Understanding Diverse Model Projections of Future Extreme El Niño

SAMANTHA STEVENSON,[a] ANDREW T. WITTENBERG,[b] JOHN FASULLO,[c] SLOAN COATS,[d] AND BETTE OTTO-BLIESNER[c]

[a] *Bren School of Environmental Sciences and Management, University of California at Santa Barbara, Santa Barbara, California*
[b] *NOAA Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey*
[c] *Climate and Global Dynamics Division, National Center for Atmospheric Research, Boulder, Colorado*
[d] *Department of Earth Sciences, University of Hawai'i at Mānoa, Honolulu, Hawaii*

(Manuscript received 27 December 2019, in final form 8 September 2020)

ABSTRACT: The majority of future projections in the Coupled Model Intercomparison Project (CMIP5) show more frequent exceedances of the 5 mm day^{-1} rainfall threshold in the eastern equatorial Pacific rainfall during El Niño, previously described in the literature as an increase in "extreme El Niño events"; however, these exceedance frequencies vary widely across models, and in some projections actually decrease. Here we combine single-model large ensemble simulations with phase 5 of the Coupled Model Intercomparison Project (CMIP5) to diagnose the mechanisms for these differences. The sensitivity of precipitation to local SST anomalies increases consistently across CMIP-class models, tending to amplify extreme El Niño occurrence; however, changes to the magnitude of ENSO-related SST variability can drastically influence the results, indicating that understanding changes to SST variability remains imperative. Future El Niño rainfall intensifies most in models with 1) *larger* historical cold SST biases in the central equatorial Pacific, which inhibit future increases in local convective cloud shading, enabling more local warming; and 2) *smaller* historical warm SST biases in the far eastern equatorial Pacific, which enhance future reductions in stratus cloud, enabling more local warming. These competing mechanisms complicate efforts to determine whether CMIP5 models under- or overestimate the future impacts of climate change on El Niño rainfall and its global impacts. However, the relation between future projections and historical biases suggests the possibility of using observable metrics as "emergent constraints" on future extreme El Niño, and a proof of concept using SSTA variance, precipitation sensitivity to SST, and regional SST trends is presented.

KEYWORDS: ENSO; El Nino; Extreme events; Climate change; Climate variability

1. Introduction

El Niño–Southern Oscillation (ENSO) dominates interannual climate variability, with severe impacts on many socioeconomic sectors, as well as on marine and terrestrial ecosystems (Hoegh-Guldberg 1999; Ainsworth et al. 2016; Di Lorenzo and Mantua 2016). Some of the largest effects are felt during strong El Niño events, which are characterized by warming in the eastern equatorial Pacific, reductions in equatorial upwelling, equatorward migration of the intertropical convergence zone (ITCZ; Ropelewski and Halpert 1986), and increased convective precipitation throughout the central and eastern equatorial Pacific. However, the properties of individual El Niño events are a function of multiple interacting atmosphere–ocean feedback processes (Collins et al. 2010; Capotondi et al. 2015; Timmermann et al. 2018), many of which will be influenced by future anthropogenically driven changes in the tropical Pacific (Collins et al. 2010; DiNezio et al. 2012). As such, constraining future changes in ENSO remains an outstanding research challenge.

Climate model projections do not presently agree on either the magnitude or sign of twenty-first-century changes in the amplitude of ENSO-driven SST anomalies (Collins et al. 2010; Vecchi and Wittenberg 2010; Stevenson et al. 2012; Stevenson 2012; Bellenger et al. 2014; Chen et al. 2017). Determining which projections are most likely to be realistic is difficult, in part because of systematic biases in the simulation of tropical Pacific climate. For example, most climate models overestimate the strength of the equatorial trade winds, leading to overly cold temperatures in the central/eastern Pacific (Guilyardi et al. 2009b; Bellenger et al. 2014). This leads to a westward shift in the ENSO "center of action" for atmospheric deep convection, rainfall, and winds, which has been hypothesized to affect the properties of El Niño in coupled models (Capotondi et al. 2015; Choi et al. 2015; Graham et al. 2017). Models also struggle to capture the magnitude of relevant feedbacks, including those relating sea surface temperature (SST) to changes in evaporative cooling and shortwave heating (Bellenger et al. 2014; Guilyardi et al. 2009a,b), which mediate future changes in ENSO (Guilyardi et al. 2009a) and the background climate (Xie et al. 2010; Huang et al. 2015; Ying and Huang 2016; Chung et al. 2019; Seager et al. 2019).

Since there is little agreement on future changes in ENSO-driven SST variability, recent work has instead focused on changes in atmospheric measures of El Niño impacts (Power et al. 2013; Cai et al. 2014, 2015), which are thought to be more robust due to the consistency of projected increases in global-mean temperature and saturation specific humidity (Held and Soden 2006; Vecchi and Soden 2007). "Extreme El Niño" events, when defined according to precipitation amounts exceeding certain thresholds in the eastern equatorial Pacific, become more frequent in coupled climate model projections of the twenty-first century (Cai et al. 2014, 2017). Several mechanisms have been proposed for the projected amplification of

Corresponding author: Samantha Stevenson, stevenson@bren.ucsb.edu

DOI: 10.1175/JCLI-D-19-0969.1

TABLE 1. Models and ensemble sizes from the CMIP5 experiments used in the present study. A plus sign (+) indicates that a model is a member of the INCR subsample; a minus sign (−) indicates that a model is a member of DECR, for a given extreme El Niño identification method. AP = anomaly percentile method; Pmn = absolute precipitation method; Panom = precipitation anomaly method.

Model	No. of members	INCR/DECR		
		Pmn	Panom	AP
ACCESS1.0	1	−		
ACCESS1.3	1			
BCC-CSM1.1	1			
BCC-CSM1.1-m	1	−		−
BNU-ESM	1			
CanESM2	5		−	−
CCSM4	6			−
CESM1-BGC	1			
CESM1-CAM5	3	+	+	+
CESM1-CAM5.1-FV2	1	+	+	+
CESM1-WACCM	3	−	−	−
CMCC-CESM	1	+	+	+
CMCC-CM	1			
CMCC-CMS	1	+	+	+
CNRM-CM5	5			
CSIRO-Mk3.6.0	10			
EC-EARTH	10			+
FGOALS-g2	1		−	−
FIO-ESM	3		−	−
GFDL CM3	1	+		
GFDL-ESM2G	1			
GFDL-ESM2M	1	−	−	−
GISS-E2-H	5	+		
GISS-E2-H-CC	1		+	+
GISS-E2-R	5	−		
GISS-E2-R-CC	1			
HadGEM2-AO	1	+		
HadGEM2-CC	3			
HadGEM2-ES	4			
INM-CM4	1			−
IPSL-CM5A-LR	4			
IPSL-CM5A-MR	1			
IPSL-CM5B-LR	1	+		
MIROC5	5	+	+	+
MIROC-ESM	1			−
MIROC-ESM-CHEM	1			−
MPI-ESM-LR	3			+
MPI-ESM-MR	1			+
MRI-CGCM3	1			+
MRI-ESM1	1			
NorESM1-M	1	−	−	−
NorESM1-ME	1			

ENSO-related precipitation anomalies, including enhanced rainfall due to Clausius–Clapeyron effects and an eastward shift of precipitation anomalies during El Niño (Bonfils et al. 2015; Huang 2016, 2017), nonlinear effects on precipitation anomalies due to changes in convective initiation (Huang and Xie 2015; Johnson and Xie 2010; Power et al. 2013), and changes in the meridional gradient of eastern Pacific SST near the equator (Cai et al. 2014). However, the relative importance of these processes varies widely across models (Huang 2017). Likewise, little has been done to address the question of how either anthropogenically driven or internal (unforced natural) variability in ENSO-induced SST patterns (Wittenberg 2009; Stevenson et al. 2010, 2012; Wittenberg et al. 2014) may affect near-term projections of precipitation extremes.

Previous studies have generally focused on the Coupled Model Intercomparison Project (CMIP) simulations, which provide consistent external forcing across multiple models to quantify model structural uncertainty. Since the completion of phase 5 of CMIP (CMIP5; Taylor et al. 2009), however, there has been increased recognition of the importance of internal variability in affecting intermodel comparisons, as well as model comparisons with observations (Deser et al. 2016; Stevenson 2012). This has motivated the development of "large ensembles," sets of many (often 20 or more) simulations with the same model differing only in their initial conditions (Deser et al. 2020).

To fully evaluate the roles of both model structural differences and contributions from internal variability, the ideal approach would be to examine large ensembles run with each of the individual CMIP5 models. However, relatively few modeling centers have generated such ensembles to date, although the number is growing (Deser et al. 2020). Here we use two representative large ensembles: the CESM1 Large Ensemble (hereafter CESM LENS; Kay et al. 2015; 33 members) and an analogous large ensemble run with the GFDL ESM2M (hereafter ESM2M LENS; Dunne et al. 2012; Rodgers et al. 2015; 30 members). Both cover the twentieth century and the 2006–2100 period under the RCP8.5 climate change scenario, and faithfully reproduce many aspects of twentieth-century climate (Knutti et al. 2013; Rodgers et al. 2015). Additionally, the two large ensembles exhibit diametrically opposing twenty-first-century ENSO projections, making them useful "end members" for comparison with CMIP5. The combination of these LENS suites with the CMIP5 (Table 1) then allows us to characterize for the first time the role of model biases in generating intermodel differences in extreme El Niño events.

2. Extreme El Niño definitions

The performance of CMIP5 models at simulating ENSO and its teleconnections has been discussed extensively elsewhere (Guilyardi et al. 2012; Bellenger et al. 2014; Chen et al. 2017). Likewise, several studies have focused on the definition of an extreme El Niño event, particularly with respect to rainfall (Cai et al. 2014, 2017). Here we present multiple identification methods for extreme El Niño, and examine strategies for isolating these extreme events across models that differ in their simulation of mean climate and ENSO behavior.

