics of ENSO rather than the data coverage and quality of the observations. Their results show that forecast skill in the first decades of the record was similarly good as that achieved during the most recent decades, thus justifying extension of the hindcasts back to the early 1900s. In this context, the strength of coupling between the extratropics and tropics during those epochs could also influence the prediction skill (Pivotti and Anderson 2021).

Next, we quantify the state-dependency of ENSO prediction skill (Fig. 2) by computing the ACC scores separately for forecasts initialized in Nov^0 at near peak El Niño (denoted as EN_All, Nov^0) and La Niña (denoted as LN_All, Nov^0) based on the events listed in Table 1. Forecasts are skillful up to 18 months when initialized near peak El Niño (Fig. 2a, red

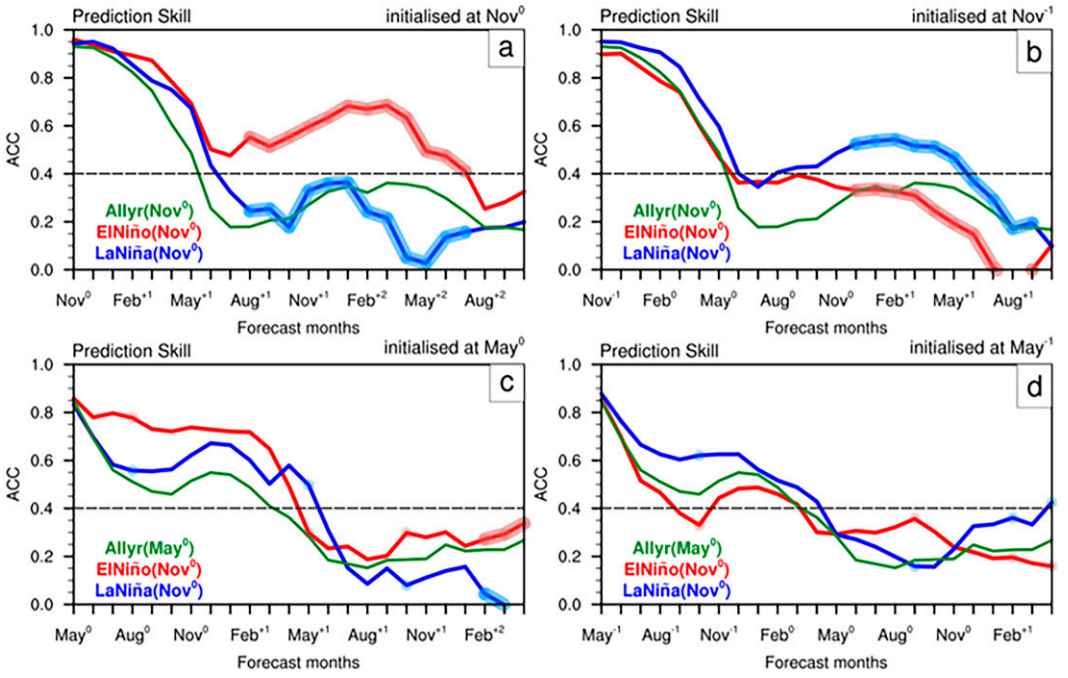

FIG. 2. ACC score-based skill dependency of model forecasts when initialized at (a) El Niño (Nov^0, EN_All; red) and La Niña (Nov^0, LN_All; blue) states for a 1 November start (Nov^0), and corresponding prediction skill of El Niño (Nov^0; red) and La Niña (Nov^0; blue) when initialized from (b) Nov^{-1} (12-month lead), (c) May^0 (6-month lead), and (d) May^{-1} (18-month lead) as a function of forecast lead months. The ACC scores for the whole period (all years, 1901–2010; as in Fig. 1) for both November and May starts are shown in green for reference. Skill differences between El Niño and La Niña significantly different from zero at the 5% level are highlighted with shading. El Niño (Nov^0) and La Niña (Nov^0) denote the near peak El Niño and La Niña state, respectively.

curve), with the skill bouncing back after the BSPB (i.e., beyond lead time 7 months). In contrast, a larger drop in ACC score below the 0.4 threshold is evident during boreal spring for the forecasts initialized from La Niña conditions (Fig. 2a, blue curve) and does not recover afterward. The skill difference between El Niño and La Niña initial states is statistically significant at lead times > 7 months (skill difference at 5% significance level are highlighted). Here the effective sample size has been estimated following Bretherton et al. (1999) and the significance of skill difference is computed using a t test. We further analyze the results with RMSE score, which matches well with the overall ACC score (supplemental Fig. S5). When the forecasts are initialized in May^0 from an El Niño or La Niña state, the differences in skill between El Niño and La Niña events are not statistically significant (supplemental Fig. S6). Note that, in this case, the subsets of El Niño and La Niña initial states are identified based on a similar normalized Niño-3.4 index but for the month of May, and mostly are at the decaying/onset stage. In both cases, the ACC skill recovery after the BSPB can sustain ACC values larger than 0.4 for about 12 months.

In Fig. 2b, we further quantify the skill for predicting the development of La Niña and El Niño in Nov^0 (and onward) from initial conditions in the previous year (Nov^{-1}). Results show that predicting the onset and development of La Niña in Nov^0 from Nov^{-1} is more skillful (over 18 months lead) than predicting the development of an El Niño, consistent with the fact that La Niña generally more robustly follows an El Niño event rather than vice versa. A similar assessment was made using May initialized forecasts (May^0) to predict peak El Niño and La Niña in Nov^0 at 6-month lead (Fig. 2c) and 18-month lead (from May^{-1}; Fig. 2d). The results indicate that May initialized forecasts are skillful up to a 12-month lead. At short lead times (Fig. 2c), the skill is significantly larger for predicting mature El Niño (Nov^0) than La Niña (Nov^0), with the former having a weaker BSPB. In contrast, the BSPB of the second year is stronger for El Niño conditions. The latter is consistent with the increased skill for predicting La Niña (Nov^0) during the first 12 months in forecasts initialized in May^{-1} in Fig. 2d. These results indicate that SEAS5-20C has high predictive skill for ENSO events with a lead time beyond a year when initialized at the 1 November near mature El Niño state.

b. *Predictability of El Niño and La Niña in Nov^0 forecasts*

To illustrate the results of the skill scores discussed above, Fig. 3 shows the typical capability of SEAS5-20C 24-month forecasts in predicting the deterministic temporal evolution of Niño-3.4 SST anomalies when forecasts are initialized at Nov^0 near the peak of El Niño (EN_All; Fig. 3a, red) and of La Niña (LN_All; Fig. 3b, blue). The figure shows composites of

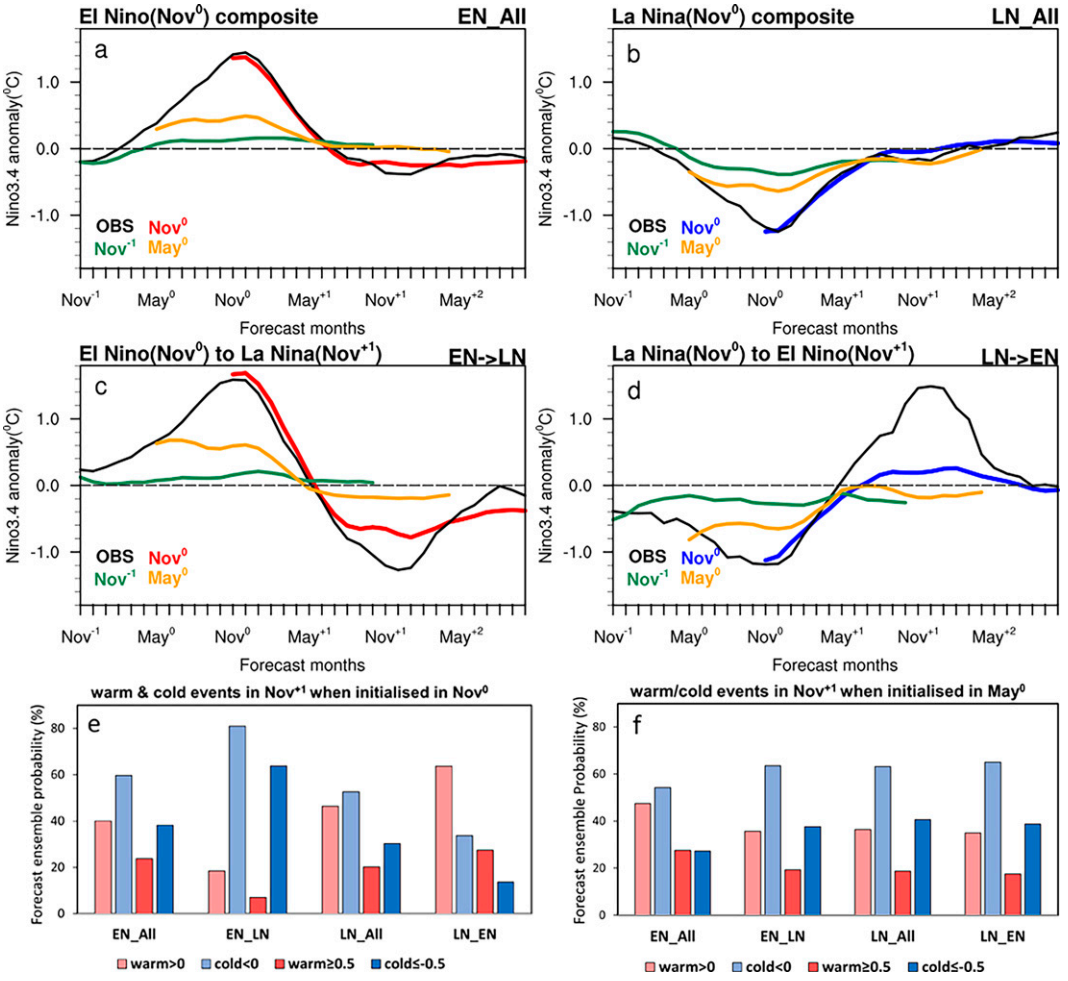

FIG. 3. Time evolution of Niño-3.4 index (°C) in observations (black curves) and ensemble-mean forecasts (colored curves) composited for (a) EN_All and (b) LN_All, and the (c) EN → LN and (d) LN → EN transitions for the period 1901–2010. The forecasts are initialized in Nov^{-1} (green), May^{0} (yellow), and Nov^{0} (red for EN; blue for LN). Nov^{0} denotes the peak of identified El Niño and La Nina events. (e),(f) Probabilistic aspects of forecast ensembles in predicting Nov^{+1} Niño-3.4 index from all ensemble forecast members of each composite for forecast initialized in Nov^{0} and May^{0}, respectively, are expressed in percentages.

observations and the corresponding ensemble-mean forecasts for the identified ENSO years (Table 1a). The ensemble mean will only capture the deterministic component of the predictability, so even a perfect forecast could underestimate the amplitude of the observations. In Fig. 3a, the composited time series of observed Niño-3.4 SST anomalies shows that El Niño events typically decline after their peak and evolve to near La Niña conditions within 6 months in the next year. We find that the model predicts well the duration of the warm event from peak to termination, as well as the onset of La Niña conditions in the next year, providing successful predictions of the ENSO cycle beyond 1.5 years, consistent with Fig. 2a.

