The equatorial Pacific cold tongue bias in CESM1 and its influence on ENSO forecasts

Xian Wu, a Yuko M. Okumura, b Pedro N. DiNezio, c Stephen G. Yeager, a and Clara Deser a

a Climate and Global Dynamics Division, National Center for Atmospheric Research, Boulder, Colorado

b Institute for Geophysics, Jackson School of Geosciences, The University of Texas at Austin, Austin, Texas

c Department of Atmospheric and Oceanic Sciences, University of Colorado Boulder, Boulder, Colorado

Journal of Climate
Submitted June XX, 2021

Corresponding author address: Dr. Xian Wu, Climate and Global Dynamics Division, National Center for Atmospheric Research, 1850 Table Mesa Drive, Boulder, Colorado, 80305

Email: xianwu@ucar.edu
Abstract

The mean-state bias and the associated forecast errors of the El Niño-Southern Oscillation (ENSO) are investigated in a suite of two-year retrospective forecasts conducted with the Community Earth System Model, version 1. The equatorial Pacific cold tongue in the forecasts is too strong and extends excessively westward due to a combination of the model’s inherent climatological bias, initial state-model imbalance, and errors in the oceanic data used to initialize the forecasts. The forecasts show a stronger cold tongue bias in the first year than that inherent to the model due to the imbalance between initial subsurface oceanic states and model dynamics. The cold tongue bias affects not only the pattern and amplitude but also the duration of ENSO in the forecasts. The predicted sea surface temperature anomalies related to ENSO extend to the far western equatorial Pacific during boreal summer when the cold tongue bias is strong, and ENSO anomalies are too weak in the central-eastern equatorial Pacific. The forecast errors of pattern and amplitude subsequently lead to errors in ENSO phase transition by affecting the amplitude of the negative thermocline feedback in the equatorial Pacific and tropical interbasin adjustments during the mature phase of ENSO. These ENSO forecast errors further degrade the predictions of atmospheric teleconnections and rainfall anomalies over the Northern Hemisphere. These mean-state and ENSO-related forecast biases are more pronounced in forecasts initialized in boreal spring and summer than other seasons due to the seasonal intensification of the Bjerknes feedback.
1. Introduction

The tropical Pacific Ocean exhibits zonal contrast in mean sea surface temperature (SST) between the western warm pool and the eastern equatorial cold tongue. The eastern equatorial Pacific cold tongue controls climatological SST and rainfall patterns across a large area of the tropics. The cold tongue simulated by the successive generations of Coupled Model Intercomparison Project (CMIP) models is commonly too cold and extends excessively westward compared to observations (Mechoso et al. 1995; Davey et al. 2002; Zheng et al. 2012; Brown et al. 2014; Bellenger et al. 2014; Planton et al. 2020; Guilyardi et al. 2020; Jiang et al. 2021). This cold tongue SST bias in models has been attributed to the misrepresentations of dynamic and thermodynamic processes in the equatorial Pacific, including too strong equatorial zonal surface winds, too strong oceanic horizontal and vertical temperature advection, too little SST damping due to erroneously positive cloud-shortwave radiation feedback, and misrepresentations of subseasonal variability, such as tropical instability waves (Mechoso et al. 1995; Davey et al. 2002; Wittenberg et al. 2006; Zheng et al. 2012; Vannière et al. 2013; Ray et al. 2018; Siongco et al. 2020). The cold tongue SST bias in coupled models is also linked with the bias of the intertropical convergence zone (ITCZ) simulated in atmospheric models, which influences the surface wind bias in the equatorial Pacific (Lin 2007; de Szoeke and Xie 2008; Li and Xie 2014).

The climatological cold tongue bias has been suggested to affect the simulations of tropical climate variations, especially the leading mode of interannual climate variability – the El Niño-Southern Oscillation (ENSO). In association with the westward extension of the cold tongue, equatorial SST, surface wind, and precipitation variations related to ENSO events are shifted too far west in climate models compared to observations (Wittenberg et al. 2006; Taschetto et al. 2014; Graham et al. 2017; Planton et al. 2020; Jiang et al. 2021). This is primarily because the enhanced
climatological zonal temperature gradient induces a stronger zonal advective feedback in the western equatorial Pacific and favors the westward extension of SST anomalies associated with ENSO events (Graham et al. 2017; Jiang et al. 2021). Besides affecting the oceanic feedback, the cold tongue bias could shift the rising branch of Walker circulation westward and make the convection and wind more sensitive to the SST anomalies in the western equatorial Pacific than observed (i.e., too strong positive SST-wind feedback; Ham and Kug 2012; Bayr et al. 2018). The cold tongue bias could also act to reduce damping of ENSO SST anomalies in the western equatorial Pacific by weakening the negative cloud-shortwave radiation feedback (Graham et al. 2017; Bayr et al. 2018). The influence of mean-state bias on ENSO amplitude is complex and masked in many models due to the error compensation of ENSO feedbacks. For example, a model can still simulate realistic ENSO amplitude when underestimating both the positive wind-SST feedback and negative heat flux-SST feedback (Bellenger et al. 2014; Bayr et al. 2019a).

The cold tongue bias can also degrade the models’ ability to predict ENSO. Models with a cold tongue bias show errors in predicting the pattern of ENSO; in particular, they show negative correlation skill in their ENSO forecasts of SST anomalies in the western Pacific (Ham et al. 2014). Forecast errors of ENSO amplitude are also found to be related to the biases of SST and upper ocean temperature in the equatorial Pacific (Manganello and Huang et al. 2009; Kim et al. 2017). Besides the forecast errors of pattern and amplitude of ENSO events, a recent study by Wu et al. (2021a) notes that the low skill in predicting the phase transition of ENSO events, especially of El Niño events, is related to the strong cold tongue bias of the retrospective forecasts conducted with Community Earth System Model, version 1 (CESM1). However, the errors regarding the phase transition of El Niño are absent in the ‘perfect model’ prediction experiments performed with the
same model (Wu et al. 2020b) because such idealized forecasts are not affected by the issues that arise in the retrospective forecasts summarized below.