The choice of a 5 mm day^{-1} threshold in absolute precipitation along the equator at the peak of SSTA during El Niño (hereafter, the "absolute precipitation method") has been previously used to define whether an El Niño is "extreme" (Cai et al. 2014). The justification for this method has been that 5 mm day^{-1} represents the latent heat release necessary for

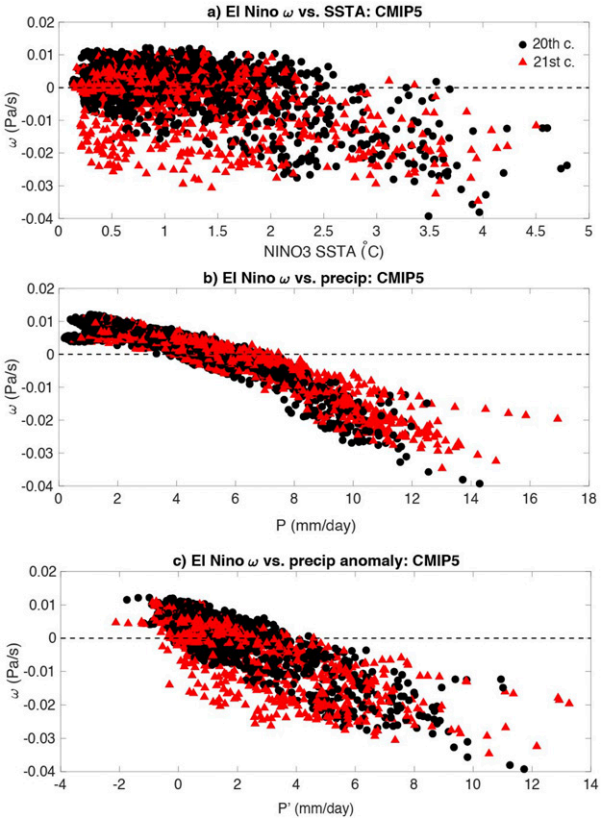

FIG. 1. Relationship of vertical pressure velocity ω with (a) SSTA, (b) precipitation, and (c) precipitation anomaly during DJF in which El Niño events peak in the CMIP5 simulations. Here values are calculated over the Niño-3 region, and for ω all pressure levels above 500 hPa are considered. Negative ω values indicate ascent. The twentieth century is defined as 1950–2005, and the twenty-first century as 2006–2100, for all models. El Niño events are defined as DJF periods where Niño-3 SSTA, deseasonalized with a 30-yr centered moving window, exceeds 0.5σ.

triggering the large-scale atmospheric reorganization needed to generate strong teleconnections to remote locations. We have evaluated the robustness of this assumption by diagnosing the upper-level (above 500 hPa) atmospheric vertical motion in the Niño-3 region over the historical period (1950–2005; Fig. 1a). Indeed, vertical motion does seem to occur preferentially in CMIP5 models when equatorial precipitation exceeds roughly 5 mm day^{-1} (Fig. 1b), with lower scatter than exists in the ω–SSTA relationship (Fig. 1a). However, changes in this relationship under global warming appear highly sensitive to the SSTA magnitude; from Fig. 1a it is apparent that the increase in ascent (decrease in ω) at a given SSTA is most significant for weak El Niño events, while the response to strong SSTA hardly changes (i.e., it saturates). Additionally, the ability of a given magnitude of vertical ascent to generate extreme precipitation increases in the future (see P values associated with particular ω in Fig. 1b).

There are further limitations to the absolute 5 mm day^{-1} definition of extreme El Niño when applied in a multimodel twenty-first-century context. Most notably, the implicit inclusion of mean-state and seasonal-cycle changes in precipitation in the response to anthropogenic forcing contributes strongly to changes in the frequency of absolute threshold crossings; mean precipitation shifts were shown to account for nearly 50% of the total precipitation change during extreme El Niño by Cai et al. (2017). This suggests that to explicitly quantify ENSO-related changes in interannual precipitation variability, a different definition is required. We therefore examine an additional extreme El Niño definition based on threshold crossing of DJF-averaged precipitation anomalies (hereafter, the "precipitation anomaly" method). The threshold for extreme El Niño is set at a 2 mm day^{-1} precipitation anomaly, which corresponds to the value, on average, above which CMIP5-simulated ω becomes negative (i.e., upward; Fig. 1c).

We next consider one final definition of extreme El Niño, which accounts for the substantial differences between the distributions of precipitation in various regions and models (hereafter, the "anomaly percentile method"). This considers precipitation anomalies during DJF larger than the 90th percentile of the twentieth-century distribution at each spatial grid point as extreme. Although this definition has less bearing on large-scale tropical atmospheric reorganization, it is more relevant to the characterization of local rainfall impacts of El Niño. In other words, by examining the extremes of the local precipitation distribution, we can compare events most likely to be considered "unusual" across locations (and models) with very different climatological mean states.

The identification of extreme El Niño events in observations is shown in Fig. 2 for all three definitions. Here, the CPC Merged Analysis of Precipitation (CMAP/A, which incorporates both rain gauge and satellite data) and NOAA's Extended Reconstructed SST (ERSSTv5) datasets are used. The mean precipitation and precipitation anomaly thresholds capture events in which there is a significant migration of the ITCZ onto the equator. Both choices of threshold result in the 1982/83 and 1997/98 El Niño events being selected; the 2015/16 El Niño is extreme according to the anomaly definition, but not the mean definition. In the anomaly percentile definition, the list of events is similar, but also includes the 1991/92 El Niño. We note that this analysis only includes events since 1979, as precipitation datasets over the ocean prior to the start of the satellite era are known to have large uncertainties.

3. Twenty-first-century changes to extreme El Niño

The next task is to assess how the occurrence frequency of extreme El Niño changes in the twenty-first century relative to the twentieth, across multiple event definitions and model projections. Here we define the 1950–2005 period as the twentieth century; this is chosen to ensure full temporal coverage across all CMIP5 models and large ensembles. The 2006–2100 period is then considered as the twenty-first century. All event frequencies have been normalized via dividing the total number of events by the appropriate epoch length. The results are shown in Fig. 3, from which it is clear that there is a wide range of projected changes simulated in CMIP5, in all three choices of event definition. The majority of models do indeed project an increase in extreme El Niño frequency; however, a

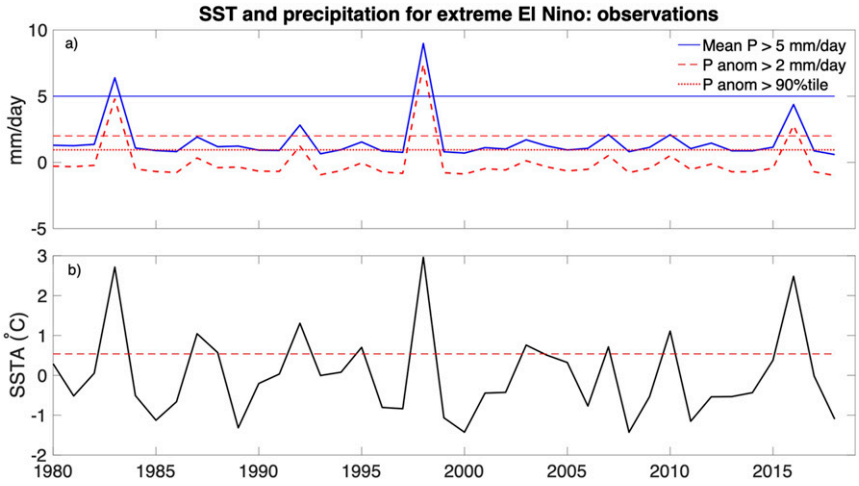

FIG. 2. Extreme El Niño event definitions in observations (CMAP and ERSSTv5). (a) Precipitation and precipitation anomaly (mm day^{-1}). (b) SSTA (°C). Anomalies are computed relative to the 1979–2016 climatology. Horizontal lines in (a) indicate the thresholds for the mean precipitation (5 mm day^{-1}), precipitation anomaly (2 mm day^{-1}), and anomaly percentile (90th percentile) extreme El Niño definitions. Horizontal line in (b) indicates the SSTA threshold used to define an El Niño event (0.5σ). All metrics are DJF averages, computed over the Niño-3 region.

substantial minority of models project a decrease, and the magnitude of the increase is quite small in others.

Event frequency changes do depend on the event definition. In some models, the mean Niño-3 precipitation (precipitation anomaly) during El Niño is less than 5 mm day^{-1} (2 mm day^{-1}) in either the twentieth or twenty-first centuries, and the change in event frequencies can therefore not be computed (missing data in Figs. 3a,b). This is consistent with previous CMIP5 evaluations (Cai et al. 2014), although our calculations use a slightly different time period for the twentieth and twenty-first centuries, and therefore do not correspond exactly. The event frequency changes are generally larger using the absolute precipitation method relative to the precipitation anomaly method, consistent with previous assessments of mean-state contributions to extreme El Niño frequency changes (Cai et al. 2017). In some models, the removal of mean precipitation can even flip the sign of the frequency change (Fig. 3a vs Fig. 3b; e.g., CanESM2, BCC-CSM1.1-m, FGOALS-g2, and FIO-ESM). The anomaly percentile method includes a larger number of models in both the increasing and decreasing categories, since all models can be assessed using this metric. The degree of relative spread between models projecting increases and decreases in event frequency, however, is comparable to the other methods. Interestingly, some models project very similar results regardless of the definition chosen. For instance, GFDL-ESM2M and NorESM1-M project decreases in all three definitions, while CESM1-CAM5, CESM1-CAM5.1-FV2, CMCC-CMS, and MIROC5 always project increases (Table 1). This suggests that the physical drivers of extreme El Niño changes in these models are independent of the details of event definition. The consistency of the CESM1 and ESM2M projections across extreme El Niño definitions, and their strong disagreement overall with one another, makes them ideal end members for further study, and we have thus included the CESM and ESM2M large ensembles (Kay et al. 2015; Rodgers et al. 2015) in our analysis as well.

Previous studies of extreme El Niño in CMIP models have focused on a subsample selected based on their ability to correctly simulate twentieth-century mean climate. For instance, precipitation skewness greater than 1 and Niño-3 rainfall greater than 5 mm day^{-1} during at least one boreal winter over the 1891–2090 period were the criteria imposed by Cai et al. (2014). We have deliberately not imposed such a subselection here, to enable a complete assessment of the impact of model bias on changes in extreme El Niño frequency. However, for comparison we have flagged those models selected by Cai et al. (2014) in Fig. 3 (red outlined bars). Notably, in all three event classifications, models selected by Cai et al. (2014) do not appear substantially more consistent with one another than the set of CMIP5 models as a whole. This suggests that even among models that do capture important aspects of the precipitation climatology correctly, there is still room to understand how physical changes across models affect the resulting twenty-first-century projections.