The figure also shows the composite evolution for forecasts initialized in May^{0} (6-month lead; yellow) and Nov^{-1} (12-month lead; green) for transitions to peak El Niño (Nov^{0}) (Table 1). For these longer lead times, the ensemble mean can capture the sign of the warm anomaly at Nov^{0}, but it largely underestimates the observed amplitude of El Niño, consistent with a substantial random component in the onset of a warm event (we will come back to this point later). The ensemble mean struggles to capture the transition to cold conditions in the second year. This result fits the interpretation that the transition to a cold state is conditioned to the intensity of a preceding warm event, which is not easy to predict in a deterministic sense from May^{0} and Nov^{-1} initial conditions.

The composites of observed Niño-3.4 SST anomalies for LN_All events shown in Fig. 3b indicate that in general La Niña decays more slowly than El Niño, remaining in cold-neutral conditions for the following year, whereas El Niño tends to transition to a cold state within the first year. This is consistent with previous studies (e.g., Kessler 2002; Okumura and Deser 2010), and points to a clear asymmetry in the ENSO evolution. The forecasts initialized around peak La Niña

provide good predictions up to ~10 lead months, after which SST anomalies remain neutral and so there is no signal to predict. This can explain the lower correlation values of the blue curve in Fig. 2a. Forecasts initialized in May^0 and Nov^{-1} before the peak La Niña (Nov^0), overall capture the negative sign of the SST anomalies, a noticeable difference with the equivalent composite for equivalent lead times for EN_All events in Fig. 2a.

We further use a subset of ENSO events that directly transition to the opposite phase (Table 1b). Compared with EN_All, for the transition subset El Niño (Nov^0) to La Niña (Nov^{+1}) (denoted as EN → LN) the model predicts better the amplitude of SST anomalies associated with the typical evolution of mature El Niño (Nov^0) transitioning to La Niña (Nov^{+1}) and its continuation beyond 1.5 years lead (Fig. 3c). We repeat the exercise for the LN → EN events (although the number of cases is small) initialized during La Niña events (Nov^0) with direct transition into El Niño (Nov^{+1}) (Fig. 3d). For these subsets, the model is able to capture the change of sign from cold to warm conditions but struggles to capture the timing of the rapid termination of La Niña and the transition to El Niño (Nov^{+1}) in the subsequent boreal spring. It largely underestimates the amplitude of warm phase SST anomalies in the second year, although arguably not as much as for the LN_All events (e.g., Nov^{-1} starts in Fig. 3a).

The probabilistic performance of the model for the different composites is illustrated in Figs. 3e and 3f for May^0 and Nov^0 initialization. The figure shows the probability (relative number of ensemble members) of having a cold or warm event for the different composites in Nov^{+1}. The forecast produces higher probabilities of cold events in Nov^{+1} when starting from El Niño conditions, and successfully captures the increased probability of cold conditions in the EN → LN subset with respect to the EN_All subset, both for Nov^0 and May^0 initialization. In contrast, for the La Niña initialized forecasts, the distribution between cold and warm anomalies in Nov^{+1} is more similar, although again the model successfully increases the probability of warm anomalies in the LN → EN subset compared with those in LN_All when initialized in Nov^0. Notably, this is not the case for forecasts initialized in May^0, where the model tends to produce an increased probability of cold anomalies in Nov^{+1}. While high probabilities of cold anomalies in the subset LN_All (initialized in May^0) can be a manifestation of the skewness of the ENSO cycle, with La Niña condition lasting longer than 1 year, the higher probabilities of cold anomalies in LN → EN subset show a model deficiency.

The results confirm that forecasts initialized near the mature El Niño state have the longest-range prediction skill, with the model being able to predict the occurrence of La Niña beyond 12 months. In contrast, when initialized from La Niña conditions, the deterministic prediction skill of the amplitude of warm events at lead months longer than 12 appears modest, although at these lead times the forecasts initialized in Nov^0 are able to predict the termination of the cold conditions with probability larger than 60% (table not shown). In the following subsections, we will consider only the forecasts initialized in Nov^0. Given that a robust probabilistic verification of the skill from different subsets is difficult with the limited number of samples and ensemble members, we will focus the discussions on the processes leading to the different levels of predictability.

c. *Physical processes contributing to the predictability in forecasts initialized in the mature ENSO phase (Nov^0)*

Both oceanic and atmospheric processes play roles in the predictability of various aspects of ENSO. Here we explore the interplay between these processes in the predictability of onset, amplitude, and duration of ENSO events, and we explore the model's capacity to predict the oceanic and atmospheric climate processes that could contribute to the differences in long-lead predictability for El Niño and La Niña events in Nov^0 initialized forecasts. Figures 4 and 5 compare the composited spatial maps of 2-month-averaged observed (left column) and predicted (right column) fields for forecasts initialized from the mature phases of El Niño (EN_All) and La Niña (LN_All) conditions, respectively. The top panels show the 2-month-averaged SST and surface wind anomalies (Figs. 4a and 5a) and the bottom panels show the upper ocean heat content (OHC300) (Figs. 4b and 5b) anomalies initialized at 1 November, both at lead month 0 (Nov^0/Dec^0) and lead month 12 (next year, Nov^{+1}/Dec^{+1}).

For both El Niño and La Niña initialized states, the SST anomalies in the tropical Pacific and adjacent oceans are well captured in Nov^0/Dec^0 (top panels of Figs. 4a and 5b) although there is a westward shift in the ENSO SST anomalies and weaker wind anomalies. The westward displacement of ENSO-related equatorial anomalies is associated with the erroneous westward penetration of the Pacific cold tongue in the SEAS5-20C model simulation, a common issue in the current state-of-the-art climate models (see, e.g., Wu et al. 2022). The model also predicts well the recharge/discharge dipole in the OHC300 anomalies between western and eastern equatorial Pacific region (top panels of Figs. 4b and 5b). At lead month 12, the model could predict the structure of changes in the oceanic and wind anomalies in transitioning to La Niña–like conditions in the next year Nov^{+1}Dec^{+1} from an El Niño initialized state Nov^{0}Dec^0 (bottom panels of Fig. 4a), consistent with the model being able to capture the duration of the warm phase through the first year (see also Fig. 3). However, when initialized from La Niña conditions (bottom panels of Figs. 5a,b), the model seems to transition to neutral condition faster than observations, with faster than observed release of heat from the equatorial western Pacific to the central/eastern Pacific, which manifests on weaker than observed positive OHC300 anomalies in the western Pacific, and stronger than observed warm conditions in the equatorial eastern Pacific in both OHC300 and SST. This can be seen more clearly by following the time sequence of spatial maps from Nov^{0}Dec^0 to Nov^{+1}Dec$^{+1,}$ shown in supplemental Figs. S7 and S8.

We further assess the equatorial physical processes contributing to the predictability difference between the time evolution of El Niño and La Niña initialized states. Figure 6 shows the composited spatiotemporal evolution of SST, surface