Errors in operational forecasts arise not only from the inherent model biases discussed above, but also from the initialization methods, including uncertainties in reproducing the true initial conditions and incompatibilities between the initial states and the model dynamics. When forecasts are initialized with full-field oceanic states derived from observations, the forecasts will drift toward the model’s own climatology as the forecasts progress (Misra et al. 2008; Magnusson et al. 2013). In addition to the gradual model drift, the full-field initialization may lead to a rapid initial adjustment of the model if there is a large imbalance between the prescribed initial conditions and the model dynamics, known as the initialization shock problem (Magnusson et al. 2013).

Understanding the causes of cold tongue bias and related ENSO forecast errors is important from the standpoint of seasonal prediction skill in ENSO-related tropical rainfall and extratropical teleconnections (Bayr et al. 2019b; Ding et al. 2020). In this paper, we analyze the climatological bias of the equatorial Pacific and its impact on ENSO forecasts in a suite of multi-year full-field initialized forecasts conducted with the CESM1. First, we examine in detail the origins of cold tongue bias in the CESM1 forecasts, including the inherent climatological bias of CESM1, imbalance between the initial conditions and model dynamics, and errors in the ocean initial conditions. Then we investigate the forecast errors in ENSO characteristics and teleconnections, and their relationship with the cold tongue bias of the forecasts and initial condition error. The rest of this paper is organized as follows. Section 2 describes the model, experiments, and analysis methods. Section 3 presents the analyses of cold tongue bias and ENSO errors in the forecasts and
the dynamical processes underlying these errors. Section 4 summarizes the results and discusses
the implications for improving the forecast skill of ENSO events.

2. Model experiments and analysis methods

a. CESM1

The CESM1 is a state-of-the-art climate model consisting of the atmosphere, ocean, land
surface, and sea ice components linked by a flux coupler (Hurrell et al. 2013). The atmospheric
component, the Community Atmosphere Model, version 5, uses a finite-volume dynamical core
at a horizontal resolution of 0.9° latitude × 1.25° longitude with 30 levels in vertical (Neale et al.
2012). The land model is the Community Land Model, version 4 (Lawrence et al. 2011) that is run
on the same horizontal grid as the atmosphere model. The oceanic component, the Parallel Ocean
Program, version 2 (POP2; Smith et al. 2010), has meridional resolutions increasing from 0.65° at
60°N to 0.27° at the equator and 60 levels in the vertical. The sea ice model, the Los Alamos
National Laboratory Community Ice Code, version 4 (CICE4; Hunke and Lipscomb 2008), uses
the same horizontal grid as the ocean model.

The CESM1 produces one of the most realistic simulations of the ENSO phenomenon
among global climate models (Bellenger et al. 2014). It reproduces the broad spectral peak of
ENSO in the 3–6-year band, the asymmetries in amplitude and duration between El Niño and La
Niña, and the wide range of durations of ENSO events (DiNezio et al. 2017a; Wu et al. 2019). The
CESM1 shows high skill in predicting the duration of El Niño and La Niña with lead times up to
two years when initialized from particular ocean states (DiNezio et al. 2017a,b; Wu et al. 2021a,b).
However, the free-running preindustrial control simulations of CESM1 overestimate the amplitude
of ENSO by about 20% and show an excessive extension of ENSO anomalies into the western
equatorial Pacific compared to observations (DiNezio et al. 2017; Capotondi et al. 2020). The
initialized forecasts of CESM1 show large biases in predicting ENSO pattern and duration especially when the climatological cold tongue bias is strong in the forecasts initialized in particular seasons (Wu et al. 2021a).

b. CESM1 retrospective forecasts and uninitialized simulations

Three sets of multiyear CESM1 ensemble forecasts initialized on March 1, June 1, and November 1 of each year from 1954 to 2015 are analyzed. The November-initialized forecasts are part of the CESM Decadal Prediction Large Ensemble (Yeager et al. 2018) and the ensembles initialized in March and June are from Wu et al. (2021a). For each year and calendar month, the ensemble forecasts are initialized with identical ocean and sea ice conditions, and ensemble spread is generated by adding small perturbations of an order of $10^{-14}$ to the atmospheric initial conditions. The ensembles initialized in March, June, and November have 10, 20, and 40 members and are integrated over 30, 27, and 34 months, respectively. The initialization months and forecast length were selected to investigate the predictability of ENSO events with lead times up to two years and the seasonal dependence of forecast skill, and an ensemble size of 10 was sufficient to estimate the ensemble mean signal of ENSO predictions (Wu et al. 2021a). The ocean and sea ice initial conditions for all ensembles were generated by forcing the ocean (POP2) and sea ice (CICE4) component models with observed atmospheric and surface flux fields, hereafter called a “forced ocean-sea ice simulation” (FOSI). The atmosphere and land initial conditions are obtained from the CESM1 Large Ensemble Project (CESM1 LE; Kay et al. 2015) for the November-initialized forecasts (see below) and from an atmosphere-land model (CAM5-CLM4) simulation forced with monthly ocean and sea ice fields from the FOSI for the March- and June-initialized forecasts. The different techniques for initializing atmosphere and land models are not expected affect the predictions of tropical Pacific SSTs, whose predictability is mostly governed by oceanic memory.