The anomaly percentile method allows changes in locally extreme events to be computed across the entire Pacific basin (Fig. 4). Additionally, both wet and dry extremes may be considered; here an extreme dry El Niño is one with DJF precipitation below the local 10th percentile. The multimodel mean structure of extreme El Niño using the anomaly percentile method is spatially El Niño–like (Figs. 4a–c), as is the pattern of twenty-first-century changes in event frequency (Fig. 4d). Extreme dry events are most common over Southeast Asia and the Maritime Continent in the twentieth century, as well as being fairly frequent on the edges of the convergence zones. There seems to be little agreement on changes in extreme dry El Niño in the twenty-first

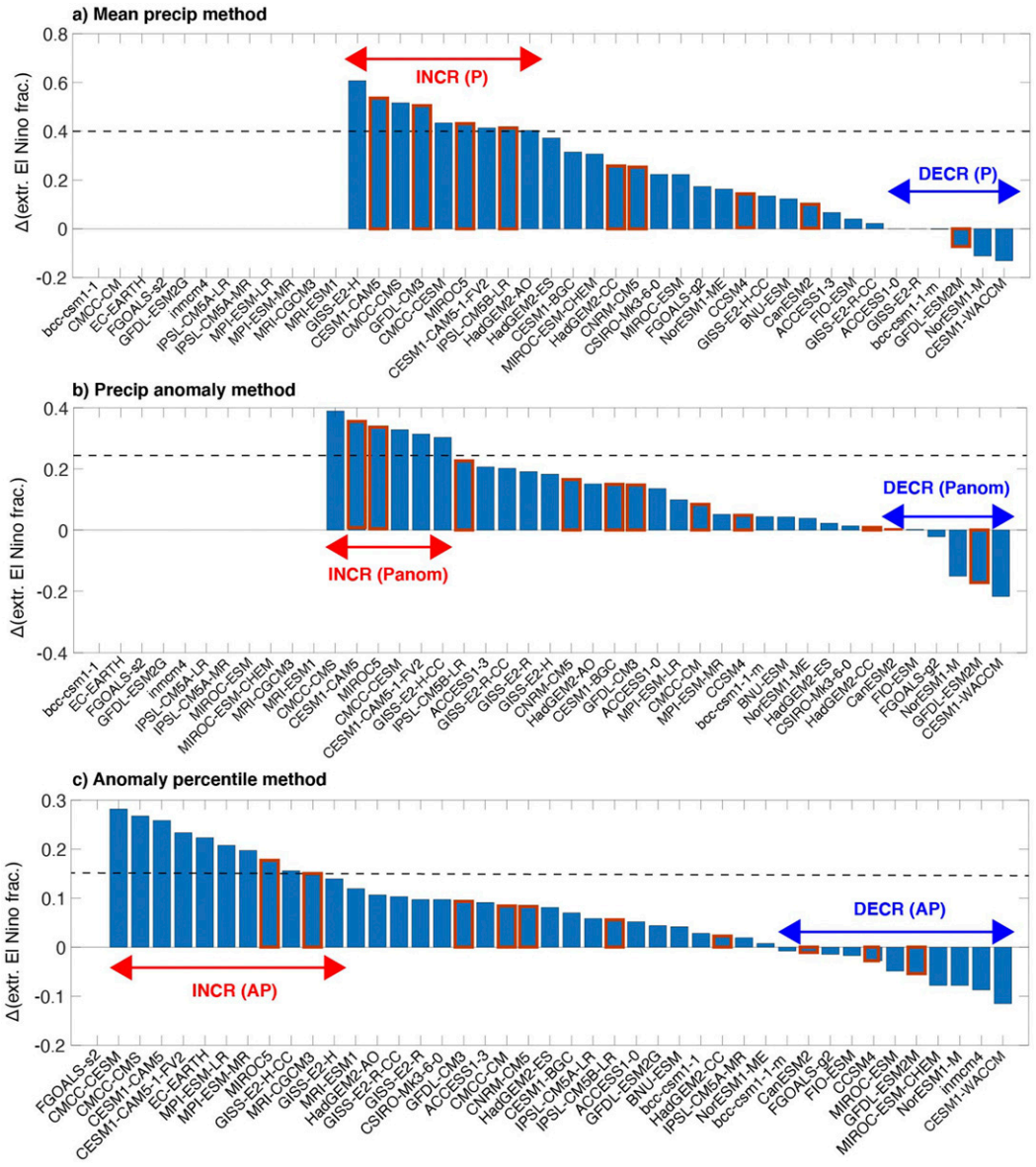

FIG. 3. Changes in proportion of extreme El Niño events over Niño-3, using the (a) absolute precipitation (*P*), (b) precipitation anomaly (Panom), and (c) anomaly percentile (AP) definitions. Red outlines indicate models selected by Cai et al. (2014) as having the capacity to best simulate El Niño rainfall and rainfall skewness. Dashed horizontal lines indicate thresholds used to select the INCR subsets of models according to all three definitions. In (a) and (b), fewer models are included since the absolute precipitation or precipitation anomaly for some models fell below the appropriate threshold for all El Niño events. Models with larger overall precipitation variability will tend to show smaller changes in (c), since the percentile changes will tend to be smaller.

century, however. Determining the exact causes for the intermodel disagreement here is beyond the scope of the present study; we hypothesize that differences in atmospheric sensitivity to SST anomaly patterns, and the expression of subtropical/midlatitude El Niño teleconnections, may play a role in setting the patterns of extreme dry El Niño.

We have created model subsamples based on each extreme El Niño definition, for later use in comparing mean-state behavior across models (section 5). The DECR subset of models are simply those that project a decrease in the frequency of extreme El Niño, while the INCR models are those that project the strongest increase. For the absolute precipitation method, this threshold is a 40% increase; for the precipitation anomaly method, a 25% increase; and for the anomaly percentile method, a 15% increase in extreme wet El Niño over the Niño-3 region. The models which belong to each of these subgroups, along with their respective numbers of ensemble members in CMIP5, are summarized in Table 1.

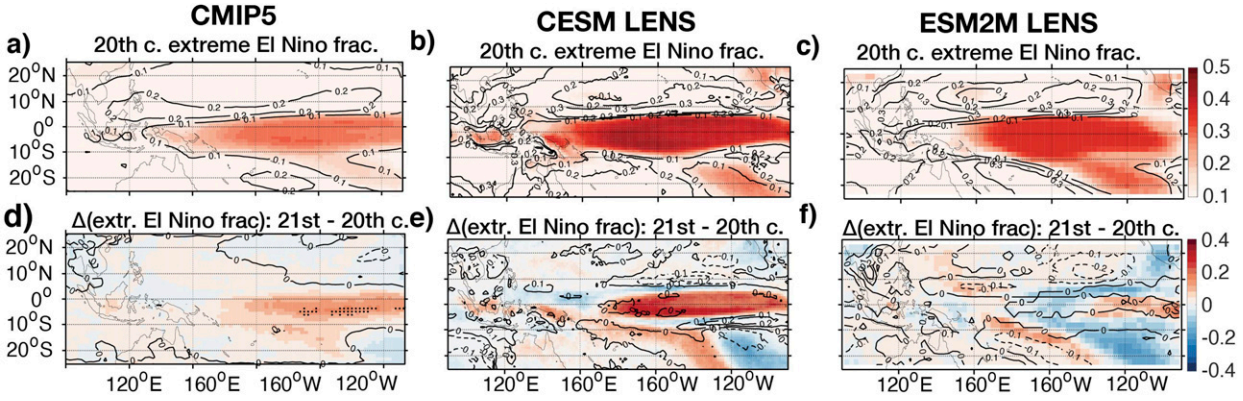

FIG. 4. Changes in local extreme El Niño event fraction using the anomaly percentile definition. (a) Twentieth-century extreme El Niño fraction in CMIP5. (b) As in (a), but for CESM LENS. (c) As in (a), but for ESM2M LENS. (d) Change in extreme El Niño fraction in the twenty-first century relative to the twentieth in CMIP5. (e) As in (d), but for CESM LENS. (f) As in (d), but for ESM2M LENS. Colors indicate changes in the frequency of precipitation anomaly exceeding the 90th percentile of twentieth-century precipitation; contours indicate changes in the frequency of precipitation anomaly falling below the 10th percentile of twentieth-century precipitation. Dashed contours indicate negative values.

4. Mechanisms for changes in precipitation during El Niño

To understand the reasons for intermodel differences in extreme El Niño projections, we begin by assessing the relative contribution from changes to twenty-first-century SST variability. This is done by statistically modeling the precipitation anomaly during El Niño as a function of SST, building on the probability distribution function (PDF)-based approach of Watanabe and Wittenberg (2012) and Watanabe et al. (2012). This method decomposes the twentieth- to twenty-first-century change in El Niño precipitation anomalies into components associated with

1) the change in magnitude of El Niño–related SSTA;
2) the change in the sensitivity of precipitation anomalies to a given SSTA; and
3) nonlinear interactions between 1 and 2 (e.g., changes in the behavior of convective initiation; Johnson and Xie 2010; Huang and Xie 2015).

This method has been previously used to diagnose the contribution of ENSO to mean precipitation (Watanabe and Wittenberg 2012) and to isolate the impact of a changing mean state on ENSO in perturbed-physics ensembles (Watanabe et al. 2012). Here, we will advance upon previous studies by examining the spatial structure of each individual source of changes in El Niño–driven precipitation anomalies. Specifically, rather than modeling regionally averaged absolute precipitation as a function of absolute temperature, we consider the (spatially varying) relationship between DJF precipitation anomalies and DJF Niño-3 temperature anomalies relative to the seasonal climatology:

$$P'_{\text{nino}} = \int f(T')C(T')\,dT'. \tag{1}$$

Here P'_{nino} is the local composite precipitation anomaly during DJF of the peak of El Niño events; $f(T')$ refers to the probability distribution function of El Niño SST anomalies over Niño-3, and $C(T')$ is the composite-mean local El Niño precipitation anomaly as a function of the Niño-3 SST anomaly. Note that $C(T')$ is constructed by selecting Niño-3 SSTA values within specified bins, and computing the average precipitation associated with those time periods. As in the rest of this analysis, the twentieth century is defined as 1950–2005 and the twenty-first as 2006–2100.