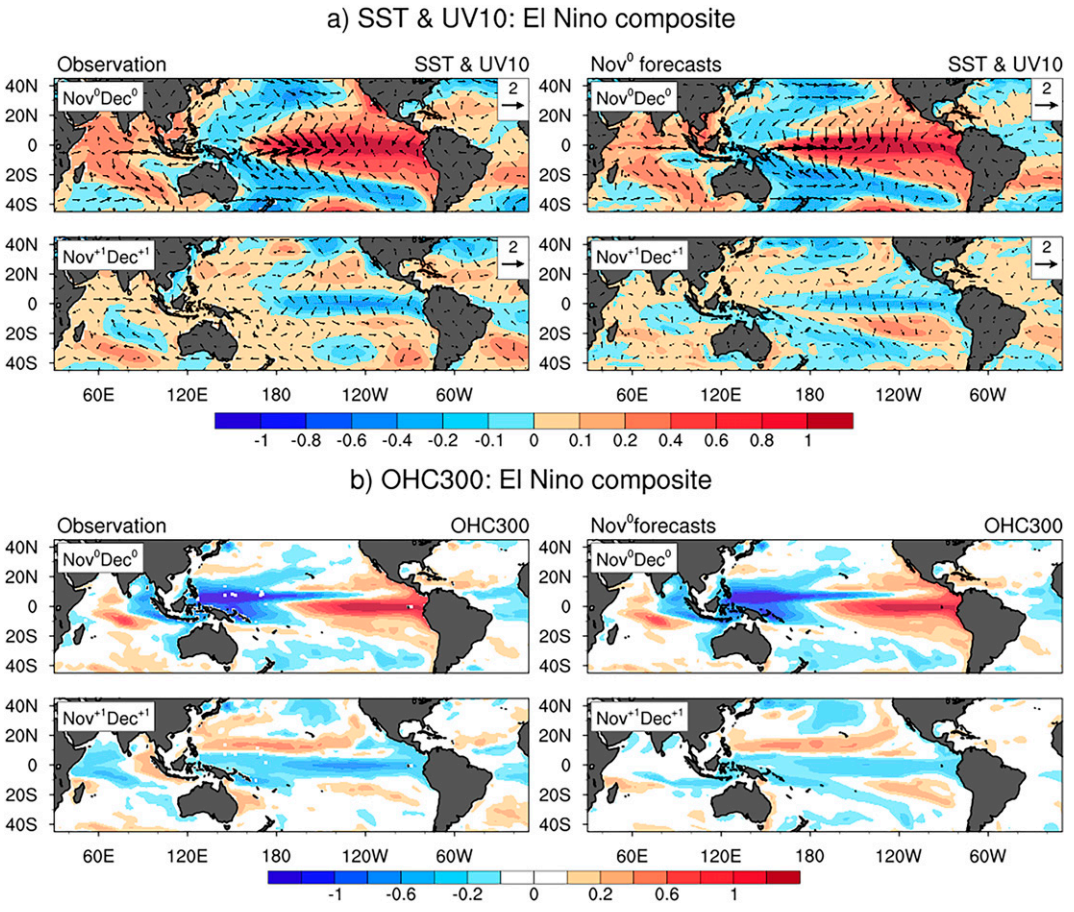

FIG. 4. Two-month-averaged (Nov^0Dec^0 and $Nov^{+1}Dec^{+1}$) maps of composited (a) SST (shading) and surface winds (vectors) and (b) OHC300 based on all El Niño events from (left) observations and (right) ensemble-mean forecasts initialized at El Niño states (Nov^0).

wind, and OHC300 anomalies averaged over the equatorial Pacific (5°S–5°N) region for El Niño events (EN_All) in observations (Figs. 6Ia,c) starting from Nov^{-1}, and Nov^0-initialized ensemble mean forecasts (Figs. 6Ib,d). Observations show the OHC300 recharge state from Nov^{-1} and gradually induced surface westerlies in the western-to-central equatorial Pacific acting as the precursors of the developing El Niño warming in the eastern equatorial Pacific (Figs. 6Ia,c). Near the peak El Niño phase (Nov^0), westerly wind anomalies around the date line are evident, consistent with the idea that by this time the Bjerknes feedback is already well established. The western Pacific has started discharging the heat, with the peak of OHC300 positive anomalies now confined to the east of the date line. The OHC300 negative anomalies will gradually propagate eastward to encompass the whole equatorial basin for the duration of the warm event, creating favorable conditions for the termination of El Niño and transition to the cold phase in the following year.

The Nov^0 initialized forecasts predict well the warmer SST anomalies and the eastward propagation of subsurface heat related to the mature stage of El Niño (Figs. 6Ic,d) and capture the transition to the cold phase the following year. However, the western Pacific heat discharge is weaker. We speculate that this weak discharge is associated with the forecast error in the first few months, related to the model weak interannual variability of the zonal wind (which appears in the first month; see supplemental Fig. S4) and/or the westward displacement of ENSO SST/wind anomalies and warmer OHC300 in the warm pool region (supplemental Fig. S8). Thus, a westward displacement of the SST anomaly in the forecasts weakens the westerly anomaly around the date line, which can weaken the OHC300 discharge in the western Pacific during the first year and results in warmer equatorial conditions in the second year. We note that the westward displacement of the anomaly error is visible in other models (e.g., Wu et al. 2021a, 2022) and is likely related to the mean-state error of westward displacement of the cold tongue and easterly wind bias (supplemental Fig. S8). The too-strong easterly wind bias may also be related to the weak wind variability. Thus, errors in the zonal winds and cold-tongue position in the first months of the forecasts can affect the prediction skill beyond year one, as discussed in Wu et al. (2022).

Given these general forecast deficiencies associated with the cold tongue and weak El Niño discharge, can we expect

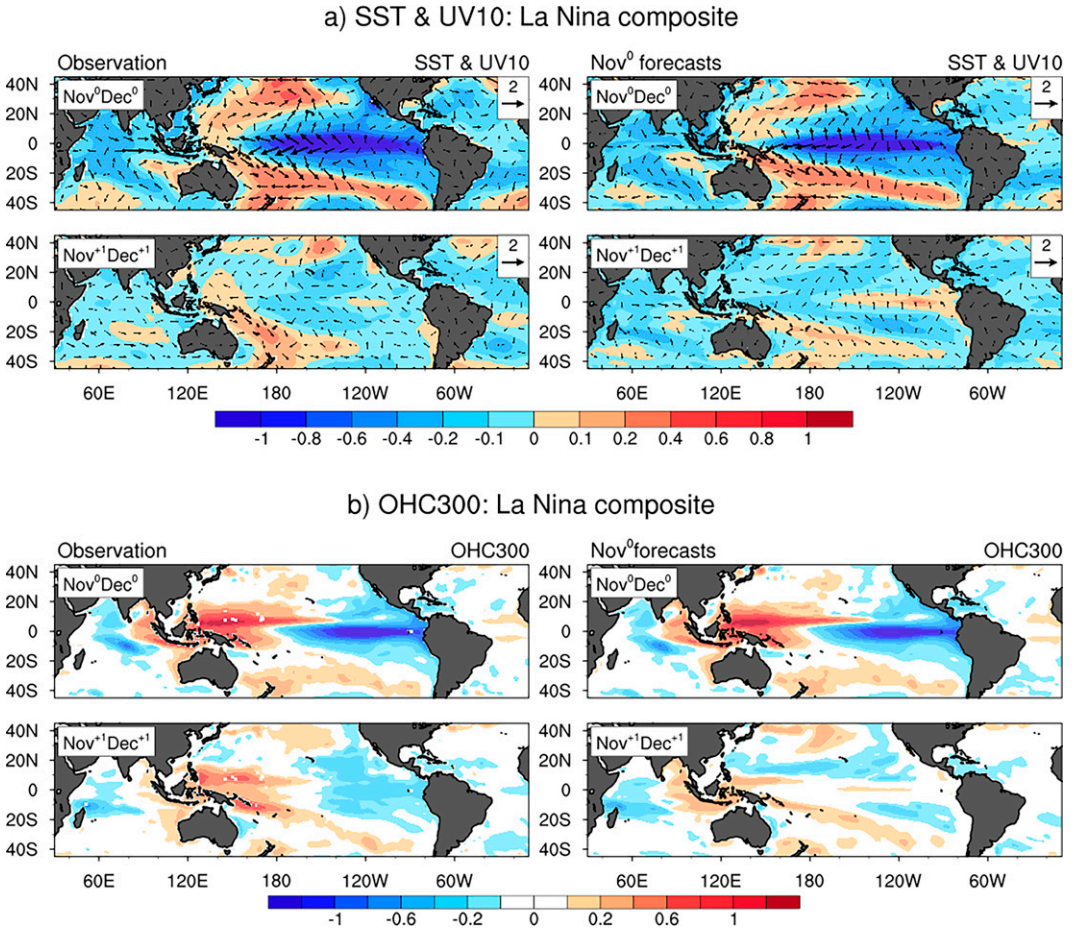

FIG. 5. (a) Two-month-averaged ($Nov^{0}Dec^{0}$ and $Nov^{+1}Dec^{+1}$) maps of composited SST (shading) and surface winds (vectors) and (b) OHC300 based on all La Niña events from observations and ensemble-mean forecasts initialized at La Niña states (Nov^{0}).

the forecast to be able to predict the transition from El Niño to La Niña? In Fig. 3c above we have seen that the model composites for the ENSO indices behave better for these subsets of cases. Figure 6II shows the composites of El Niño events transitioning to La Niña events (EN → LN events in Table 1b). Compared to the observed EN_All composite, the observed EN → LN composite (Fig. 6IIa) shows stronger SST anomalies in the first year and stronger and faster heat discharge in the equatorial western Pacific. The forecasts can capture the difference between the EN_All and EN → LN, being able to predict the stronger amplitude of the El Niño, its termination and transition to La Niña in the next year, preceded by strong OHC discharge associated in the western Pacific region and subsequent eastward propagation. The ability of the model to distinguish between transition and standard events implies that this information is contained in the initial conditions. Indeed, there are noticeable differences in the precursors of the El Niño to La Niña subset (EN → LN) with respect to EN_All events in the Nov^{0} initial conditions: both the negative OHC300 anomaly over the western Pacific OHC300 and the positive SST anomalies east of the date line are stronger than in the EN_All. These conditions seem to have been developing during the preceding year (from Nov^{-1}). These precursors favor the occurrence of a stronger than average El Niño event from Nov^{0} to May^{+1} and appear to affect the predictability of La Niña (Nov^{+1}) in the second year. Based on this information in the initial conditions, the forecasts can discern the properties of these strong El Niño leading to La Niña from the standard moderate El Niño events. But we also note that the predicted La Niña SST anomalies (Nov^{+1}) in the eastern Pacific are weaker than observed. This can be a consequence of the probabilistic nature of the forecasts (the ensemble mean amplitude can be weaker than observations, particularly in the most extreme cases), or a consequence of model error (e.g., weak discharge associated with western displacement of the cold tongue, or weak interannual wind variability in the model, as discussed above).