To assess the errors inherent to the model in simulating the mean state and ENSO variability, we make use of the 40-member CESM1 LE during the forecast period 1954–2015. The first member of the CESM1 LE is initialized with oceanic and atmospheric conditions of year 401 of the CESM1 preindustrial control simulation and integrated from 1850 to 2100 under the historical (prior to 2006) and RCP 8.5 (2006–2100) forcings. The other 39 members are branched from the first member on 1 January 1920 with small perturbations applied to the atmospheric initial conditions (order of $10^{-14}$ K).

c. Observational datasets

We compare the forecasts, uninitialized simulations, and oceanic initial conditions used for initializing the forecasts against several observational and reanalysis datasets. The SST dataset is the Hadley Centre Sea Ice and SST dataset (HadISST; Rayner et al. 2003) available for 1870–2019 on a 1° grid. The thermocline depth is defined as the depth of maximum vertical ocean temperature gradient derived based on European Centre for Medium-Range Weather Forecasts (ECMWF) Ocean Reanalysis System 4 (ORAS4; Balmaseda et al. 2013) available for 1958–2017 on a 1° grid with 42 levels in the vertical. Several other ocean reanalysis datasets are also used to validate the estimation of observed equatorial Pacific thermocline depth (Fig. S1), including the Simple Ocean Data Assimilation reanalysis, version 2.2.4 (SODA; Carton and Giese 2008), version 4 of the Met Office Hadley Centre ‘EN4’ (EN4; Good et al. 2013) with two different sets of bias corrections by Levitus et al. (2009; EN4I09) and Gouretski and Reseghetti (2010; EN4g10), Estimating the Circulation and Climate of the Ocean Version 4, Release 4 (ECCO4R4; Forget et al. 2015), global ocean data assimilation system (GODAS; Behringer and Xue 2004), and the global
objective analyses of ocean data from the International Comprehensive Ocean and Atmosphere Data Set and the Kobe Collection (ICOADS_Kobe; Ishii et al. 2005). For surface wind components, precipitation, sea level pressure (SLP), and 200 hPa geopotential height (Z200), we make use of the National Centers for Environmental Prediction-National Center for Atmospheric Research (NCEP-NCAR) reanalysis (Kalnay et al. 1996) available for 1948–2017 on a 2.5° grid. Observed monthly climatology is calculated for 1958–2015 and monthly anomalies are quadratically detrended. We regrid the observational data to the same grid as the model output before calculating their difference.

d. Analysis methods

The drifting climatology of the ensemble forecasts initialized at a particular calendar month is calculated by averaging the ensemble mean forecasts across the initialization years from 1958-2015 for each lead time. We define forecast anomalies as deviations from the drifting climatology. The effect of external radiative forcing is removed by subtracting a quadratic trend computed from the ensemble mean anomalies for each lead time. The observations are treated in a similar manner, with the monthly climatology calculated for 1958–2015, and quadratically detrending applied to the monthly anomalies. The monthly climatology inherent to the CESM1 is calculated using the 40-member ensemble mean of CESM1 LE for the forecast period of 1958-2015. The effect of radiative forcing is removed by subtracting the monthly ensemble mean from individual ensemble members.

The temporal evolution of ENSO events is tracked using the Niño-3.4 index defined as SST anomalies averaged over the Niño-3.4 region (170°W–120°W, 5°S–5°N). Following Wu et al. (2019; 2021a), observed ENSO events are defined when the absolute value of the Niño-3.4 index smoothed with a 3-month running mean filter is greater than 0.75 standard deviations in any
month from October to February. The year when an ENSO event first develops is denoted as year 0 and the months of the year as Jan$^0$, Feb$^0$, …, and Dec$^0$. We composite ensemble forecasts initialized in Nov$^{-1}$, Mar$^0$, Jun$^0$, Nov$^0$, Mar$^{+1}$, and Jun$^{+1}$ of observed ENSO events. ENSO events are further classified into 2-yr events when the smoothed absolute value of the Niño-3.4 index remains above 0.5 standard deviations in any month from Oct$^{+1}$ to Feb$^{+2}$, and otherwise into 1-yr events. During 1954-2017, we identify ten 1-yr El Niño (1963, 1965, 1972, 1982, 1991, 1994, 1997, 2002, 2006, and 2009), five 2-yr El Niño (1957, 1968, 1976, 1986, and 2014), four 1-yr La Niña (1964, 1988, 1995, and 2005), and eight 2-yr La Niña events (1954, 1970, 1973, 1983, 1998, 2007, 2010, and 2016), with the list of years indicating year 0. We define the onset and termination of ENSO events in the forecasts when the absolute value of Niño-3.4 first exceeds and drops below 0.5°C, respectively. Readers are referred to Wu et al. (2021a) for further details of the analysis methods. The ENSO variability in the CESM1 LE is evaluated based on the ENSO events identified using the same method as for observations applied to all 40 members during 1954-2015. We compare the climatology and composite ENSO events in the initialized CESM1 forecasts with those in observations and the CESM1 LE. The statistical significance of the composite ENSO anomalies is assessed using a two-tailed Student’s $t$-test at the confidence level of 95%.

3. Results

a. Climatological errors in the equatorial Pacific

The climatological seasonal cycle of SST, surface wind, and thermocline depth in the equatorial Pacific for the forecasts initialized in November, March, and June are compared with observations and CESM1 LE (Fig. 1). The three sets of forecasts capture the seasonal intensification of the SST cold tongue during boreal summer and fall (Figs. 1a-d). In both the forecasts and observations (Figs. 1a-d), the SST cooling intensifies in the eastern equatorial Pacific
when the equatorial southerly winds start to develop in late boreal spring to early summer (May–June) in association with the seasonal northward migration of the ITCZ (Mitchell and Wallace, 1992; Xie 1994). During summer and early autumn (June–September), the eastern equatorial Pacific SST cooling and southeasterly winds extend to the central-western equatorial Pacific. The intensification and westward extension of the Pacific cold tongue and associated climatological easterly winds are overestimated in the forecasts (Fig. 1e). Climatological SSTs in the forecasts show a cooling bias of up to -4.3°C around Sep⁰ in the first year of the March-initialized forecasts, together with an excessive westward shift of the Pacific warm pool edge (28°C) by up to 35° of longitude during Sep⁰–Mar⁺¹ of the March- and June-initialized forecasts relative to observations (Fig. 1e). The SST cooling and easterly wind biases over the western-central equatorial Pacific linger into early boreal spring in the forecasts, resulting in a 2-month delay of the warmest season in the equatorial eastern Pacific (February–April in observations and April–June in all forecast ensembles). These climatological SST and wind biases are partly inherent to the model’s climatology (Fig. 1f), since the deviations of forecasts from the CESM LE (Fig. 1g) are smaller than that from the observations (Fig. 1e). The biases in all three sets of forecasts are statistically significant at the 99.5% confidence level despite the different ensemble sizes, according to the bootstrap method across the ensemble members (Fig. S2).