As in previous work using similar approaches (Watanabe et al. 2012), we consider precipitation anomaly changes relative to a reference (in this case, the twentieth-century precipitation anomalies during El Niño, hereafter P'_{nino20}). Differences between the twenty-first- and twentieth-century values of a quantity are hereafter referred to as Δ. The change in precipitation anomaly $\Delta P'_{\text{nino}}$ is then modeled using the separate contributions of changes in $f(T')$, $C(T')$, and their interactions:

$$\Delta P'_{\text{nino}} = \underbrace{\int \Delta f C_0(T')\,dT'}_{\text{Changes to El Niño SSTA}} + \underbrace{\int f_0 \Delta C(T')\,dT'}_{\text{Changes to } P \text{ sensitivity to SSTA}} + \underbrace{\int \Delta f \Delta C(T')\,dT'}_{\text{Nonlinear interactions}}. \tag{2}$$

The twentieth-century reference PDF f_0 and the reference composite $C_0(T')$ are computed individually for each CMIP5 model (for all ensemble members, in the case of multiple simulations from that model), and the change in PDF $\Delta f(T')$ and change in composite $\Delta C(T')$ computed from the deviations relative to those references. Multimodel mean contributions are then computed by averaging the values of each term in (2) over the CMIP5 simulations and the CESM/ESM2M large ensembles. It should be noted that *all* El Niño events are included here (i.e., all events where DJF Niño-3 SSTA exceeds 0.5σ regardless of rainfall amount), rather than only those qualifying as extreme according to previous sections.

The PDF method successfully reproduces the overall twenty-first-century changes in precipitation anomaly during El Niño, both for individual CMIP5 models and across the CMIP5

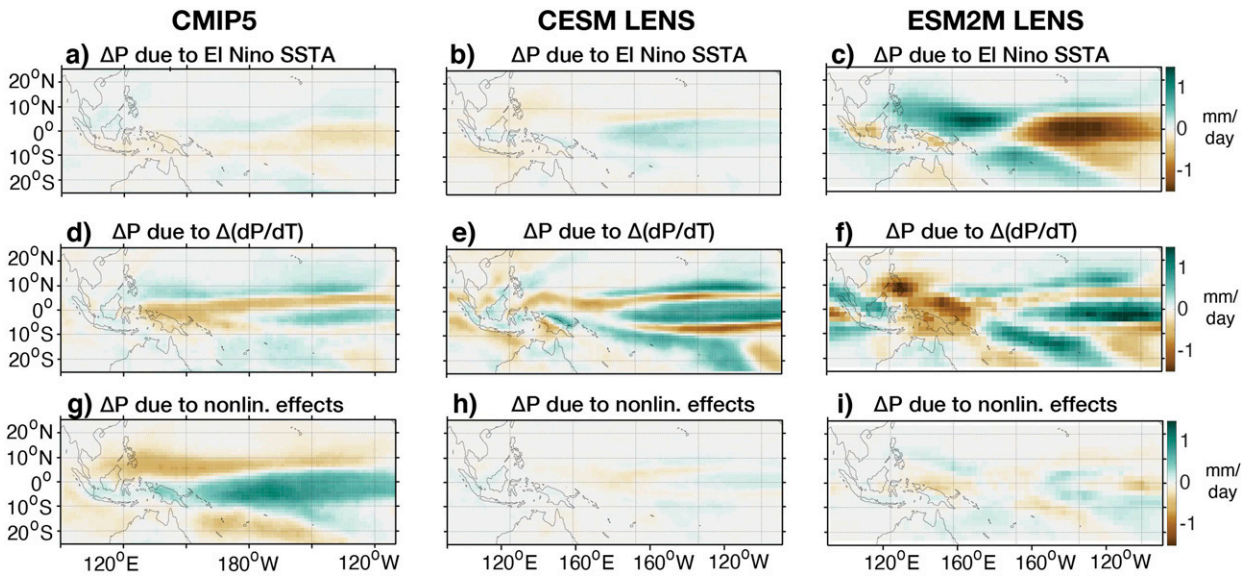

FIG. 5. Mechanisms for changes in El Niño precipitation. (a) Change in twenty-first- vs twentieth-century El Niño precipitation due to changes in El Niño SSTA in CMIP5, as estimated by the PDF method [Eq. (2)]. (b) As in (a), but for CESM LENS. (c) As in (a), but for the ESM2M LENS. (d) As in (a), but for precipitation changes due to changes in sensitivity of precipitation to SSTA in CMIP5. (e) As in (d), for CESM LENS. (f) As in (d), but for ESM2M LENS. (g) Changes in precipitation due to nonlinear interactions between SSTA and precipitation sensitivity to SSTA in CMIP5. (h) As in (g), but for CESM LENS. (i) As in (g), but for the ESM2M LENS. The twentieth century is defined as 1950–2005, and the twenty-first century as 2006–2100, for all models. El Niño events are defined as DJF periods where Niño-3 SSTA, deseasonalized with a 30-yr centered moving window, exceeds 0.5σ. In both CESM and ESM2M, the twenty-first-century portion of the simulations follows RCP8.5.

archive as a whole (not pictured). Due to projected changes in the sensitivity of precipitation to Niño-3 SSTA [$\Delta C(T')$, second term in Eq. (2)], the CMIP5 multimodel archive shows substantial increases in precipitation during El Niño along the equator in the central/eastern Pacific, with reductions to the north and south (Fig. 5d). This may be interpreted as an increase in the tendency for warm equatorial SSTAs to drive convective precipitation events in the future. In the multimodel mean, SSTA changes have a slight tendency to *decrease* El Niño precipitation in the eastern equatorial Pacific (Fig. 5a). However, the strongest contributor to CMIP5 multimodel mean El Niño precipitation changes is the nonlinear term (Fig. 5g), which drives a large precipitation increase in the eastern equatorial Pacific.

We now turn to the two large ensembles, and apply the same analysis. CESM and ESM2M form an especially useful contrast since they are both models with highly consistent extreme El Niño frequency changes across all definitions, yet their responses are diametrically opposed to one another (Table 1, Fig. 3) as are their projected changes in ENSO amplitude (Fasullo et al. 2018). As in CMIP5, both CESM and ESM2M show increases in equatorial precipitation during El Niño due to precipitation sensitivity to SSTA, with reductions off the equator (Figs. 5e,f). Interestingly, although CESM shows a substantial increase in SSTA during El Niño (Stevenson et al. 2017), this does not appear to play a major role in driving increased precipitation extremes in this model (Fig. 5b). This contrasts with the large contribution of changing SSTA to precipitation in ESM2M, where the reduction in SSTA variance strongly reduces associated precipitation (Fig. 5c). In both cases, the nonlinear term is relatively unimportant, with the exception of a small region in the central/eastern Pacific that is likely to be near the threshold for convective initiation (Figs. 5h,i).

The contrast between the CMIP5 and LENS results in terms of the importance of nonlinear effects may arise from the small ensemble size of most CMIP5 submissions. With only a few simulations per model in most cases, deviations from the ''reference'' distributions of SSTA and precipitation may be preferentially partitioned into the nonlinear term due to undersampling during construction of those distributions. This work cannot definitively attribute the reason for this outcome; regardless, these results highlight the importance of atmospheric sensitivity to SSTA, via the precipitation sensitivity to SSTA or the nonlinear term, in setting the ENSO precipitation response to climate change (He et al. 2018).

Figure 6 gives a more detailed breakdown of contributions from the terms in the PDF method when averaged over the Niño-3 region; note that here, all terms are computed using spatial gridpoint values prior to regional averaging. The majority of CMIP5 models show a positive but weak contribution from precipitation sensitivity to SSTA (Fig. 6b); however, there is a large degree of scatter in this relationship. A stronger relationship is seen with the other two terms (Figs. 6a,c)—strikingly, the nonlinear contribution is positive in nearly all CMIP5 models (Fig. 6c), whereas the SSTA contribution is roughly evenly split

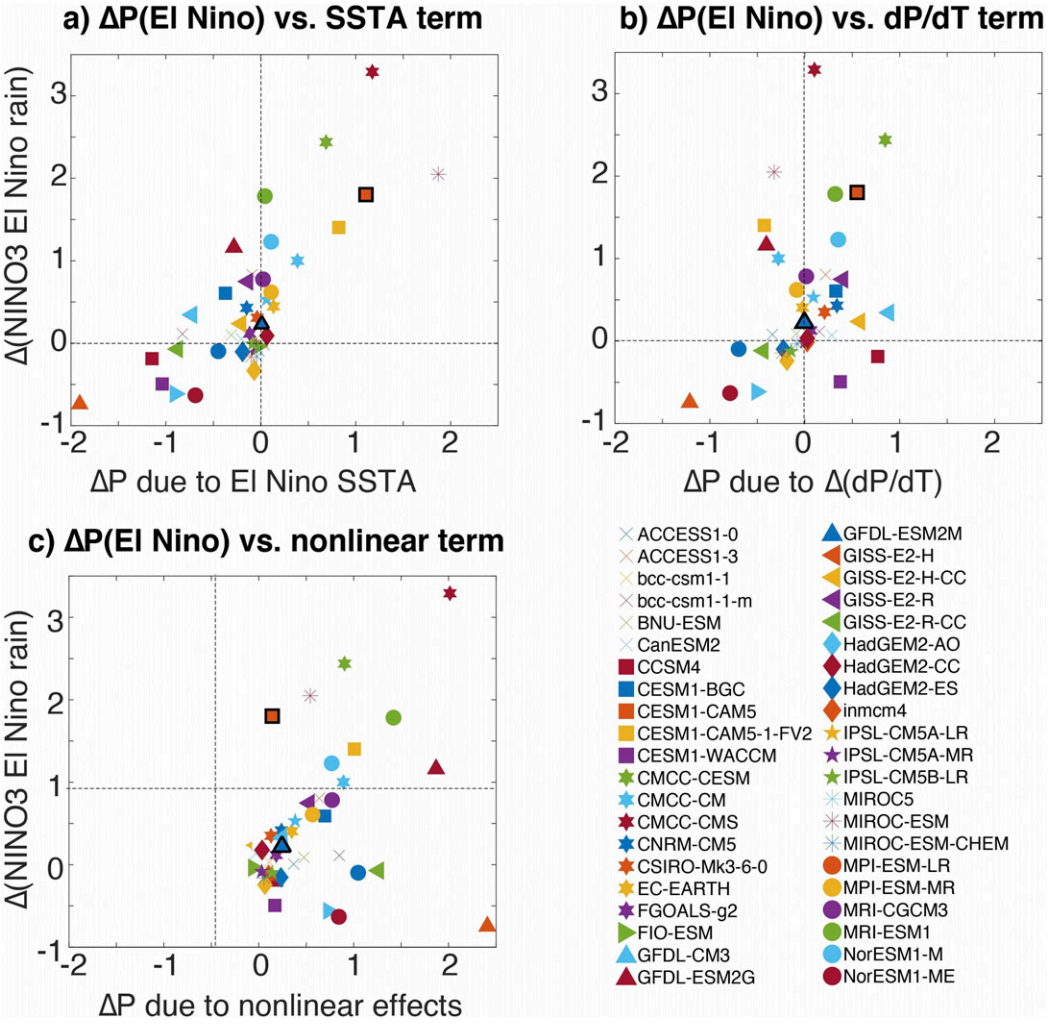

FIG. 6. Mechanisms for changes in El Niño precipitation over Niño-3. (a) Changes in El Niño precipitation vs contributions of SSTA to precipitation changes, computed from the PDF decomposition. (b) As in (a), but for contributions of the sensitivity of precipitation to SSTA from the PDF decomposition. (c) As in (a), but for contributions of the nonlinear term from the PDF decomposition. CESM and ESM2M CMIP5 simulations are indicated as black-outlined symbols in (b) and (c).