We further assess what controls the predictable component in Nov^{0} forecasts from La Niña initialized states. Composites of the LN_All cases and subset of LN → EN are shown in

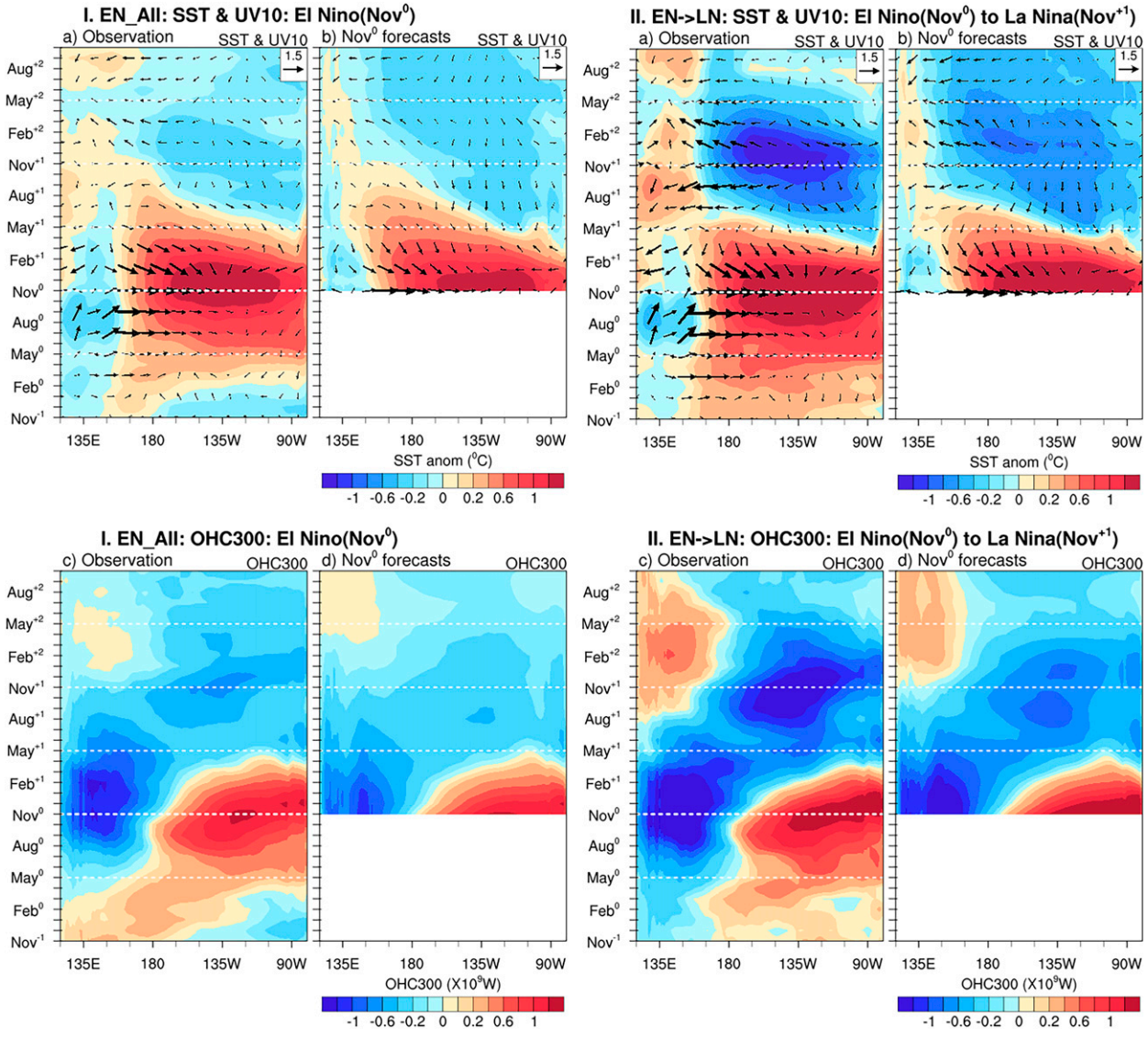

FIG. 6. Longitude–time evolution of composited (a),(b) SST (shading) and surface winds (vectors) and (c),(d) OHC300 averaged over equatorial region (5°S–5°N) for (I) all El Niño events (EN_All) and (II) transition from El Niño (Nov^0) to La Niña (EN → LN) in observation (starting from previous year) and ensemble-mean forecasts initialized at El Niño states (Nov^0).

Fig. 7. Composited LN_All observations (Figs. 7Ia,c) show that La Niña tends to develop around May^0 in the central-eastern equatorial Pacific, and easterly wind anomalies over the western-central equatorial Pacific force an eastward-propagating signal, inducing cold OHC300 anomalies in the eastern equatorial Pacific around the peak of La Niña. The Nov^0 initialized forecasts (Figs. 7Ib,d) predict well the cold SST and OHC anomalies related to the mature stage of La Niña in the central-to-eastern Pacific and positive OHC anomalies in the western Pacific in the first year. Compared with the observations, the model predicts a more westward displacement of cold ENSO anomalies in the second year, as well a premature initiation of the warm phase (consistent with Fig. 3b), associated with a faster recharge of the equatorial Pacific heat content.

We further investigate the LN → EN subset, comprising La Niña events transitioning to El Niño (Fig. 7II). Observations show that in these cases, El Niño is preceded by more persistent and stronger than normal La Niña events (Fig. 7IIa). These La Niña events are accompanied by enhanced recharge of OHC300 in the equatorial Pacific (generally via meridional Sverdrup transport; Meinen and McPhaden 2000), which leads to deepening of the thermocline, the appearance of stronger westerly wind events, and development of anomalous SST warming through recharge process (Jin 1997), thus transitioning to El Niño in the following year through the Bjerknes feedback. The Nov^0 forecasts predict well the initial La Niña–related SST anomalies in the equatorial Pacific in year 1 but underestimate the magnitude of oceanic and atmospheric conditions while

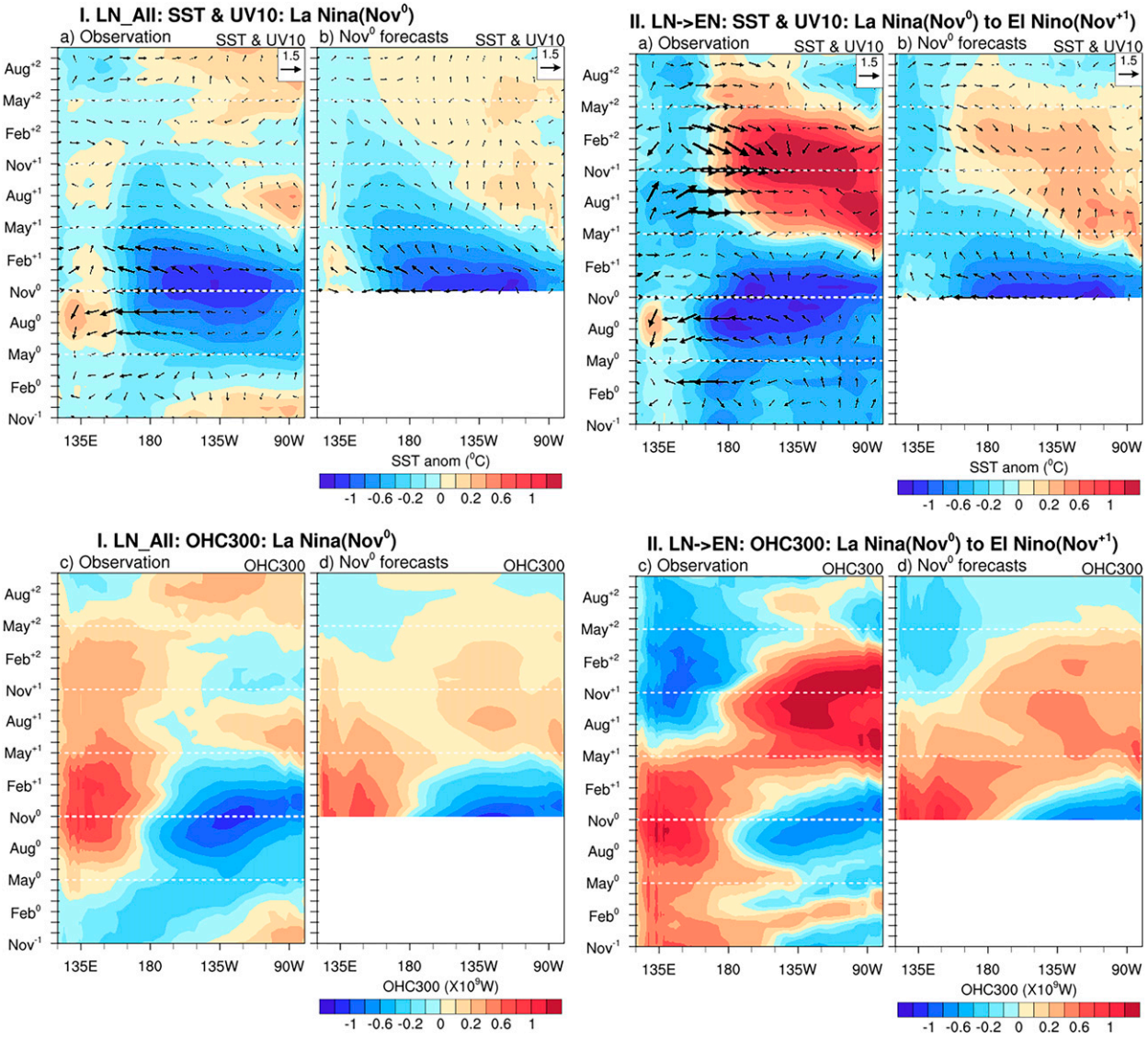

FIG. 7. As in Fig. 6, but for La Niña (Nov^0) initialized forecasts.

evolving to a warm phase a year after. The warm SST anomalies in the central-eastern Pacific during Nov^{+1} is largely underestimated, affected by weaker westerlies in the western-central equatorial Pacific around May^{+1}–Aug^{+1}, coupled with weaker OHC recharge in the eastern equatorial OHC300. The weaker amplitude of the forecast compared with the observed composite is consistent with El Niño events being triggered by stochastic forcing during boreal spring (e.g., Dommenget et al. 2013; Planton et al. 2021). This will lead to reduction in amplitude of the ensemble mean El Niño–related SST anomalies in the equatorial Pacific and hence weaker long-lead deterministic predictability of ENSO states from peak La Niña initialized conditions. Still, compared with the standard La Niña cases in Fig. 7I, when the forecasts are initialized from long-lasting La Niña conditions, they favor the occurrence of warmer conditions in the second year, so the model is able to differentiate between the LN_All and LN → EN subsets.