The development of the SST cooling bias in the eastern equatorial Pacific is preceded by a negative (shallow) thermocline depth bias that propagates eastward from the western-central equatorial Pacific after the initialization (Fig. 1e). The initial “shallow” thermocline bias is rooted in errors in the initial condition data from FOSI, which simulates shallower climatological thermocline in the western-central Pacific compared to the ORAS4 (Fig. 2) and other six ocean reanalysis datasets (Fig. S1a). The equatorial Pacific SST simulated by the FOSI is overall warmer
than the observations, suggesting that the quick development of cold tongue SST bias in the March- and June-initialized forecasts is not an amplification of the initial SST errors. The causes of FOSI errors in simulating the observed thermocline depth in the equatorial Pacific require further examination but could be partly caused by the CESM1’s inherent bias in simulating the mean depth and slope of thermocline in the equatorial Pacific (Fig. 1f) and/or the errors in the prescribed surface winds in the equatorial and off-equatorial Pacific used to force the FOSI (not shown).

The cold tongue SST bias in year 0 is larger in the March- and June-initialized forecasts compared to the November-initialized forecasts (Fig. 1e). The dependence of cold tongue bias on the initialization month of the forecasts is associated with the strong seasonality of thermocline-SST feedback in the equatorial eastern Pacific (Fig. 3). In both observations and the CESM1 LE, the strength of the feedback, measured by the regression coefficient of SST on the thermocline depth in the eastern equatorial Pacific, is smallest during boreal spring (March–May; 0.02 °C m\(^{-1}\) in CESM1 LE) and largest in autumn (September–November 0.12 °C m\(^{-1}\) in CESM1 LE). The seasonality of the SST sensitivity to thermocline depth is associated with the seasonal change in climatological upwelling in the eastern equatorial Pacific (not shown). In the March- and June-initialized forecasts, the initial thermocline shoaling bias in the western-central equatorial Pacific propagates into the eastern basin during boreal summer-autumn, inducing large negative SST bias (Fig. 1e). In the November-initialized forecasts, by contrast, the thermocline shoaling bias arriving in boreal spring produces a smaller SST bias in the eastern equatorial Pacific. These results are consistent with a recent study by Siongco et al. (2020), which also shows that the cold tongue bias of CESM1 hindcasts develops faster during boreal summer-autumn than other seasons due to the strengthening of vertical temperature advection in the eastern equatorial Pacific.
The three sets of forecasts, especially those initialized in March and June, display a pronounced cooling bias in the first year (year 0) relative to subsequent years of the forecast (Fig. 1e). This strong initial cooling bias is associated with the imbalance of the initial ocean state and the model’s dynamics as shown in the difference between forecasts and CESM1 LE (Fig. 1g). The initial thermocline shoaling anomalies across the equatorial Pacific leads to an SST cooling response in year 0 of the forecasts relative to CESM1 LE. As result, the climatology of forecasts shows pronounced deviations from the model’s climatology in year 0 before stabilizing. The difference in climatological surface winds between forecasts and CESM LE is consistently very small throughout the two-year forecasts, suggesting that the initial model adjustment is mostly driven by oceanic processes.

b. Forecast errors of ENSO characteristics

Forecast errors of the temporal evolution and amplitude of ENSO events, and the dependency of error development on the initialization season, are assessed based on the predictions of Niño-3.4 SST index. Figures 4a,b show the timeseries of the Niño-3.4 index composited for all El Niño and La Niña events during 1954–2015 in observations and ensemble-mean forecasts initialized at different lead times. The average development of El Niño and La Niña (defined using a threshold of |Niño-3.4| > 0.5°C) in the summer of year 0 is not captured in the forecasts initialized in Nov⁻¹, which is probably related to the overall low predictability of the onset of ENSO events before boreal spring (e.g., Torrence and Webster 1998) rather than the model errors. One exception from this composite result is the onset of La Niña events following strong El Niño, which can be predicted in the forecasts initialized from the preceding El Niño peak (not shown). The onset is first seen in the Mar⁰-initialized forecasts, although the Niño-3.4 SST anomalies quickly dissipate
in the following autumn. The Jun$^0$-initialized forecasts show improvement, but the predicted peak of El Niño and La Niña is weaker and occurs too early compared to observations.

After the peak, all forecasts of El Niño show a warm SST bias in year +1 of about 0.6°C, while all forecasts of La Niña capture the returning of La Niña as seen in observations (Figs. 4a,b). To assess the forecast errors of ENSO event duration, we classify El Niño and La Niña events into those lasting one year (Figs. 4c,d) and two years (Figs. 4e,f). The forecasts initialized in and after Jun$^0$ successfully capture the termination of 1-yr El Niño and 1-yr La Niña events (defined using a threshold of $|\text{Niño-3.4}| < 0.5^\circ$C), but weak Niño-3.4 SST anomalies of the same sign persist too long into year +1 (Figs. 4c,d). On the other hand, the forecasts show good performance in predicting the continuation of 2-yr ENSO events. The re-intensification of 2-yr El Niño events is captured in the forecasts initialized in/after Nov$^0$, and the persistence of 2-yr La Niña events is predicted in all six sets of forecasts (Figs. 4e,f). The mechanisms driving the long-term predictability of 1-yr and 2-yr ENSO events have been explored by Wu et al. (2021a) and DiNezio et al. (2017b). Because 1-yr El Niño accounts for a large portion of El Niño events (10 out of 15) and 2-yr La Niña dominates the signal of all La Niña events (8 out of 12), the model shows better performance in predicting the overall duration of La Niña than El Niño events. In the remainder of this section, we show the results of our dynamical analysis for 1-yr El Niño events, and discuss the similarities and differences with the other types of ENSO events whose results are shown in the supplemental materials.