between positive and negative values (Fig. 6a). We thus conclude that any disagreements between projections of extreme El Niño precipitation in CMIP5 are likely to be driven by this changing SSTA contribution. This is borne out by the nearly linear relationship between changes in Niño-3 SSTA variance and changes in the frequency of extreme wet El Niño over the Niño-3 region (Figs. 7a–c). These results strongly suggest that there remains a need to understand intermodel differences in SSTA projections, even when using precipitation-based metrics to characterize future changes in ENSO.

5. Links between the mean state and extremes

We next assess the relation between mean climate and extreme El Niño. Models that project the largest future increases in extreme El Niño frequency (the INCR models) simulate systematically colder and drier historical conditions in the central to eastern equatorial Pacific relative to those in which extreme El Niño frequency decreases (DECR; Figs. 8a,c,e). The INCR models also systematically have stronger mean zonal equatorial SST gradients across the Pacific during the historical period. The twenty-first-century mean-state change is quite different in the INCR models relative to DECR as well; warming is much stronger in the INCR models over the central to eastern equatorial Pacific, and is accompanied by increases in mean precipitation (Figs. 8b,d,f). These effects are largest in the INCR models selected using the absolute precipitation definition (Fig. 8a), but are present when using all three definitions.

Taken together, the results of Fig. 8 would suggest that El Niño–like mean-state changes tend to favor increases in extreme El Niño frequency; and that colder conditions in the

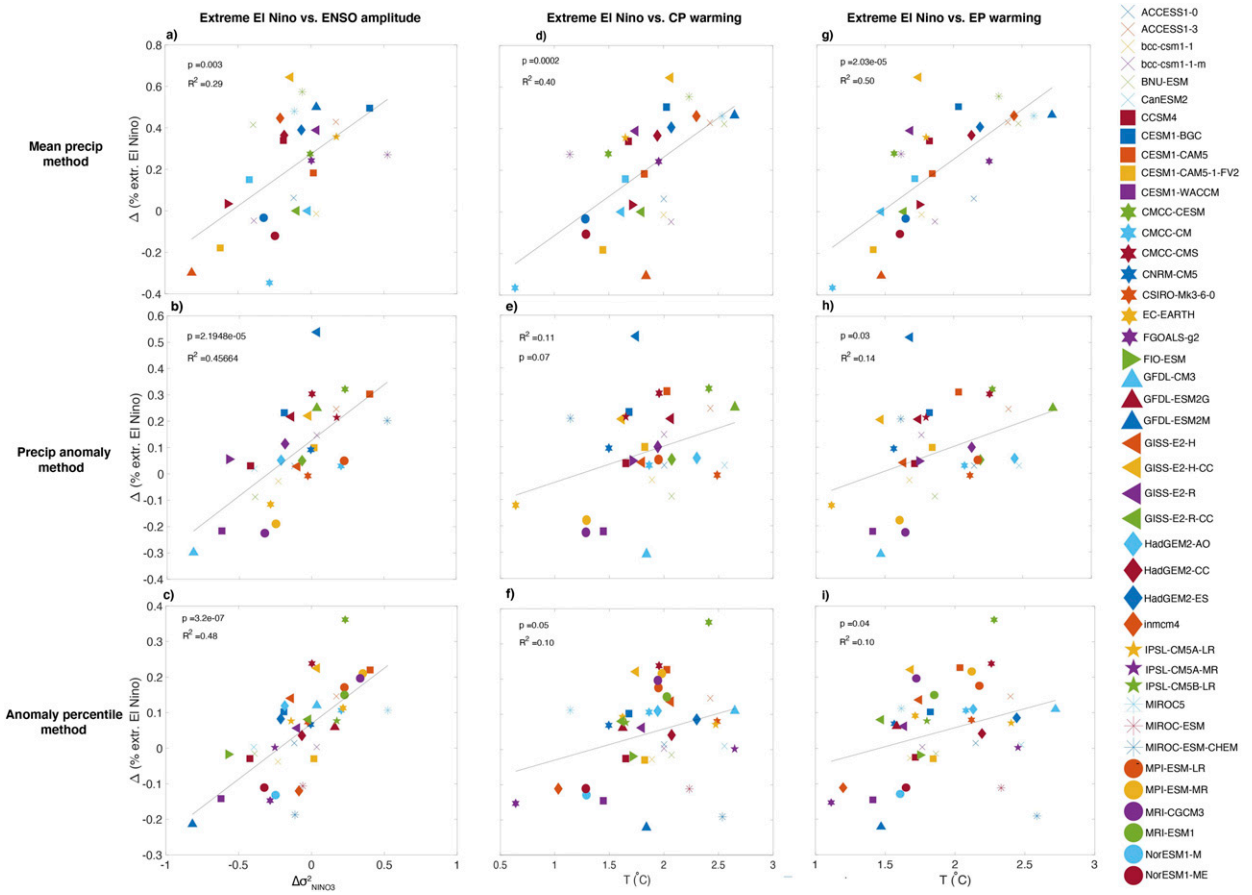

FIG. 7. Relationship of changes in extreme El Niño fraction with mean and variance changes (twenty-first century − twentieth century) in the CMIP5 models. (a) Extreme El Niño fraction vs SSTA variance changes, using the absolute precipitation method. (b) As in (a), but for the precipitation anomaly method. (c) As in (a), but for the anomaly percentile method. (d) Extreme El Niño fraction vs central Pacific SST changes, using the absolute precipitation method. (e) As in (d), but for the precipitation anomaly method. (f) As in (d), but for the anomaly percentile method. (g) Extreme El Niño fraction vs eastern Pacific SST changes, using the absolute precipitation method. (h) As in (g), but for the precipitation anomaly method. (i) As in (g), but for the anomaly percentile method. Here central Pacific SST is the average over 2.5°S–2.5°N, 160°E–120°W; eastern Pacific SST the average over 0°–10°S, 85°–105°W.

central and far eastern equatorial Pacific during the historical period, in turn, favor El Niño–like warming. To determine how well this conclusion holds across the CMIP5 archive as a whole, we examine regional changes in two locations chosen to isolate major SST biases: the far eastern equatorial Pacific (0°–10°S, 85°–105°W), where SST is typically warmer than observed in most models; and the central equatorial Pacific (2.5°S–2.5°N, 160°E–120°W), where SST is typically colder than observed (Fig. 9a). In the eastern Pacific, warming trends are robustly related to increasing extreme El Niño frequency as measured by all three event definitions (Figs. 7g–i). Figure 10 then shows the relationship between twentieth-century bias and twenty-first-century warming: for the eastern Pacific, twenty-first-century warming is negatively correlated with SST bias in that region (Fig. 10b). In the central Pacific, twenty-first-century warming is also robustly related to increases in extreme El Niño, although the variance explained by this relationship is lower than for the eastern Pacific (Figs. 7d–f). Central Pacific SST bias is negatively correlated with twenty-first-century central Pacific SST warming (Fig. 10a). Colder historical central and eastern equatorial Pacific SSTs clearly do favor El Niño–like warming, and in turn, increases in extreme El Niño frequency.

The next question is to what degree mean-state biases alter extreme El Niño projections; this requires a consideration of cloud radiative feedbacks (Ceppi et al. 2017). In the eastern equatorial Pacific near the South American coast, this leads to underestimation of the *positive* cloud–SST feedback. The far eastern Pacific hosts extensive decks of stratus clouds, associated with climatological upwelling off the South American coast. In such stratus-dominated regimes, the relationship between cloud cover and SST is generally dominated by the positive stratus cloud–shortwave radiation–SST effect (Ramanathan and Collins 1991; Ying and Huang 2016). A future climatological warming of this region's SST would tend to reduce the local climatological stratus cover, thereby increasing insolation and amplifying the