d. *The impact of western equatorial Pacific recharge–discharge on ENSO predictability*

In this section, we try to characterize the different subsets and their predictability in terms of ocean heat content variations, which can give some insight into the ENSO energy cycle and role of the ocean initial conditions. The equatorial Pacific OHC is regarded as a key ocean memory for ENSO predictability at seasonal time scales, but so far it remains unclear if it can be used as a predictor at time scales beyond 1 year. Recent studies indicate that subsurface OHC anomalies in the western equatorial Pacific better encompass the lower-frequency ENSO dynamics and could be more promising precursors of ENSO events beyond a year (Planton et al. 2018, 2021; Izumo et al. 2019). We first estimate the deterministic skill of both the western equatorial Pacific OHC (OHC_W; averaged over 5°S–5°N, 120°–205°E) and the entire equatorial Pacific OHC

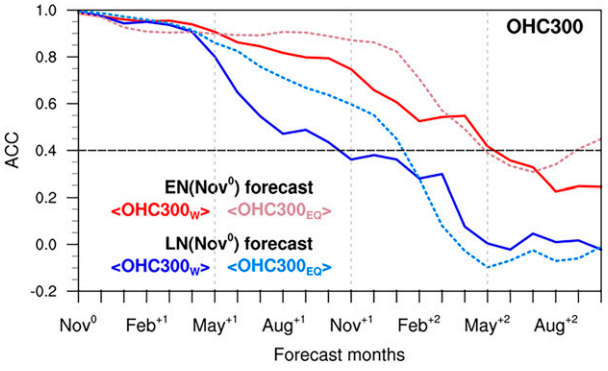

FIG. 8. ACC of ensemble-mean monthly OHC300 anomaly as a function of forecast lead months averaged over entire equatorial Pacific (OHC_{EQ}; 5°S–5°N, 120°–280°E; dashed curve) and western Pacific (OHCw; 5°S–5°N, 120°–205°E; solid curve) initialized at Nov^0 near peak El Niño state (red) and La Niña state (blue).

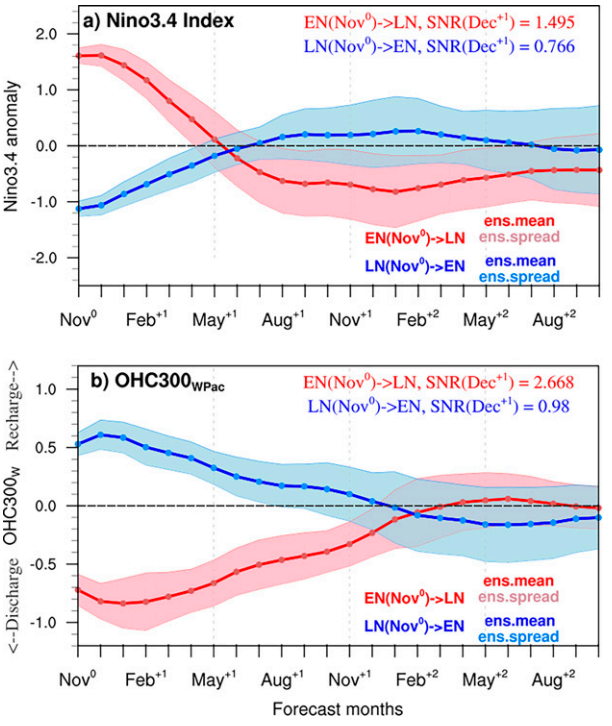

FIG. 9. Ensemble mean (solid curve) and spread (±1 standard deviation; shaded) of Nov^0 initialized forecasts for all transition years and all ensemble members from El Niño (Nov^0) to La Niña (Nov^{+1}; red) and La Niña (Nov^0) to El Niño (Nov^{+1}; blue). The ensemble spread (±1 standard deviation) from all individual members of subset EN → LN (LN → EN) events for Nov^0 forecasts is indicated in pink (blue) shading. The Nov^{+1} Niño-3.4 (12-month lead) signal-to-noise ratios (SNR) for both transitions are shown at the top-right corner of each plot.

(OHC_{EQ}; averaged over 5°S–5°N, 120°–280°E) using ACC scores for El Niño and La Niña initialized forecasts (Fig. 8). It is evident that both quantities exhibit a clear asymmetry regarding prediction skill, with forecasts initialized during peak El Niño providing significant and skillful predictions beyond 1.5 years, compared to La Niña initialized forecasts (up to 1 year), consistent with Fig. 2a. The figure also shows that the BSPB in the first year is mostly apparent in the OHC_W in forecasts initialized from peak La Niña conditions, a feature that may be related to the BSPB in SST predictions. The OHC_{EQ} does not exhibit any obvious BSPB in the first year, which is likely related with the subsequent skill recovery in SST (e.g., Balmaseda et al. 1995).

We further assess the SST and OHC forecasts of the transition subsets by using diagnostics of the SNR, which is an important factor in determining variations in predictability and prediction skill of the ENSO cycle following Larson and Pegion (2020). By applying this diagnostic to the EN → LN and LN → EN, we aim at quantifying if the forecasts show different degree of predictability, such as noticeable differences in the SNR (we remind the reader that comparing the skill independently for different subsets is not possible, given the limited number of cases in some of them). Figure 9 shows the ensemble mean (solid curves) of Niño-3.4 SST (Fig. 9a) and OHC_W anomaly (Fig. 9b), as well as ensemble spread (±1 standard deviation; shaded) as a function of lead months for the Nov^0 initialized forecasts subsets of EN → LN (in red) and LN → EN (in blue). The figures also show the SNR in Dec^{+1}, for easier comparison with Larson and Pegion (2020). The first noticeable feature is the larger SNR in OHC_W (2.7 vs 1.5 and 1.0 vs 0.8) compared with SST, a clear indication of the longer predictability of this variable. The second feature is that the SNR is consistently larger in EN → LN than in LN → EN subsets, in terms of both SST (1.5 vs 0.8) and OHC_W (2.7 vs 1.0), with the asymmetry between subsets being more noticeable in OHC_W than in SST. We also note that the differences in subsets are also present at the initial state, with larger amplitudes of SST and OHC_W anomalies in the LN → EN subset. The different estimation of predictability in our reforecast contrasts with the results by Larson and Pegion (2020), who did not find any indication of increased predictability of the EN → LN cases by applying the SNR diagnostic to a different set of seasonal reforecasts spanning a shorter period. One possible explanation is that our reforecasts allow a greater number of independent cases compared to Larson and Pegion (2020); also, they used predictions from multiple models.

To corroborate the relative role of OHC_W in affecting ENSO predictability, Fig. 10 illustrates the phase-space diagrams of observations (top panels) and forecast (Nov^0-initialized; bottom panels) trajectories in the (OHC_W, OHC_{EQ}) phase space. We use the normalized OHC anomalies of the composites for the different subsets. The left panels are for the EN_All and EN → LN subsets (Figs. 10a,c), while the corresponding diagrams for LN_All and LN → EN are shown in Figs. 10b and 10d. In these diagrams, the ENSO-related OHC cycle follows clockwise trajectories, where the x axis (y axis) indicates recharge/discharge of OHC_W (OHC_{EQ}). The observed trajectory during the preceding year (Nov^{-1}–Nov^0) is shown as gray dots in the observational panels (Figs. 10a,b).