To illustrate the dynamical processes underpinning the forecast errors in Niño-3.4 SST index, the spatiotemporal evolutions of SST, surface winds, and thermocline depth anomalies in the forecasts of 1-yr El Niño are compared with observations (Fig. 5; see Fig. S3 for the results of other types of ENSO events and Fig. S4 for the statistical significance of the composites). In the
Mar⁰- and Jun⁰-initialized forecasts (Figs. 5b,c), the SST warming and surface westerly wind anomalies associated with 1-yr El Niño quickly extend toward the far western Pacific during boreal summer (Jun⁰–Aug⁰), when the mean-state cold tongue bias intensifies and alters the oceanic and atmospheric feedbacks in the western equatorial Pacific (Fig. 1e; Ham and Kug 2012; Bayr et al. 2018; Graham et al. 2017; Jiang et al. 2021). The difference between the forecasts and observations during Jun⁰–Apr⁰⁺₁ shows a zonal dipole pattern in the equatorial Pacific, indicating a westward shift of the SST and wind anomalies (Fig. 5e). The weakened amplitude of SST warming over the central-eastern equatorial Pacific (180°–80°W) in the Mar⁰ and Jun⁰-initialized forecasts is related to the weaker zonal dipole pattern of thermocline depth anomalies compared to the observations. The weakened thermocline tilt is linked not only to the simultaneous weakened westerly wind anomalies over the western-central Pacific (160°E–180°) but also to the errors in the initial thermocline depth anomalies. The initial positive thermocline depth anomalies in the central-eastern equatorial Pacific are smaller in the Mar⁰ and Jun⁰-initialized forecasts than observations, contributing to weaker El Niño amplitude. Besides the pattern and amplitude biases of 1-yr El Niño forecasts during Jun⁰–Apr⁰⁺₁, the biased SST warming and westerly wind anomalies over the western equatorial Pacific persist after the mature phase. In association with the lingering westerly anomalies and weaker peak amplitude of 1-yr El Niño, the thermocline shoaling over the western equatorial Pacific is too weak and does not propagate into the eastern equatorial Pacific after the mature phase in the Mar⁰ and Jun⁰-initialized forecasts, resulting in the failure of predicting the La Niña state in year +1.

In contrast, the biases of spatial pattern, peak amplitude, and phase transition of 1-yr El Niño events are reduced in the Nov⁰-initialized forecasts (Figs. 5d,e and S4), for which the mean-state cold tongue bias is weaker and the initialization is closer to the El Niño peak than the Mar⁰-
and Jun$^0$-initialized forecasts (Fig. 1e). In particular, the phase transition from El Niño to La Niña is captured in the Nov$^0$-initialized forecasts, although the amplitude of subsequent La Niña is underestimated presumably due to the smaller initial thermocline depth anomalies in FOSI. The forecasts of the other three types of ENSO events also show westward-shifted pattern bias and weaker amplitude of anomalies over the central-eastern Pacific, and these biases are stronger in the Mar$^0$- and Jun$^0$-initialized forecasts than the Nov$^0$-initialized forecasts (Fig. S3). Similar to 1-yr El Niño, the Mar$^0$ and Jun$^0$-initialized forecasts fail to capture the phase transition of 1-yr La Niña but show biased weak equatorial Pacific cooling throughout year +1. In contrast, the continuation of 2-yr El Niño is predicted by the Nov$^0$-initialized forecasts, and the continuation of 2-yr La Niña events is predicted in all three sets of forecasts, despite the pattern and amplitude biases in year +1.

To understand why the surface wind anomalies over the far western equatorial Pacific tend to linger into the second year and prolong equatorial Pacific warming in the 1-yr El Niño forecasts, we examine SST anomalies in the key ocean basins that could influence these wind anomalies. Figure 6 compares the longitude-time sections of SST and surface wind anomalies over the equatorial Pacific, the tropical Indian Ocean (10°S–0°), and the tropical North Atlantic (0°–20°N) composited for 1-yr El Niño events in the observations, forecasts initialized in Jun$^0$ and Nov$^0$, and CESM1 LE (see Fig. S5 for the results of other types of ENSO events and Fig. S6 for the statistical significance of the composites). In observations, the tropical Indian and Atlantic Oceans warm up during the mature phase (Dec$^0$–Feb$^{+1}$) of 1-yr El Niño due to tropical atmospheric adjustments to the equatorial Pacific SST warming (Fig. 6a; Xie and Carton 2004; Schott et al. 2009). The warming of tropical Indian and Atlantic Oceans during Dec$^0$–Feb$^{+1}$ is, however, weaker in the Jun$^0$-initialized forecasts than observations presumably due to the weaker El Niño warming over
the eastern equatorial Pacific (Fig. 6b,e; the Mar⁰ forecasts show similar results). The weaker warming of the Indian and Atlantic Oceans in the Jun⁰-initialized forecasts might be insufficient to reduce the interbasin SST gradient towards the Pacific and to weaken the surface westerly wind anomalies over the western Pacific. The influence of Indian and Atlantic Oceans on El Niño termination has been investigated in many previous studies (see Cai et al. 2019 for a review and references therein).