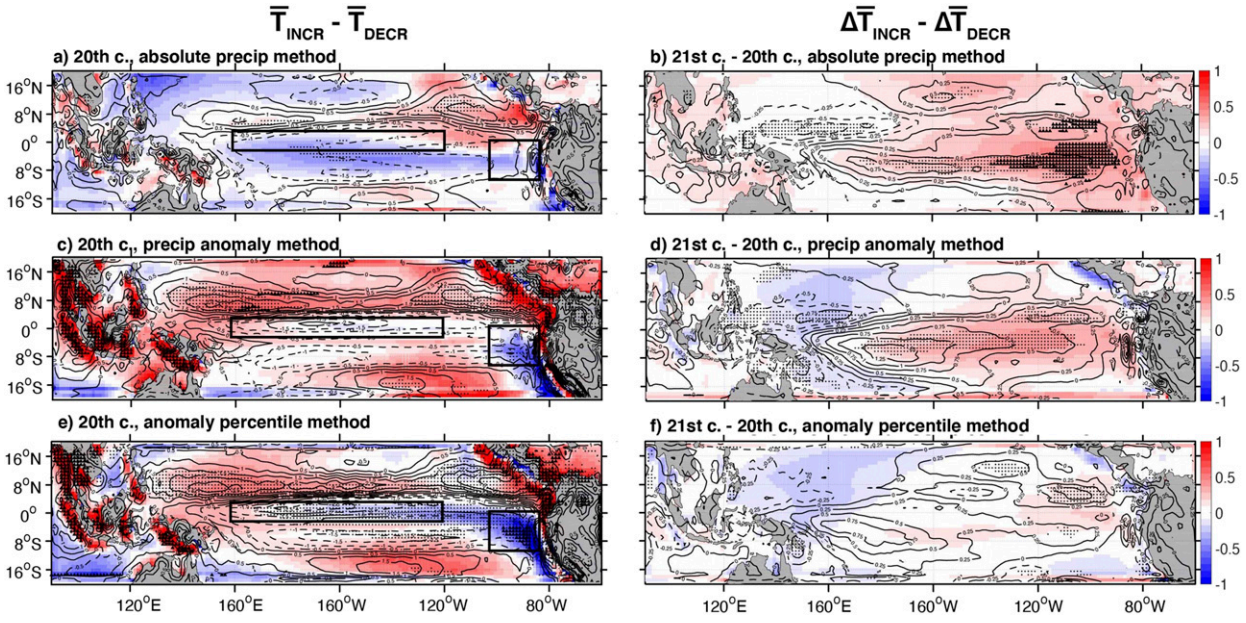

FIG. 8. Relationships between mean responses and changes to extreme El Niño frequency in CMIP5. (a) Difference in twentieth-century mean SST and precipitation between the INCR and DECR model subsets, selected using the absolute precipitation method. (b) Difference in the twenty-first-century − twentieth-century change in mean SST and precipitation between the INCR and DECR model subsets, selected using the absolute precipitation method. (c) As in (a), but for model subsets selected using the precipitation anomaly method. (d) As in (b), but for model subsets selected using the precipitation anomaly method. (e) As in (a), but for model subsets selected using the anomaly percentile method. (f) As in (b), but for model subsets selected using the anomaly percentile method. Colors indicate changes in mean SST; contours indicate changes in mean precipitation in the twenty-first century relative to the twentieth. Dashed lines indicate negative values. Stippling indicates differences are statistically significant at 90% using a rank-sum test on ensemble means: triangles show significant SST differences and circles show precipitation differences. Boxes in (a), (c), and (e) indicate the eastern and central Pacific regions used in Figs. 7 and 10.

climatological warming. This is a positive feedback that tends to amplify the future warming of climatological SST in the eastern equatorial Pacific; it will be too strong in models that start with excessive climatological stratus in the eastern equatorial Pacific, leading to excessive climatological warming of the surface. In the central equatorial Pacific, a convective regime dominates; here, the feedback between shortwave radiation and SST is negative, since warmer SSTs lead to stronger convection, thicker cloud cover, and a reduction in surface shortwave heating. A future climatological warming of *this* region's SST would tend to *increase* the local climatological cumulus cover, thereby decreasing insolation and reducing the climatological warming. This negative feedback will tend to be too weak in models that start with overly cold SST and therefore insufficient climatological cumulus in the central equatorial Pacific; without enough cumulus cover present to mitigate future warming, climatological SST will increase excessively. Since both the central and eastern equatorial Pacific are colder in the INCR models, this will lead to both weaker convective cloud generation and a thicker stratus deck, which have distinct influences on future projections.

To illustrate the importance of cloud feedbacks to extreme El Niño in CMIP5, we follow the methods of Ying and Huang (2016) to examine model behavior in the historical period. The "cloud shortwave feedback index" (CSFI) is computed by regressing twentieth-century (1981–2000) surface shortwave anomalies onto local SSTA, after first removing the tropical Pacific regional mean (20°S–20°N, 120°E–90°W). When this is done, the CMIP5 multimodel mean clearly shows the stratus- and cumulus-dominated regimes described above (Fig. 11a). In the INCR model subset, the negative CSFI values in the central Pacific are substantially weaker, as illustrated by the positive anomalies in Figs. 11b–d; these weaker negative feedbacks then increase local climatological warming (Figs. 8b,d,f). In the eastern Pacific, the INCR models clearly exhibit a more positive cloud–SST feedback, consistent with their smaller overall SST biases in this region; this favors the stronger twenty-first-century climatological eastern Pacific warming observed in these models.

Biases in the negative cloud–SST feedback in the central equatorial Pacific have previously been cited as causing intermodel spread in SST warming patterns in CMIP5 (Ying and Huang 2016; Zheng et al. 2016), but were hypothesized to favor La Niña–like mean changes. This was ascribed to overly strong central Pacific heating, driving a convergent wind response which cooled eastern Pacific SSTs via enhanced evaporation. Previous studies therefore concluded that *correcting* cloud–SST biases should result in a more El Niño–like mean state change. Our results, in contrast, show that INCR models tend to have *weaker* negative cloud–SST feedback biases *and* an El Niño–like warming and eastward shift in equatorial convection, suggesting that common model biases may lead to an

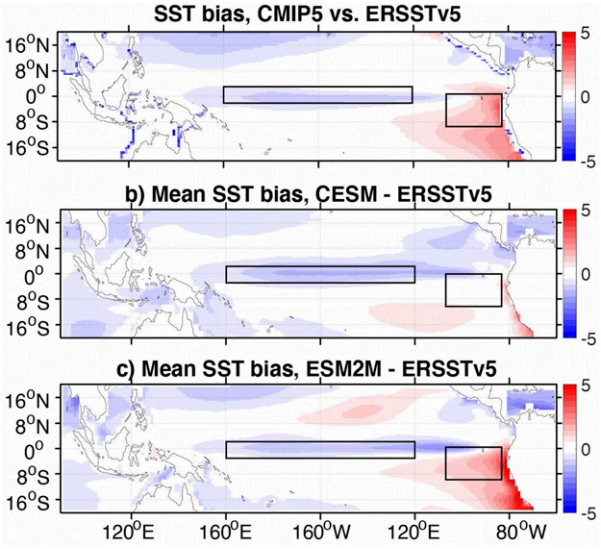

FIG. 9. Biases in SST relative to observations (ERSSTv5): (a) CMIP5 ensemble, (b) CESM LENS, and (c) ESM2M LENS. Differences are calculated over the 1950–2005 period. Boxes indicate eastern and central Pacific regions used in Figs. 7 and 10.

overestimate of future El Niño–driven precipitation extremes. However, the effect of eastern Pacific SST biases may act as a compensatory mechanism; correcting these biases would lead to colder conditions, which Fig. 10 shows is associated with a stronger twenty-first-century local warming in the region. This enhanced warming, in turn, favors an amplification of future extreme El Niño (Fig. 7). Model biases in the eastern Pacific, in other words, may lead to an underestimate of extreme El Niño; these two potential mechanisms are summarized schematically in Fig. 12.

We note that the model groupings here were constructed differently than in the Ying and Huang study; INCR models are chosen based on their relative increase in extreme El Niño, rather than the extent of historical SST biases. Our method tends to select for INCR models with larger twenty-first-century SSTA variance increases (Figs. 6 and 7), and with less twentieth-century rainfall in Niño-3 (Fig. 8); in the former case, SSTA variance amplifies rainfall variance, and in the latter, a smaller starting rainfall amount favors a larger relative increase in extremes. However, our results nonetheless point to a delicate balance between local cloud feedback effects and the remote wind response documented by Ying and Huang (2016). In particular, the zonal position of the biases in shortwave–SST feedbacks may be crucial to the overall influence on SST pattern change; if a model has a very strong stratus regime in the eastern equatorial Pacific (as in INCR), the local cloud–SST feedback will dominate the response to climate change, creating a more El Niño–like SST warming pattern. However, if a model has a less stratus-dominated eastern equatorial Pacific and insufficient convection in the central equatorial Pacific, the SST warming will be enhanced in the central Pacific relative to the east, thereby enhancing the zonal equatorial SST gradient, easterly winds, evaporation, upwelling, and cooling in the eastern equatorial Pacific. In such cases, this same feedback can lead to the previously identified coupled wind response (favoring a more La Niña–like warming pattern). The details of the wind response will depend sensitively on the location and magnitude of the SST warming; however, our results suggest that there may be multiple ''regimes'' occupied by CMIP5 models, which respond differently based on their dominant cloud and convective characteristics.

6. Emergent constraints

These results suggest that twentieth-century simulated climate may serve as a source of predictive power, by providing ''emergent constraints'' on the projected range of twenty-first-century changes to extreme El Niño. Interestingly, despite the strong association between ENSO amplitude changes and extreme El Niño frequency shifts over the full twenty-first century, the change to ENSO amplitude over the more recent

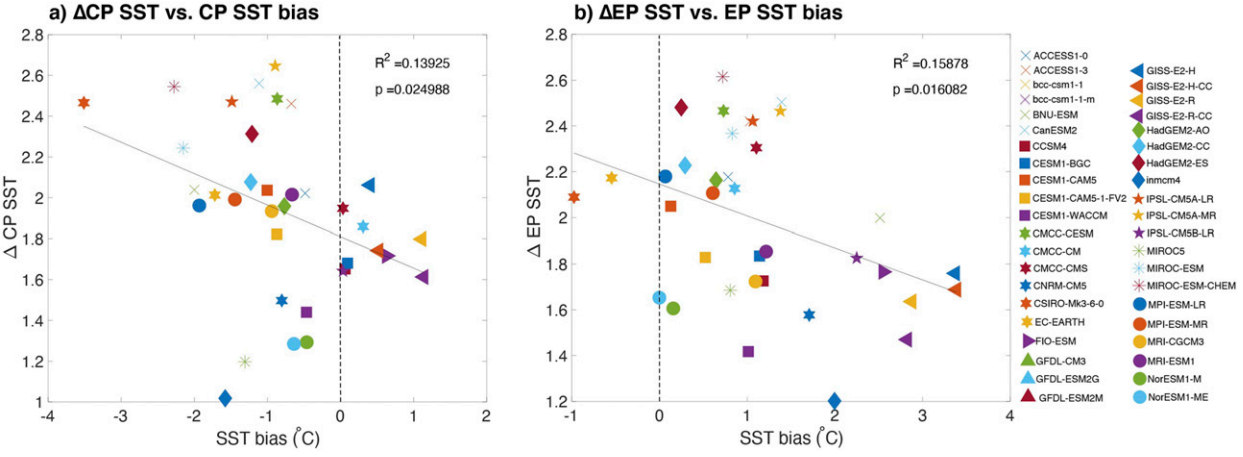

FIG. 10. Relation of twenty-first- minus twentieth-century mean SST changes with twentieth-century SST biases: (a) central Pacific (2.5°S–2.5°N, 160°E–120°W) and (b) eastern Pacific (0°–10°S, 85°–105°W).