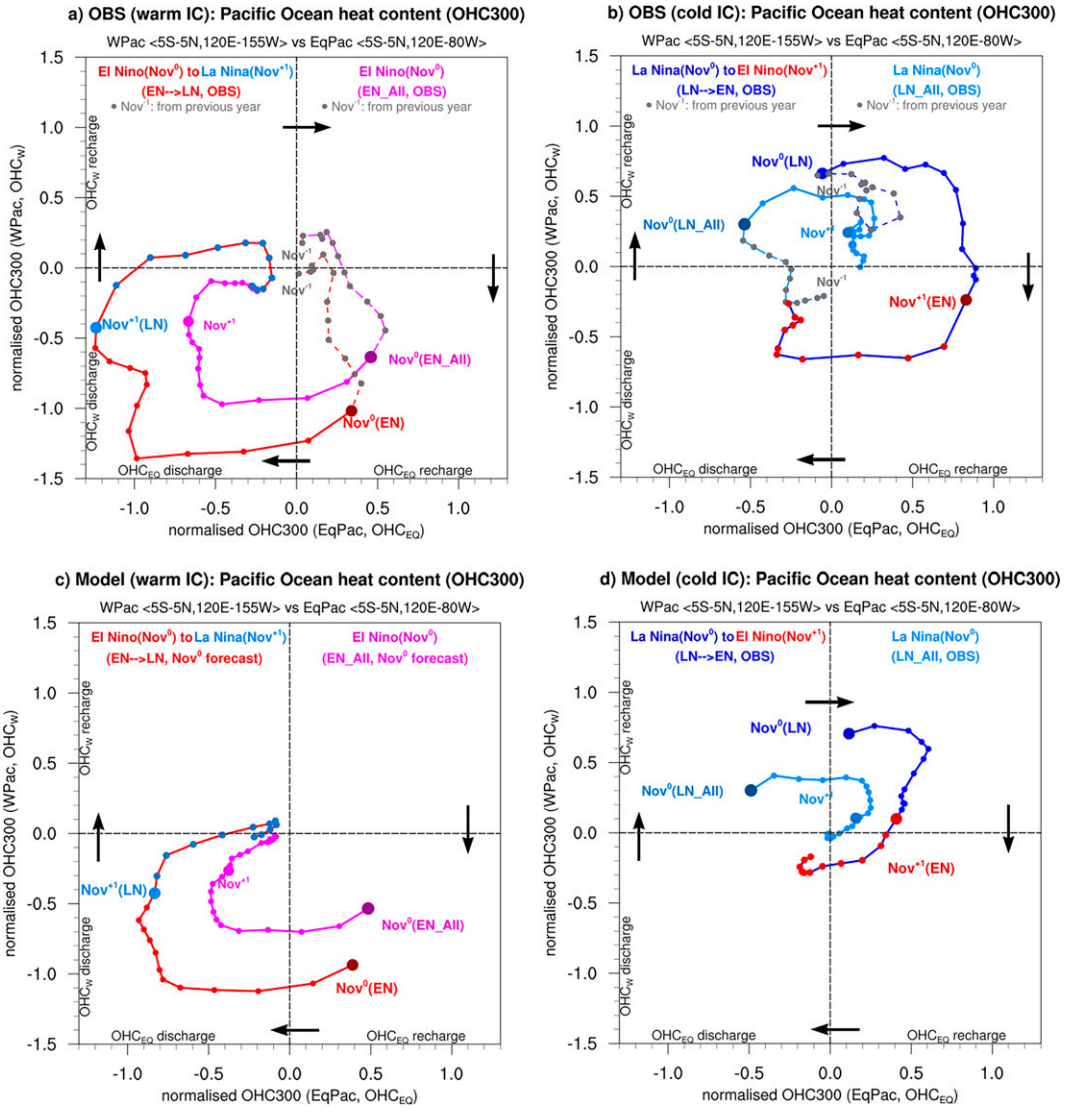

FIG. 10. Phase portrait diagram of entire equatorial Pacific OHC300 (OHC_{EQ}; 5°S–5°N, 120°–280°E) vs western Pacific OHC300 (OHC_W; 5°S–5°N, 120°–205°E) for warm/cold ENSO initial conditions–all El Niño (EN_All, Nov^0; magenta curve) and El Niño to La Niña (EN → LN, Nov^0; red curve) and all La Niña (LN_All, Nov^0; cyan curve) and La Niña to El Niño (LN → EN Nov^0; blue curve) from (top) observations and (bottom) the ensemble-mean Nov^0 initialized forecasts. The OHC anomalies are normalized by observed standard deviation at each lead month.

The observed trajectories of EN_All and EN → LN (magenta and red curves in Fig. 10a), share similar precursor in Nov^{-1} (gray dots), with a weak positive heat anomaly in both the western and equatorial regions a year prior to the El Niño (Nov^0; red dots). From Nov^{-1} to Nov^0 the two trajectories diverge, with a gradual increase (decrease) in OHC_{EQ}(OHC_W) evident in both but it is stronger in the EN → LN subset. A strong OHC_W discharge (peak around Mar^{+1}) is observed following El Niño mature phase (Nov^0–May^{+1}), while OHC_{EQ} also swings to discharge state (i.e., the trajectory moves toward the lower-left quadrant) until reaching the cold phase the following year. Until Nov^{+1}, OHC_W remains in a discharged state and starts to stabilize close to neutral afterward, while the eastern Pacific starts gaining heat. The observed OHC phase space diagram clearly shows the difference between EN_All (magenta curve) and EN → LN (red curve), in terms of initial conditions and evolution: the initial conditions for EN → LN have a more discharged state in the whole Pacific, especially noticeable in the western side. This more discharged state places the system in an outer orbit in the two dimensional phase space. The fact that the observed trajectories for EN_All and LN_All evolve in parallel orbits is indicative of the memory of the system. The model

forecasts in Fig. 10c are able to exploit the memory of the initial conditions, showing also parallel orbits for more than 1 year, which start converging (decaying) well after Nov^{+1}.

The precursors for the observed trajectories for LN_All and LN → EN (light and dark blue curves in Fig. 10b) show differences already in Nov^{-1} (gray dots): a recharged state for the LN → EN state and a weakly discharged one for LN_All for both OHC_W and OHC_{EQ}. During the subsequent year, the LN_All gradually moves from the weekly discharged state to recharged OHC_W, while increasing the discharge in the OHC_{EQ}, which reaches its minimum by the time the peak La Niña conditions (Nov^0), when the trajectory is clearly in the upper-left quadrant with a sizable discharged OHC_{EQ}, This contrasts with the Nov^0 initial conditions for the LN → EN case, where the trajectory is neutral in OHC_{EQ}, but visibly recharged in the western Pacific. So, the OHC_{EQ} appears as a distinctive feature for the transition LN → EN, which is favored by not having a heat deficit in equatorial Pacific. After Nov^0 both trajectories increase OHC_{EQ} from Nov^0 to May^{+1}, by approximately similar amounts, bringing the OHC_{EQ} to weakly recharged for LN_All and strongly recharged for LN → EN. After that, LN_All eventually decay toward the origin, while the LN → EN trajectory continues clockwise in a large orbit, transitioning to the El Niño state by Nov^{+1} (with fully recharged OHC_{EQ} but deficit in the OHC_W), and terminating the El Niño cycle by Nov^{+2}. The forecast counterparts are shown in the lower right panels. The forecast ensemble mean trajectories are similar to the observations until May^{+1}. After that, both LN → EN and LN_All decay toward zero, in contrast with the observations, where the LN → EN trajectory continued in its quasi-circular orbit. By the Nov^{+1} time the LN → EN are still distinguishable from the LN_All trajectory but compared to the observations it has lost energy (orbit with smaller radius) and has slowed down (failed to amplify/retain the OHC_{EQ} and not enough discharge of heat on the western Pacific). In summary, the difference between the transition subsets (EN → LN and LN → EN) appears clearly distinctive in the (OHC_W, OHC_{EQ}) phase space, evolving in outer orbits (e.g., more energetic) than the nontransition events. For the EN → LN transition, this difference appears in the Nov^0 initial conditions for those peak El Niño states where the negative OHC_W anomaly exceeds one standard deviation of the interannual variability, and it is well maintained by the model forecasts beyond 1 year. For LN → EN transition the positive anomaly in OHC_W is also larger than for the nontransition events, but its amplitude does not exceed the 1 standard deviation, and the forecast trajectory decays substantially faster than in observations after May^{+1}.

The OHC phase diagram illustrates the asymmetry between the peak states of EN and LN transition subsets in terms of OHC_W, with negative values during peak in El Niño twice as large as the positive values in peak La Niña. This corroborates the role of large OHC_W discharge during the peak of some El Niño events as a key component of the increased predictability of La Niña next year, which can be explained simply by energy conservation arguments: it will take stronger perturbations to put the system into a substantially different energy level. This is consistent with a recent modeling study using perfect model framework (Planton et al. 2021) that suggests discharged western Pacific initial conditions evolve more predictably into a neutral or La Niña state than recharged conditions into a neutral or El Niño state, highlighting a potential asymmetry in the link between preconditioning and ENSO predictability. Our forecasts are able to capture the increased predictability and translate it into actual prediction skill.

The OHC phase diagram also provides an interesting perspective of the phase locking of ENSO to the seasonal cycle: we note that changes of direction in the OHC_W tend to happen in spring tends to be zero in spring (e.g., OHC_W starts recharging in the lower-left quadrant by May^{+1} from peak Nov El Niño states, and it starts discharging in the upper-right quadrant in May^{+1} from peak La Niña states). In this diagram (consistent with Fig. 9), the spring predictability barrier is only visible in the LN → EN forecasts (upper-left quadrant), while for the EN → LN subset the forecasts do not seem to suffer that barrier. There are several reasons that can explain that behavior, and we refer to DiNezio and Deser (2014) and DiNezio et al. (2017b) for discussion on mechanisms affecting the stability of the El Niño and La Niña states, and therefore their sensitivity to stochastic perturbations.

4. Summary and perspectives

Skillful prediction of ENSO and its related climate impacts at lead times beyond a year is crucial for effective management of climate disasters. However, predicting ENSO skillfully beyond a year remains a major challenge. Theoretically, ENSO is believed to be predictable beyond 2–3 years from the self-sustained nature of the tropical Pacific coupled ocean–atmosphere system. Despite the significant progress in predicting ENSO evolution several seasons in advance using dynamical models, predictions of ENSO beyond the 1-yr lead time remain unavailable operationally. A handful of recent studies indicate the potential of predicting ENSO beyond 1 year using experimental long-lead retrospective forecasts (e.g., Luo et al. 2008, 2017; Planton et al. 2021; Park et al. 2018; Wu et al. 2021a), longer than the typical 6–9-month lead time practiced at many operational centers. These studies also point out that the skill at these long lead times may be conditioned by the phase of ENSO.

Although there are growing demands for knowledge of the likelihood of El Niño or La Niña beyond 1 year ahead, we have limited consensus about a potential dependence of long-lead ENSO predictability on initial ENSO state in a real-time prediction scenario (Planton et al. 2021). Operational ENSO prediction centers use relatively short reforecasts periods (~30 years), and although initialized on a monthly basis, they do not cover enough samples of independent El Niño or La Niña events to assess the flow-dependent reliability of the long-term ENSO predictions. In this study, we explore ENSO predictability using the 110 years of 24-month-long 10-member ensemble retrospective forecasts (or hindcasts) from ECMWF's coupled model initialized on 1 November and 1 May with initial conditions provided from CERA-20C for the period 1901–2010. Even though the reforecast dataset contains a limited number of ensemble members per initial date (10) and are initialized

only twice a year, they offer the advantage of including more samples of independent ENSO events. Our analyses show that the model has a high predictive skill of ENSO events (based on Niño-3.4 SST anomalies) with a lead time of ~18 months when initialized on 1 November. The prediction skill drops to about 12 lead months when initialized on 1 May because the second year's boreal spring predictability barrier (BSPB) is encountered. ENSO prediction skill may also be affected by the substantial multidecadal variation and the best skill is obtained during the recent 30 years (1981–2010; Weisheimer et al. 2022); thus, the assessment of ENSO prediction skill may be sensitive to the selection of appropriate hindcast period.