Furthermore, this negative feedback from the remote tropical oceans may also become less effective when the westerly wind anomalies associated with 1-yr El Niño are displaced westward and are too strong over the western Pacific in the Jun⁰-initialized forecasts, according to the mechanisms proposed for the persistent duration of La Niña events (Okumura and Deser 2010; Okumura et al. 2011). Supporting the two notions, the biases of predicted wind and SST in year +1 are reduced in the Nov⁰-initialized forecasts when the basin-wide Indian Ocean warming becomes more comparable to observations and the anomaly pattern bias of El Niño is reduced (Fig. 6c,f). We note that the westward shift of ENSO anomaly pattern is a bias inherent to the CESM1 (Fig. 6d). Nonetheless, this model reproduces the observed frequency of 1-yr El Niño events in the control simulation (Wu et al. 2019). In CESM1 LE, the delayed warming of the tropical Indian and Atlantic Oceans associated with 1-yr El Niño is much stronger than observations and may thus compensate for the errors in simulating the SST and wind anomaly patterns in the equatorial Pacific (Fig. 6d,g).

c. Forecast errors of ENSO teleconnections

The forecast errors of tropical SST and precipitation anomalies during the peak of 1-yr El Niño also affect the wintertime atmospheric teleconnections in the Northern Hemisphere. Figure 7 compares SST, surface wind, SLP, precipitation, and Z200 anomalies over the tropics and
Northern Hemisphere during Dec⁰–Feb⁺¹ compositing for the 1-yr El Niño events in observations and the Jun⁰ and Nov⁰-initialized forecasts (see Fig. S7 for the results of other types of ENSO events and Fig. S8 for the statistical significance of the composites). In association with the tropical Pacific SST warming, SLP decreases over the North Pacific, deepening the Aleutian Low in both observations and forecasts (Figs. a-c, left). In the Jun⁰-initialized forecasts, the center of the circulation anomalies exhibits a westward shift of about 10° of longitude and a more zonally elongated pattern compared to observations (Figs. 7a,b, left).

The surface atmospheric biases are mirrored in the upper tropospheric circulation anomalies (Figs. 7a,b, right). In observations, the increased precipitation and resultant diabatic heating induce a pair of anticyclones at 200 hPa straddling the central equatorial Pacific and a Rossby wave train propagating into the Northern Hemisphere. In the Jun⁰-initialized forecasts, the response of extratropical Z200 anomalies is shifted westward and is more zonally elongated compared to observations. The pattern difference between Jun⁰-initialized forecasts and observations displays a slightly tilted zonal dipole pattern in both SLP and Z200 over the mid-latitude North Pacific (Fig. 7d). This Z200 anomaly dipole is part of a Rossby wave train originating in southeastern China, which is forced by the biased positive precipitation anomalies over the far western tropical Pacific. In contrast, the biases of extratropical atmospheric circulations are much reduced in the Nov⁰-initialized forecasts, consistent with the weaker biases in the pattern and amplitude of tropical SST and precipitation anomalies (Fig. 7e). This result is consistent with the findings by Bayr et al. (2019b), who show that those CMIP5 models with a stronger cold tongue bias tend to simulate a westward shift of atmospheric convection in the tropical Pacific and associated atmospheric teleconnections to the North Pacific during ENSO events.
Besides the biases of atmospheric responses over the extratropical North Pacific, we also observe that the anticyclonic SLP anomalies over the western North Pacific (WNP) are much weaker in the Jun^0^-initialized forecasts than observations. In observations, anticyclonic SLP anomalies develop over the western North Pacific mainly as a Rossby wave response to the SST cooling and decreased precipitation anomalies over the far western Pacific during El Niño (Fig. 7a; e.g., Wang et al. 2000). The precipitation over the western Pacific is not suppressed only by local SST cooling but also by the Indian Ocean warming (e.g., Watanabe and Jin 2002). The Jun^0^-initialized forecasts predict positive SST and precipitation anomalies over the far western equatorial Pacific as well as weaker Indian Ocean warming, resulting in a very weak response of SLP anomalies over the WNP (Fig. 7b). In observations, the surface southerly wind anomalies on the western flank of the anticyclonic SLP anomalies increase the precipitation over East Asia by bringing high moisture content from tropical oceans (Fig. 7a, e.g., Wu et al. 2003). Instead, the Jun^0^-initialized forecasts fail to capture these southerly wind anomalies and associated increased precipitation anomalies over East Asia (Fig. 7b,d). The predictions of atmospheric circulation and rainfall anomalies over East Asia are both improved in the Nov^0^-initialized forecasts when the ENSO-related tropical SST and precipitation anomalies are more realistically reproduced (Fig. 7c,e).

These forecast errors of atmospheric teleconnections may, in turn, affect surface wind variability over the western equatorial Pacific and the evolution of ENSO states in the second year. In the difference map between the Jun^0^-forecasts and observations, the biased SST warming and the southwesterly wind anomalies over the subtropical Pacific resembles the North Pacific meridional mode (NPMM) and potentially contributes to the persistence of equatorial Pacific warming in year +1 (Fig. 7d; e.g., Anderson 2003; Vimont et al. 2003; Chang et al. 2007;
Alexander et al. 2010; Fang and Yu 2020; Kim and Yu 2021). The weak development of the WNP anomalous anticyclone may also contribute to the ENSO forecast errors in year +1 because the easterly wind anomalies on the southern flank of this anticyclone were suggested as an important factor for El Niño termination (Weisberg and Wang 1997).

Excessively weak tropical interbasin SST adjustments, along with biases in the pattern and amplitude of Northern Hemisphere teleconnections, are found for the forecasts of the other three types of ENSO events (Figs. S5 and S7), although the results for 1-yr La Niña and 2-yr El Niño in observations are largely insignificant at the 95% confidence level due to the small number of events in the composites (Figs. S6 and S8). In contrast to degrading the forecasts of 1-yr ENSO events in year +1, the teleconnection errors over remote oceans during the mature phase of ENSO favor the persistence of 2-yr ENSO events. For example, in the Jun⁰ forecasts of 2-yr La Niña (Figs. S5 and S7), the too weak interbasin adjustment, biased negative NPMM, and weak WNP cyclonic SLP circulation may promote the persistence of easterly wind anomalies and SST cooling over the western equatorial Pacific in the second year.