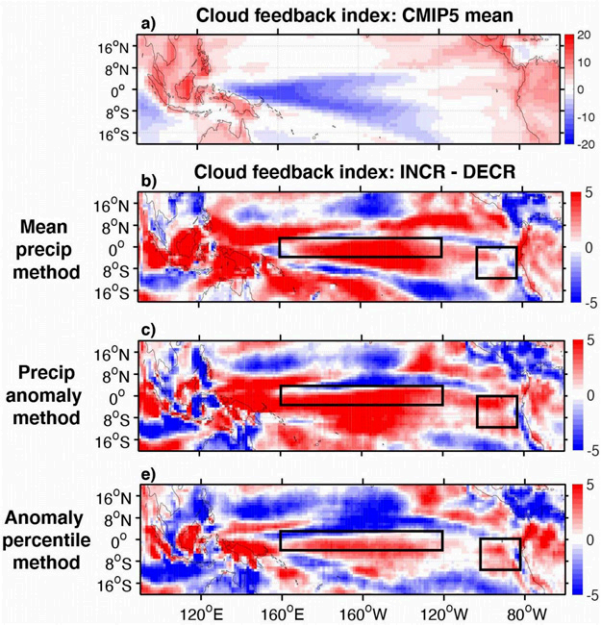

FIG. 11. (a) Cloud shortwave feedback index (CSFI; defined in main text; $\mathrm{W\,m^{-2}\,K^{-1}}$), CMIP5 multimodel mean. (b) CSFI differenced between the model subsets (INCR − DECR) as selected based on the absolute precipitation method. (b) Cloud fraction regressed on SSTA, differenced between INCR and DECR model subsets as selected based on the absolute precipitation method. (c) As in (a), but for model subsets selected based on the precipitation anomaly method. (d) As in (b), but for model subsets selected based on the precipitation anomaly method. (e) As in (a), but for model subsets selected based on the anomaly percentile method. (f) As in (b), but for model subsets selected based on the anomaly percentile method. All analyses cover the 1981–2000 period, following Ying and Huang (2016). Boxes indicate the eastern and central Pacific regions used in Figs. 7 and 10.

period does not appear to exhibit predictive power (not pictured). This is most likely due to the known large natural modulations in ENSO amplitude over 20-yr periods, as were used for the computations; natural ENSO variability has been previously noted to be substantial on time scales up to several centuries (Wittenberg 2009; Stevenson et al. 2010; Stevenson 2012). Rather, the most effective emergent constraint metrics seem to be the change in the mean equatorial Pacific zonal SST gradient ($\partial T/\partial x$) and the change in precipitation sensitivity to SST ($\partial P/\partial T$). These metrics are inspired by the terms in the PDF decomposition of section 4, and are thus clearly strongly related to changes in extreme El Niño frequency (Figs. 5, 6, and 7).

For an emergent constraint to be effective, it must have predictive power for the variable of interest (here, the change in extreme El Niño frequency), and the independent variables must be observable quantities that can be applied as benchmarks for climate model simulations. The time period of greatest interest for our independent variables is therefore the satellite era (1979–2019). The form of the emergent constraint relation is

$$\Delta_{\mathrm{dep}} f = \beta_0 + \beta_1 \Delta_{\mathrm{ind}} \frac{\partial T}{\partial x} + \beta_2 \Delta_{\mathrm{ind}} \frac{\partial P}{\partial T}, \tag{3}$$

where f indicates the occurrence frequency of extreme El Niño, and Δ values indicate the change in either independent (ind) or dependent (dep) variables over a specified time period. The time periods used for dependent and independent variables need not be identical; indeed, for a true emergent constraint, the independent variables should hold predictive power over time periods longer than those over which they are measured. As an initial test of the method, we first examine predictions made using differences between the full (postsatellite era) historical period and twenty-first century within the CMIP5 ensemble: both Δ_{ind} and Δ_{dep} are set to (2021–2100) − (1979–2020). In general, extreme El Niño changes identified using the absolute precipitation method show less skill; none of the regressions are statistically significant (Figs. 13a–c). This likely relates to the inherent complexities associated with predicting mean precipitation, where changes to simulated specific humidity and the seasonal cycle may be confounding factors. However, the regression relationship is highly significant for both the precipitation anomaly and anomaly percentile methods (Figs. 13d,g), explaining up to 50% of the variance in extreme El Niño frequency change. We conclude that this choice of predictor variable set is reasonable to use for derivation of an observational emergent constraint.

We next examine changes over the satellite era, setting Δ_{ind} and Δ_{dep} to (1979–90) − (1991–2019) (note that observational data extend only to early 2019). The choice of 1990 as a center point for dividing the observational record is made in order to avoid impacts from the early 2000s global-warming hiatus. Once again, significant skill is seen for the two anomaly methods (Figs. 13e,h), albeit explaining less variance than was the case for the full twenty-first century since the chosen epochs are shorter. Now it is possible to include observationally derived metrics (red vertical lines in Fig. 13); when the regression relationship is applied to observational estimates of $\partial T/\partial x$ and $\partial P/\partial T$, observations predict much smaller changes in extreme El Niño frequency than the CMIP5 models simulate. This is consistent with the hypothesis of section 5, indicating that correcting biases in CMIP5 models may tend to reduce the previously cited (Cai et al. 2014) increase in extreme El Niño occurrence.

Finally, we consider the case where Δ_{ind} and Δ_{dep} differ; here Δ_{ind} is kept at (1979–90) − (1991–2019) to approximate the observational era, whereas Δ_{dep} is extended to (1991–2070) − (1979–90) to form a true emergent constraint. Once again, predictive skill remains in the anomaly methods (Figs. 13f,i), and the observationally derived estimate lies within the spread of the CMIP5 model predictions. This first step toward constraining extreme El Niño projections indicates that it may be possible to use observational constraints to narrow the spread in CMIP5 projections.

7. Discussion and conclusions

This work represents a new, systematic analysis of changes to El Niño–driven precipitation and associated changes to

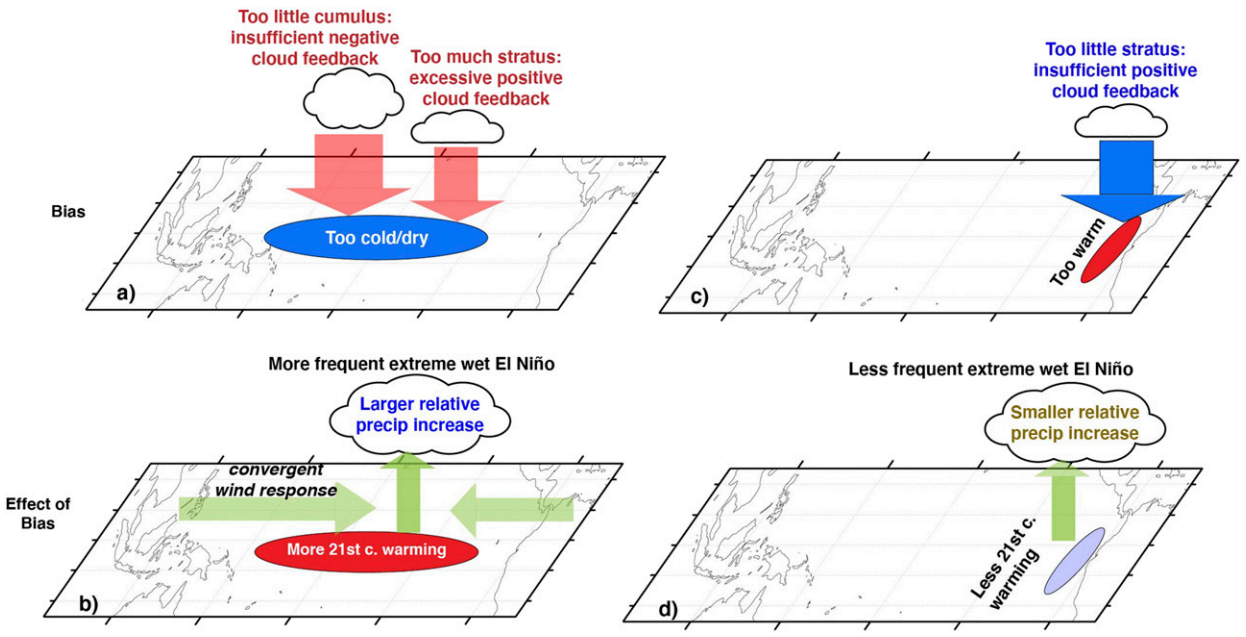

FIG. 12. Schematic relating mean-state biases to cloud feedbacks and changes in extreme El Niño occurrences. (a) Twentieth-century central Pacific bias: Weakened cumulus formation creates an insufficiently negative cloud–SST feedback; farther east; stratus enhancement leads to an overly strong positive cloud–SST feedback. (b) Twenty-first-century atmospheric changes due to central Pacific bias: both the reduced cumulus and enhanced stratus biases create positive feedback anomalies, which enhance warming in the cold tongue region. The convergent wind response leads to coastal cooling in the eastern Pacific. The dry twentieth-century bias favors a stronger relative precipitation increase, leading to larger percentage changes in extreme wet El Niño occurrence. (c) Twentieth-century eastern Pacific bias: weakened stratus cover creates an overly weak positive feedback. (d) Twenty-first-century atmospheric changes due to eastern Pacific bias: the reduced stratus creates a negative feedback anomaly, which reduces warming near the South American coast and suppresses precipitation.

"extreme El Niño" frequency across the entire suite of CMIP5 simulations. When changes to extreme El Niño are examined on a case-by-case basis, we find that there is a wide range in projections of twenty-first-century event frequency. Notably, this is true even if one considers only models previously assessed as capable of realistically simulating deep convection (Cai et al. 2014), and holds across multiple different definitions of extreme El Niño. The majority of models do indeed simulate increases to extreme El Niño frequency, but a substantial minority do not. Additionally, among models that do simulate an increase, the magnitude of that increase varies dramatically. This points to the need to better understand drivers of intermodel differences, to narrow the spread in future projections.