A systematic investigation of the general capability of the ECMWF model to simulate and predict ENSO and the controlling processes reveals some insights that might be valuable for developing future operational long-lead prediction systems. Forecasts initialized on 1 November from an El Niño state predict well the temporal evolution of the El Niño event from peak to termination, as well as evolving into La Niña condition in the subsequent year, thus providing prediction beyond 1.5 years. Forecasts initialized around the peak of La Niña (1 November start) can skillfully predict temporal evolution of Niño-3.4 anomalies and related physical processes up to ~10 lead months, at which time ENSO generally becomes neutral. Moreover, predicting the transition from La Niña to El Niño in the subsequent months beyond the first year is difficult as the model encounters the BSPB, which seems to be stronger for the transition to El Niño than for the transition to La Niña. Our results confirm the asymmetry in the long-lead prediction skill between El Niño and La Niña initialized forecasts, where forecasts initialized on 1 November at El Niño states have better skill at lead times beyond 1 year than those initialized at La Niña states. Hence, in the second year it is easier to predict the development of a La Niña than the transition to El Niño.

The enhanced skill for the transition from El Niño to La Niña appears related to the existing different preconditioned OHC in the western (rather than equatorial) Pacific states, that is, a stronger ocean heat content deficit over the western equatorial Pacific following peak El Niño compared to the heat gain following peak La Niña. Thus, the large deficit of heat in the western Pacific is a key precursor for predicting La Niña in the second year. Our results agree with those of Planton et al. (2018, 2021), who indicated the role of oceanic preconditioning for higher 1-yr lead predictability when starting from discharged rather than from recharged initial states from analyses of CMIP5 models and perfect-model experiments.

In addition to the ocean heat content in the western Pacific, the heat storage in the equatorial Pacific also appears as a distinctive feature to enable transition from La Niña to El Niño, which is favored by not having a heat deficit in equatorial Pacific. But for the predictions starting from peak La Niña, the BSPB is more apparent in forecasts starting with a recharged state and evolving toward El Niño than in those starting from a discharged state evolving toward La Niña, consistent with the discussion offered in DiNezio and Deser (2014) and DiNezio et al. (2017b).

Furthermore, our results suggest that the prediction of ENSO in the second year in the ECMWF system may be further improved by correcting the representation of key processes: the shorter duration of the cold events and the weak discharge during the warm events; and the correct level of wind variability at interannual time scales. These processes are likely related to the model systematic tendency of producing a westward shift of the cold tongue and easterly wind bias, as discussed by Wu et al. (2022). We note that these errors in the second-year forecasts are consistent with those of the coupled ECMWF model versions in long coupled preindustrial integrations, as reported by Roberts et al. (2018). Roberts et al. (2018) also noted more symmetric distributions of warm and cool event durations compared to the observed, which is consistent with the findings in this study using initialized reforecasts. It was speculated that the ECMWF model might have a tendency to underestimate the persistence of longer-lasting warm and cool events (>36–42 months) in the tropical Pacific. The challenge of simulating multiyear La Niña events is especially prominent at this current time when the third boreal winter La Niña season in a row has been announced and the ECMWF SEAS5 predictions failed the prolonged cold conditions in the previous years. Therefore, more research should prioritize reducing model biases, and increasing process-oriented understanding of climate predictability that could lead to enhanced predictive skill and higher degree of confidence in future real-time operational forecasts. Long reforecast records of 24-month-long predictions constitute an invaluable test bed for further model developments and understanding of ENSO predictability, mechanisms, and model errors. Since the samples of El Niño and La Niña events used in this study includes both 1-yr and multiyear events, we will further assess the SEAS5-20C's ability to predict multiyear ENSO events using these same 110-yr hindcasts following Wu et al. (2021a) in a future paper. Regardless of the caveats discussed above, this comprehensive study provides crucial insights into some fundamental issues of long-lead ENSO predictability and dependence of ENSO prediction skill on the initial state, which will be valuable for developing operational long-lead ENSO forecasts beyond 1 year.

Acknowledgments. We thank the editor Prof Y. Okumura, Dr. Sarah Larson, and three anonymous reviewers for constructive comments that helped improve the quality of the manuscript. This research contributes to the Northern Australia Climate Program (NACP) project, funded by the Meat and Livestock Australia (MLA), the Queensland Government through the Drought and Climate Adaptation Program (DCAP), and the University of Southern Queensland (UniSQ). The authors thank Drs. Eun-Pa Lim, Matthew Wheeler, and David Jones for their helpful internal feedback at the Bureau. SS also thanks Dr. S. Abhik for initial draft feedback and useful discussion on significance tests of skill difference. The assistance of resources from the National Computational Infrastructure (NCI) supported by the Australian Government is acknowledged. The authors have no conflicts of interest to declare.

Data availability statement. All observational and reanalysis data are publicly available. The hindcast datasets used in

this study are available from https://doi.org/10.21957/fzf9-te33 (ECMWF 2021). The data in this study were analyzed and plotted using NCAR Command Language V6.6.2 (www.ncl.ucar.edu).

REFERENCES

Alves, O., M. A. Balmaseda, D. Anderson, and T. Stockdale, 2004: Sensitivity of dynamical seasonal forecasts to ocean initial conditions. *Quart. J. Roy. Meteor. Soc.*, **130**, 647–667, https://doi.org/10.1256/qj.03.25.

Balmaseda, M. A., and D. Anderson, 2009: Impact of initialization strategies and observations on seasonal forecast skill. *Geophys. Res. Lett.*, **36**, L01701, https://doi.org/10.1029/2008GL035561.

——, M. K. Davey, and D. L. T. Anderson, 1995: Decadal and seasonal dependence of ENSO prediction skill. *J. Climate*, **8**, 2705–2715, https://doi.org/10.1175/1520-0442(1995)008<2705:DASDOE>2.0.CO;2.

Barnston, A. G., M. K. Tippett, M. L. L'Heureux, S. Li, and D. G. Dewitt, 2012: Skill of real-time seasonal ENSO model predictions during 2002-11: Is our capability increasing? *Bull. Amer. Meteor. Soc.*, **93**, 631–651, https://doi.org/10.1175/BAMS-D-11-00111.1.

——, ——, M. Ranganathan, and M. L. L'Heureux, 2019: Deterministic skill of ENSO predictions from the North American Multimodel Ensemble. *Climate Dyn.*, **53**, 7215–7234, https://doi.org/10.1007/s00382-017-3603-3.

Bretherton, C. S., M. Widmann, V. P. Dymnikov, J. M. Wallace, and I. Bladé, 1999: The effective number of spatial degrees of freedom of a time-varying field. *J. Climate*, **12**, 1990–2009, https://doi.org/10.1175/1520-0442(1999)012<1990:TENOSD>2.0.CO;2.

Cane, M. A., S. E. Zebiak, and S. C. Dolan, 1986: Experimental forecasts of El Niño. *Nature*, **321**, 827–832, https://doi.org/10.1038/321827a0.

Chakravorty, S., R. C. Perez, B. T. Anderson, B. S. Giese, S. M. Larson, and V. Pivotti, 2020: Testing the trade wind charging mechanism and its influence on ENSO variability. *J. Climate*, **33**, 7391–7411, https://doi.org/10.1175/JCLI-D-19-0727.1.

Chen, D., and M. A. Cane, 2008: El Niño prediction and predictability. *J. Comput. Phys.*, **227**, 3625–3640, https://doi.org/10.1016/j.jcp.2007.05.014.

——, ——, A. Kaplan, S. E. Zebiak, and D. Huang, 2004: Predictability of El Niño over the past 148 years. *Nature*, **428**, 733–736, https://doi.org/10.1038/nature02439.

DiNezio, P. N., and C. Deser, 2014: Nonlinear controls on the persistence of La Niña. *J. Climate*, **27**, 7335–7355, https://doi.org/10.1175/JCLI-D-14-00033.1.

——, ——, A. Karspeck, S. Yeager, Y. Okumura, G. Danabasoglu, and G. A. Meehl, 2017a: A 2 year forecast for a 60–80% chance of La Niña in 2017–2018. *Geophys. Res. Lett.*, **44**, 11 624–11 635, https://doi.org/10.1002/2017GL074904.

——, ——, Y. Okumura, and A. Karspeck, 2017b: Predictability of 2-year La Niña events in a coupled general circulation model. *Climate Dyn.*, **49**, 4237–4261, https://doi.org/10.1007/s00382-017-3575-3.

Dommenget, D., T. Bayr, and C. Frauen, 2013: Analysis of the non-linearity in the pattern and time evolution of El Niño southern oscillation. *Climate Dyn.*, **40**, 2825–2847, https://doi.org/10.1007/s00382-012-1475-0.

ECMWF, 2021: Global monthly reforecast sea surface temperatures from 1901 to 2010 with 24-month lead time. ECMWF, https://apps.ecmwf.int/research-experiments/expver/guxf/.

Gonzalez, P. L. M., and L. Goddard, 2016: Long-lead ENSO predictability from CMIP5 decadal hindcasts. *Climate Dyn.*, **46**, 3127–3147, https://doi.org/10.1007/s00382-015-2757-0.

Hendon, H. H., E. Lim, G. Wang, O. Alves, and D. Hudson, 2009: Prospects for predicting two flavors of El Niño. *Geophys. Res. Lett.*, **36**, L19713, https://doi.org/10.1029/2009GL040100.

Huang, B., and Coauthors, 2017: Extended Reconstructed Sea Surface Temperature, version 5 (ERSSTv5): Upgrades, validations, and intercomparisons. *J. Climate*, **30**, 8179–8205, https://doi.org/10.1175/JCLI-D-16-0836.1.

Hudson, D., and Coauthors, 2017: ACCESS-S1 The new Bureau of Meteorology multi-week to seasonal prediction system. *J. South. Hemisphere Earth Syst. Sci.*, **67**, 132–159, https://doi.org/10.1071/ES17009.