4. Summary and discussions

This study investigates the climatological bias in the equatorial Pacific and associated ENSO forecast errors in three sets of two-year CESM1 retrospective forecasts initialized in March, June, and November during 1954–2015. The forecasts show a strong cooling bias of the climatological SSTs in the eastern Pacific cold tongue, along with a westward shift of the edge of the Pacific warm pool and excessive climatological easterly winds over the western-central Pacific. This cold tongue bias is especially large in the first year of the forecasts before the model stabilizes at its own climatology at longer lead times. This initial large forecast bias is caused by the imbalance between the observed (FOSI) initial states and the model’s own dynamics. The model
is initialized with a shallower climatological thermocline depth in the western-central equatorial Pacific than its own climatology, which induces subsequent SST cooling in the first year. The magnitude of the cold tongue bias also depends on the initialization timing due to the seasonality of SST-thermocline feedback in the equatorial Pacific. The cold tongue bias is larger in the March- and June- initialized forecasts than in the November-initialized forecasts, because the thermocline depth errors in the initial condition data propagate into the eastern equatorial Pacific during summer-autumn when the SST-thermocline feedback strengthens and can more effectively induce the SST cooling bias.

Related to the climatological biases of the cold tongue, larger ENSO forecast errors are identified in the March- and June-initialized forecasts than in the November-initialized forecasts. In the March- and June-initialized forecasts, SST and wind anomalies associated with ENSO migrate rapidly toward the far western equatorial Pacific in the summer season when the strong climatological cold tongue bias could overamplify zonal advective feedback in the western equatorial Pacific (Graham et al. 2017; Jiang et al. 2021). The pattern bias results in weakened SST anomalies over the central-eastern Pacific and hence weaker ENSO amplitude (as captured by the Niño-3.4 SST index). The underestimation of ENSO amplitude in the forecasts is also associated with weaker initial thermocline anomalies simulated by the FOSI compared to the ocean reanalysis data. The pattern and amplitude biases of ENSO further affect the forecasts of the processes that are important to the termination of ENSO events. In particular, too weak El Niño in the forecasts results in weak negative thermocline feedback in the equatorial Pacific and weak tropical Indian and Atlantic SST warming around the mature phase; the weaker SSTs in turn are insufficient to cause a substantial reduction in the strength of the westerly wind anomalies over the western equatorial Pacific. Due to the pattern bias, the wind anomalies associated with ENSO
are displaced too far to the west and therefore may make the wind variability over the western
equatorial Pacific less susceptible to the negative feedback induced by the tropical Indian and
Atlantic Oceans. These biased oceanic and atmospheric processes may work together to cause the
forecast errors of ENSO phase transition.

The ENSO forecast errors further influence the predictions of atmospheric teleconnections
and rainfall in the Northern Hemisphere during the mature phase of ENSO. Related to the
westward displacement of ENSO SST and precipitation anomalies over the tropical Pacific, the
tropospheric and surface atmospheric circulation anomalies over the North Pacific show a
westward shift and become more zonally elongated in forecasts compared to observations.
Additionally, the pattern bias of ENSO anomalies leads to SLP errors over the western Pacific and
to precipitation errors over East Asia. These extratropical atmospheric errors during the mature
phase of ENSO events may, in turn, affect the predictions of ENSO duration in the following year
via the NPMM and WNP anomalous anticyclones/cyclones. In particular, these errors tend to
prolong the duration of ENSO, degrading the predictions of ENSO termination but contributing to
the high forecast skill of ENSO continuation for the wrong reason. Further studies are needed to
examine the forecast errors of ENSO teleconnections in other seasons and the interactions between
the tropical and extratropical processes.

Our study shows that errors in climatologies and ENSO characteristics in full-field
initialized forecasts can be even larger than those in uninitialized simulations of the same model.
The initialization imbalance between the ‘observed’ (FOSI) oceanic initial conditions and model
dynamics is amplified by coupled feedbacks in the equatorial Pacific, resulting in a stronger cold
tongue bias in the forecasts than that inherent to the model. The traditional method of removing
the drifting climatology of forecasts cannot eliminate the influence of biased ocean-atmosphere
feedbacks on ENSO variability. As a result, these initialized forecasts show ENSO errors that are absent from the uninitialized simulations. For example, the CESM1 forecasts tend to overestimate the duration of ENSO events, while the uninitialized simulations of the same model show realistic proportions of ENSO events lasting one and two years (Wu et al. 2019). It will be interesting to compare ENSO forecast skill in CESM1 with that in its successor, CESM2, which shows a reduced cold tongue bias (Capotondi et al. 2020). More generally, the relationship between mean state biases and ENSO biases warrants further examination using multi-model ensemble forecasts.

Our results reveal opportunities to improve the forecast skill of ENSO events by improving the model’s mean state and the accuracy of initial condition data. Despite upgraded physics and spatial resolution across the generations of CMIP models, systematic model errors in tropical Pacific climatology and ENSO variability persist (see a recent review by Guilyardi et al. 2020). The method of ‘flux adjustment’ was adopted as an intermediate step to improve the simulation of the tropical Pacific mean state and its seasonal variation by prescribing a seasonally varying nudging climatology (e.g., Ray et al. 2018). Another method termed ‘anomaly initialization’ has been used to reduce the model drift and initialization shock, by which forecasts are initialized from the observed anomalies added to the model’s climatology. However, it remains unclear and model-dependent which initialization method can produce more skillful ENSO and seasonal climate predictions (Magnusson et al. 2013; Smith et al. 2013; Hu et al. 2020). The ocean initial condition data used for our forecasts show errors in reproducing both the climatology and variability of thermocline depth in the equatorial Pacific. These initial conditions were obtained by forcing the ocean model of CESM1 with atmospheric forcing from the reanalysis data, which are hence subject to the model errors and uncertainties in the prescribed wind forcing. Integrated efforts aimed at
improving model simulations, observing systems and data assimilation methods are needed to make progress on prediction skill of ENSO and its worldwide climate impacts.