Izumo, T., M. Lengaigne, J. Vialard, I. Suresh, and Y. Planton, 2019: On the physical interpretation of the lead relation between warm water volume and the El Niño Southern Oscillation. *Climate Dyn.*, **52**, 2923–2942, https://doi.org/10.1007/s00382-018-4313-1.

Jin, E. K., and Coauthors, 2008: Current status of ENSO prediction skill in coupled ocean–atmosphere models. *Climate Dyn.*, **31**, 647–664, https://doi.org/10.1007/s00382-008-0397-3.

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

Johnson, S. J., and Coauthors, 2019: SEAS5: The new ECMWF seasonal forecast system. *Geosci. Model Dev.*, **12**, 1087–1117, https://doi.org/10.5194/gmd-12-1087-2019.

Kessler, W. S., 2002: Is ENSO a cycle or a series of events? *Geophys. Res. Lett.*, **29**, 2125, https://doi.org/10.1029/2002GL015924.

Kirtman, B. P., and P. S. Schopf, 1998: Decadal variability in ENSO predictability and prediction. *J. Climate*, **11**, 2804–2822, https://doi.org/10.1175/1520-0442(1998)011<2804:DVIEPA>2.0.CO;2.

——, and Coauthors, 2014: The North American multimodel ensemble: Phase-1 seasonal-to-interannual prediction; phase-2 toward developing intraseasonal prediction. *Bull. Amer. Meteor. Soc.*, **95**, 585–601, https://doi.org/10.1175/BAMS-D-12-00050.1.

Laloyaux, P., and Coauthors, 2018: CERA-20C: A coupled reanalysis of the twentieth century. *J. Adv. Model. Earth Syst.*, **10**, 1172–1195, https://doi.org/10.1029/2018MS001273.

Larson, S. M., and B. P. Kirtman, 2017: Drivers of coupled model ENSO error dynamics and the spring predictability barrier. *Climate Dyn.*, **48**, 3631–3644, https://doi.org/10.1007/s00382-016-3290-5.

——, and K. Pegion, 2020: Do asymmetries in ENSO predictability arise from different recharged states? *Climate Dyn.*, **54**, 1507–1522, https://doi.org/10.1007/s00382-019-05069-5.

Latif, M., T. P. Barnett, M. A. Cane, M. Flügel, N. E. Graham, H. von Storch, J. S. Xu, and S. E. Zebiak, 1994: A review of ENSO prediction studies. *Climate Dyn.*, **9**, 167–179, https://doi.org/10.1007/BF00208250.

Luo, J.-J., S. Masson, S. K. Behera, and T. Yamagata, 2008: Extended ENSO predictions using a fully coupled ocean–atmosphere model. *J. Climate*, **21**, 84–93, https://doi.org/10.1175/2007JCLI1412.1.

——, J.-Y. Lee, C. Yuan, W. Sasaki, S. Masson, S. Behera, Y. Masumoto, and T. Yamagata, 2015: Current status of

intraseasonal–seasonal-to-interannual prediction of the Indo-Pacific climate. *Indo-Pacific Climate Variability and Predictability*, S. K. Behera and T. Yamagata, Eds., World Scientific, 63–107.

——, G. Liu, H. Hendon, O. Alves, and T. Yamagata, 2017: Inter-basin sources for two-year predictability of the multi-year La Niña event in 2010–2012. *Sci. Rep.*, **7**, 2276, https://doi.org/10.1038/s41598-017-01479-9.

McPhaden, M. J., A. Santoso, and W. Cai, 2020: *El Niño Southern Oscillation in a Changing Climate*. *Geophys. Monogr.*, Vol. 253, Amer. Geophys. Union, 528 pp., https://doi.org/10.1002/9781119548164.

Meinen, C. S., and M. J. McPhaden, 2000: Observations of warm water volume changes in the equatorial Pacific and their relationship to El Niño and La Nina. *J. Climate*, **13**, 3551–3559, https://doi.org/10.1175/1520-0442(2000)013<3551:OOWWVC>2.0.CO;2.

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

Park, J. H., J. S. Kug, T. Li, and S. K. Behera, 2018: Predicting El Niño beyond 1-year lead: Effect of the Western Hemisphere warm pool. *Sci. Rep.*, **8**, 14957, https://doi.org/10.1038/s41598-018-33191-7.

Philander, S. G., 1990: *El Niño, La Niña, and the Southern Oscillation*. Academic Press, 293 pp.

Pivotti, V., and B. T. Anderson, 2021: Transition between forced and oscillatory ENSO behavior over the last century. *J. Geophys. Res. Atmos.*, **126**, e2020JD034116, https://doi.org/10.1029/2020JD034116.

Planton, Y., J. Vialard, E. Guilyardi, M. Lengaigne, and T. Izumo, 2018: Western Pacific oceanic heat content: A better predictor of La Niña than of El Niño. *Geophys. Res. Lett.*, **45**, 9824–9833, https://doi.org/10.1029/2018GL079341.

——, ——, ——, ——, and M. J. McPhaden, 2021: The asymmetric influence of ocean heat content on ENSO predictability in the CNRM-CM5 coupled general circulation model. *J. Climate*, **34**, 5775–5793, https://doi.org/10.1175/JCLI-D-20-0633.1.

Plummer, N., and Coauthors, 2018: Climate services: For informing decisions and managing risk. *Bridging Science and Policy Implication for Managing Climate Extremes*, World Scientific, 49–64, https://doi.org/10.1142/9789813235663_0004.

Poli, P., and Coauthors, 2016: ERA-20C: An atmospheric reanalysis of the twentieth century. *J. Climate*, **29**, 4083–4097, https://doi.org/10.1175/JCLI-D-15-0556.1.

Roberts, C. D., R. Senan, F. Molteni, S. Boussetta, M. Mayer, and S. P. E. Keeley, 2018: Climate model configurations of the ECMWF Integrated Forecasting System (ECMWF-IFS cycle 43r1) for HighResMIP. *Geosci. Model Dev.*, **11**, 3681–3712, https://doi.org/10.5194/gmd-11-3681-2018.

Tang, Y., and Coauthors, 2018: Progress in ENSO prediction and predictability study. *Natl. Sci. Rev.*, **5**, 826–839, https://doi.org/10.1093/nsr/nwy105.

Taschetto, A. S., C. C. Ummenhofer, M. F. Stuecker, D. Dommenget, K. Ashok, R. R. Rodrigues, and S.-W. Yeh, 2020: ENSO atmospheric teleconnections. *El Niño Southern Oscillation in a Changing Climate*, *Geophys. Monogr.*, Vol. 253, Amer. Geophys. Union, 309–335, https://doi.org/10.1002/9781119548164.ch14.

Timmermann, A., and Coauthors, 2018: El Niño–Southern Oscillation complexity. *Nature*, **559**, 535–545, https://doi.org/10.1038/s41586-018-0252-6.

Titchner, H. A., and N. A. Rayner, 2014: The Met Office Hadley Centre sea ice and sea surface temperature data set, version 2: 1. Sea ice concentrations. *J. Geophys. Res. Atmos.*, **119**, 2864–2889, https://doi.org/10.1002/2013JD020316.

Vecchi, G. A., and D. E. Harrison, 2000: Tropical Pacific sea surface temperature anomalies, El Niño, and equatorial westerly wind events. *J. Climate*, **13**, 1814–1830, https://doi.org/10.1175/1520-0442(2000)013<1814:TPSSTA>2.0.CO;2.

Vimont, D. J., J. M. Wallace, and D. S. Battisti, 2003: The seasonal footprinting mechanism in the Pacific. *J. Climate*, **16**, 2668–2675, https://doi.org/10.1175/1520-0442(2003)016<2668:TSFMIT>2.0.CO;2.

Weisheimer, A., D. J. Befort, D. MacLeod, T. Palmer, C. O'Reilly, and K. Strømmen, 2020: Seasonal forecasts of the twentieth century. *Bull. Amer. Meteor. Soc.*, **101**, E1413–E1426, https://doi.org/10.1175/BAMS-D-19-0019.1.

——, M. Balmaseda, T. Stockdale, M. Mayer, E. de Boisseson, R. Senan, and S. Johnson, 2021: Retrospective two-year ENSO predictions during the 20th century. *ECMWF Newsletter*, No. 169, Reading, United Kingdom, 7–8, https://www.ecmwfint/sites/default/files/elibrary/2021/20225-newsletter-no-169-autumn-2021.pdf.

——, M. A. Balmaseda, T. N. Stockdale, M. Mayer, S. Sharmila, H. Hendon, and O. Alves, 2022: Variability of ENSO forecast skill in 2-year global reforecasts over the 20th century. *Geophys. Res. Lett.*, **49**, e2022GL097885, https://doi.org/10.1029/2022GL097885.

Wu, X., Y. M. Okumura, C. Deser, and P. N. DiNezio, 2021a: Two-year dynamical predictions of ENSO event duration during 1954–2015. *J. Climate*, **34**, 4069–4087, https://doi.org/10.1175/JCLI-D-20-0619.1.

——, ——, and P. N. DiNezio, 2021b: Predictability of El Niño duration based on the onset timing. *J. Climate*, **34**, 1351–1366, https://doi.org/10.1175/JCLI-D-19-0963.1.

——, ——, ——, S. G. Yeager, and C. Deser, 2022: The equatorial Pacific cold tongue bias in CESM1 and its influence on ENSO forecasts. *J. Climate*, **35**, 3261–3277, https://doi.org/10.1175/JCLI-D-21-0470.1.

Xue, Y., M. Chen, A. Kumar, Z. Z. Hu, and W. Wang, 2013: Prediction skill and bias of tropical Pacific sea surface temperatures in the NCEP Climate Forecast System version 2. *J. Climate*, **26**, 5358–5378, https://doi.org/10.1175/JCLI-D-12-00600.1.

Zhao, M., H. H. Hendon, O. Alves, G. Liu, and G. Wang, 2016: Weakened eastern Pacific El Niño predictability in the early twenty-first century. *J. Climate*, **29**, 6805–6822, https://doi.org/10.1175/JCLI-D-15-0876.1.