Acknowledgments. The observational datasets used in this study are obtained online: HadISST from the Met Office Hadley Center (http://www.metoffice.gov.uk/hadobs/), ECMWF Ocean Reanalysis System 4 from ftp://ftp-icdc.cen.uni-hamburg.de/EASYInit/ORA-S4, and NCEP-NCAR reanalysis from https://psl.noaa.gov/data/gridded/data.ncep.reanalysis.html. The CESM1 Large Ensemble simulations used in this study are available from https://www.cesm.ucar.edu/projects/community-projects/LENS/. This study is supported by the National Oceanic and Atmospheric Administration (NOAA) Climate Program Office Modeling, Analysis, Predictions, and Projections Program (NA17OAR4310149 and NA17OAR4310145) and the National Science Foundation (NSF) Physical Oceanography Program (OCE-1756883). X. Wu is supported by an Advanced Study Program postdoctoral fellowship from then National Center for Atmospheric Research (NCAR). The CESM project is supported primarily by the NSF. This material is based upon work supported by the NCAR, which is a major facility sponsored by the NSF under Cooperative Agreement No. 1852977.
References


Figure 1. Longitude-time sections of climatological monthly SST (°C; shading), thermocline depth (contours at intervals of 30 m), and surface wind (m s⁻¹; vectors) in the equatorial Pacific (3°S–3°N) for the ensemble forecasts initialized in (a) November, (b) March, and (c) June.
superscripts on the months along the y-axis indicate the first year when all forecasts overlap), (d)

observations (HadISST, ORAS4, and NCEP-NCAR) and (f) CESM1 LE during 1958–2015. Note

that the climatologies of the forecasts are a function of lead time, whereas those of the observations

and CESM1 LE are simply repeated for clarity. The deviations of the forecast climatologies from

(e) observations and (g) CESM1 LE for SST (°C; shading), thermocline depth (contours at

intervals of 10 m; zero contours thickened and negative contours dotted), and surface wind (m s⁻¹;

vectors). The thermocline depth is smoothed with a 9-point running-mean filter in the longitudinal
direction. The statistical significance of Fig. 1e is shown in Fig. S2.
Figure 2. (a) Longitude-time sections of climatological SST (°C; shading) and thermocline depth (contours at intervals of 30 m) in the equatorial Pacific (3°S–3°N) during 1958–2015 for the FOSI. (b) The difference of climatological SST (°C; shading) and thermocline depth (contours at intervals of 10 m; zero contours thickened and negative contours dotted) between the FOSI and observations (HadISST and ORAS4). The thermocline depth is smoothed with a 9-point running-mean filter in the longitudinal direction.
Figure 3. Scatterplots of thermocline depth (m) vs SST (°C) in the eastern equatorial Pacific (3°S–3°N; 150°–80°W) averaged over March–May (orange), June–August (blue), September–November (purple), and December–February (green) in the (a) observations (HadISST and ORAS4) and (b) CESM1 LE during 1958–2015. The numbers at the bottom right of each panel indicate the regression coefficients of SST on thermocline depth (°C per m) in each season.
Figure 4. Time series of the Niño-3.4 index (°C) composited for (a) all El Niño, (b) all La Niña, (c) 1-yr El Niño, (d) 1-yr La Niña, (e) 2-yr El Niño, and (f) 2-yr La Niña events in the observations (HadISST; black curves) and ensemble-mean forecasts (colored curves) during 1954-2015. The forecasts are initialized in Nov^{-1} (pink), Mar^{0} (red), Jun^{0} (yellow), Nov^{0} (green), Mar^{+1} (light blue), and Jun^{+1} (dark blue). Dec^{0} denotes the first peak of El Niño/La Niña events. The dashed horizontal lines denote ± 0.5°C threshold. The solid portions of the colored curves indicate that the composite forecasts are significantly different from zero at the 95% confidence level.
Figure 5. Longitude-time sections of SST (°C; shading), thermocline depth (contours at intervals of 10 m; zero contours thickened and negative contours dashed), and surface wind (m s\(^{-1}\); vectors) anomalies in the equatorial Pacific (3°S–3°N) composited for 1-yr El Niño in the (a) observations (HadISST, ORAS4, and NCEP-NCAR), ensemble-mean forecasts initialized in (b) Mar\(^0\), (c) Jun\(^0\), (d) Nov\(^0\), and (e) the difference between the forecasts and observations. The thermocline depth anomalies are smoothed with a 1-2-1 filter in the time direction and a 9-point running-mean filter in the longitudinal direction in both observations and forecasts. The composites for other types of ENSO events are shown in Fig. S3, and the statistical significance of these anomalies is shown in Fig. S4.
Figure 6. Longitude-time sections of SST (°C; shading) and surface wind (m s\(^{-1}\); vectors) anomalies in the equatorial Pacific (3°S–3°N), the Indian Ocean (10°S–0°), and the Atlantic (0°–20°N) composited for 1-yr El Niño in the (a) observations (HadISST and NCEP-NCAR), (b) Jun\(^0\) forecasts, (c) Nov\(^0\) forecasts, (d) CESM1 LE, and the deviations of (e) Jun\(^0\) forecasts, (f) Nov\(^0\) forecasts and (g) CESM1 LE relative to the observations. The composites for other types of ENSO events are shown in Fig. S5, and the statistical significance of these anomalies is shown in Fig. S6.
Figure 7. Maps of (left) SST (°C; shading), surface wind (m s$^{-1}$, vectors), SLP (contours at intervals of 0.6 hPa; zero contours thickened and negative contours dashed in gray), (right) precipitation (mm day$^{-1}$; shading), and Z200 (contours at intervals of 15 m; zero contours thickened and negative contours dashed) anomalies composited for 1-yr El Niño in the (a) observations (HadISST, ORAS4, and NCEP-NCAR), (b) Jun$^0$ forecasts, (c) Nov$^0$ forecasts, and deviations of the (d) Jun$^0$ forecasts and (e) Nov$^0$ forecasts from the observations. The composites for other types of ENSO events are shown in Fig. S7, and the statistical significance of these anomalies is shown in Fig. S8.