Using a new application of a PDF-based decomposition method (Watanabe and Wittenberg 2012), we have shown that changes to El Niño–driven precipitation are affected by several factors. The previously identified tendency for precipitation to become more sensitive to SSTA in a warming climate (Power et al. 2013; Bonfils et al. 2015) contributes substantially, creating an increase in equatorial precipitation during El Niño. However, changes in the SSTA variance also contribute strongly to precipitation responses, and our results show that SSTA changes dominate intermodel diversity in El Niño–driven precipitation projections. As SSTA variance projections continue to vary widely in current-generation models (Stevenson 2012; Bellenger et al. 2014; Fasullo et al. 2018), our work indicates that constraining the response of interannual SSTA variance to future climate change should remain a priority.

Our work also demonstrates the power of combining CMIP-style model intercomparisons with large single-model ensembles. The CESM and ESM2M large ensembles were run using the CMIP5 configuration of each model (Kay et al. 2015; Rodgers et al. 2015) and have diametrically opposed responses of ENSO variability and extreme El Niño changes, allowing them to function as "end members" for the full CMIP5 analysis. The contribution of SSTA variance to El Niño precipitation differs between the ensembles, tending to slightly amplify anomalies in CESM and strongly damp them in ESM2M. The sensitivity of extreme El Niño frequency to mean-state change is also different between the ensembles; in CESM, the zonal SST gradient response to climate change is tightly coupled with the frequency of extreme El Niño occurrence, whereas this dependence is much weaker in ESM2M. Notably, in both models the zonal SST gradient appears to have a larger influence than the previously cited effect of the meridional SST gradient (Cai et al. 2014). We hypothesize that the ITCZ migration mechanism proposed by Cai et al. (2014), while operative in some models, has a smaller controlling influence on emergent extreme El Niño frequency changes than zonally asymmetric feedbacks along the equator.

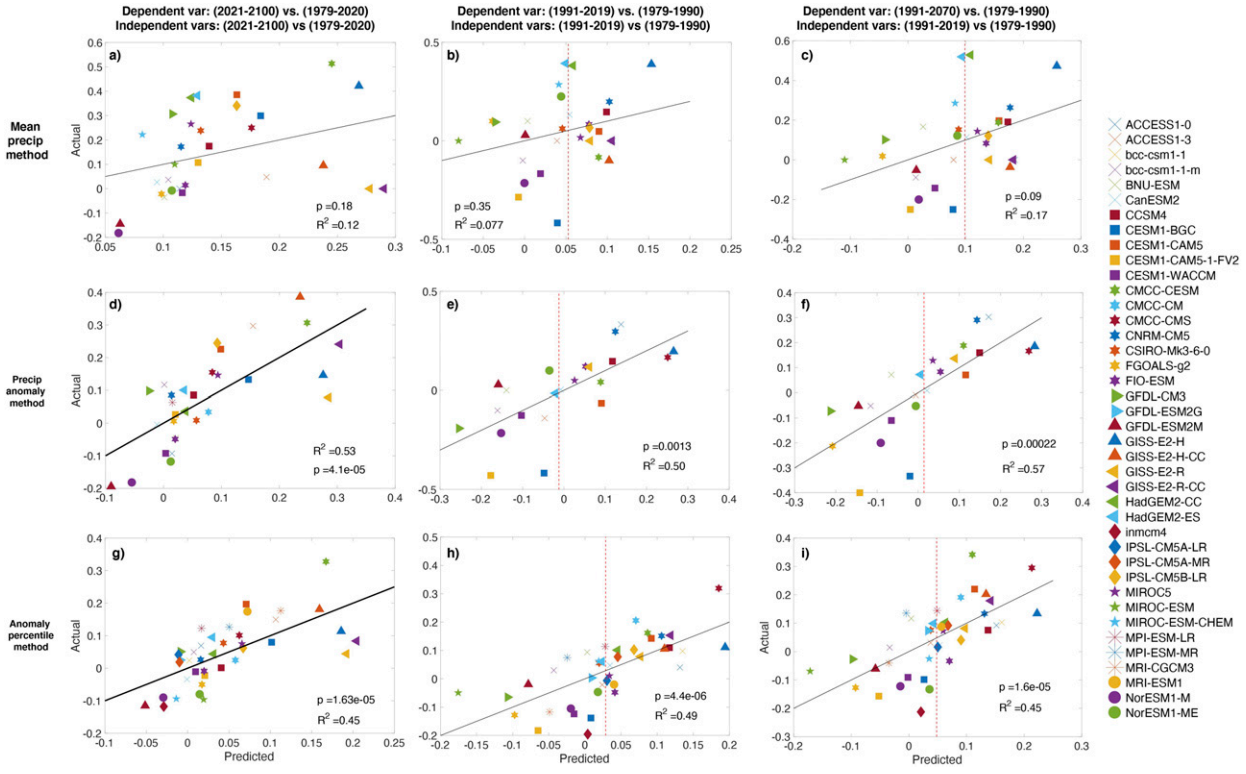

FIG. 13. Prediction of change in extreme El Niño fraction using observable metrics over the modern era: the zonal equatorial gradient of SST ($\Delta dT/dx$), and the sensitivity of precipitation to SSTA ($\Delta dP/dT$; regression of precipitation on SSTA over Niño-3). (a) Prediction of changes to extreme El Niño identified using the mean precipitation method, for both independent and dependent variable differences computed between 2021–2100 and 1979–2020. (b) As in (a), but for independent and dependent variables differenced between 2001–19 and 1979–2000. (c) As in (a), but for independent variables differenced between 2001–2019 and 1979–2000, and dependent variables differenced between 2001–2070 and 1979–2000. (d)–(f) As in (a)–(c), but for extreme El Niño events identified using the precipitation anomaly method. (g)–(i) As in (a)–(c), but for extreme El Niño events identified using the anomaly percentile method. Black solid lines indicate best-fit regressions; bold indicates a statistically significant relationship. Vertical red dashed lines indicate the predicted value generated using observational data.

The connection between the zonal SST gradient and extreme El Niño is expressed in part through the action of cloud shortwave feedbacks. Models that simulate larger increases in extreme El Niño frequency tend to have colder and drier twentieth-century climates, with overall larger zonal SST gradients. This enhances biases toward overly strong stratus coverage in the far eastern equatorial Pacific and weak cumulus coverage nearer the date line. Both of these effects reduce the negative surface shortwave feedback, facilitating stronger relative warming in the eastern part of the basin. In model subsets selected according to their precipitation responses, this effect seems to dominate over previously identified remote wind responses, which would counteract warming in the east Pacific (Ying and Huang 2016).

We note that there may be other influences affecting mean-state responses to climate change, not all of which could be addressed in the present study. Of particular interest is the potential for off-equatorial warming to create remote responses in the tropics, owing to a weakening of the Hadley circulation and oceanic subtropical cells with associated changes in equatorial SST (Stuecker et al. 2020). This mechanism may indeed affect the sensitivity of the cold tongue to future warming, through modification of the equatorial thermocline and associated SST anomalies communicated with the surface via vertical advection. However, even in a case where subtropical processes play a significant role, equatorial processes are expected to be fundamental in setting the regional behavior of SST pattern formation (Xie et al. 2010; Ying and Huang 2016).

We have demonstrated a first proof of concept of the use of precipitation and temperature-based metrics as emergent constraints on changes to extreme El Niño frequency. Changes in the amplitude of Niño-3 SSTA variability, the sensitivity of precipitation to SST, and the zonal SST gradient over the modern observational era (1980–2020) are used as predictors for overall twenty-first-century changes to extreme El Niño frequency, and the relationship is statistically significant for the two anomaly-based methods. The variance explained for the absolute precipitation method is much lower, likely a consequence of confounding factors complicating the prediction of mean-state increases. When observational data are used to estimate the predicted value of ''true'' future extreme El

Niño changes, the values lie near the median of the range simulated by CMIP5. This may indicate that CMIP5 models reliably capture the potential spread in extreme El Niño frequency; however, the presence of multiple confounding mechanisms related to cloud shortwave feedbacks presents a cautionary note and motivates future process-based investigations of twenty-first-century ENSO behavior.

Acknowledgments. This work is supported by an NSF EaSM grant, AGS-1243125 to NCAR and AGS-1243107 to the University of Arizona. SS was also supported by the U.S. Department of Energy, DE-SC0019418. The CESM project is supported primarily by the National Science Foundation (NSF). This material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the NSF under Cooperative Agreement 1852977. Computing and data storage resources, including the Cheyenne supercomputer (https://doi.org/10.5065/D6RX99HX), were provided by the Computational and Information Systems Laboratory (CISL) at NCAR.

REFERENCES

Ainsworth, T. D., S. F. Heron, J. C. Ortiz, P. J. Mumby, A. Grech, D. Ogawa, C. M. Eakin, and M. Leggat, 2016: Climate change disables coral bleaching protection on the Great Barrier Reef. *Science*, **352**, 338–342, https://doi.org/10.1126/science.aac7125.

Bellenger, H., E. Guilyardi, J. Leloup, M. Lengaigne, and J. Vialard, 2014: ENSO representation in climate models: from CMIP3 to CMIP5. *Climate Dyn.*, **42**, 1999–2018, https://doi.org/10.1007/s00382-013-1783-z.

Bonfils, C. J. W., B. D. Santer, T. J. Phillips, K. Marvel, L. R. Leung, C. Doutriaux, and A. Capotondi, 2015: Relative contributions of mean-state shifts and ENSO-driven variability to precipitation changes in a warming climate. *J. Climate*, **28**, 9997–10 013, https://doi.org/10.1175/JCLI-D-15-0341.1.

Cai, W., and Coauthors, 2014: Increasing frequency of extreme El Niño events due to greenhouse warming. *Nat. Climate Change*, **4**, 111–116, https://doi.org/10.1038/nclimate2100.

——, and Coauthors, 2015: Increased frequency of extreme La Niña events due to greenhouse warming. *Nat. Climate Change*, **5**, 132–137, https://doi.org/10.1038/nclimate2492.

——, G. Wang, A. Santoso, X. Lin, and L. Wu, 2017: Definition of extreme El Niño and its impact on projected increase in extreme El Niño frequency. *Geophys. Res. Lett.*, **44**, 11 184–11 190, https://doi.org/10.1002/2017GL075635.

Capotondi, A., and Coauthors, 2015: Understanding ENSO diversity. *Bull. Amer. Meteor. Soc.*, **96**, 921–938, https://doi.org/10.1175/BAMS-D-13-00117.1.

Ceppi, P., F. Brient, M. D. Zelinka, and D. L. Hartmann, 2017: Cloud feedback mechanisms and their representation in global climate models. *Wiley Interdiscip. Rev.: Climate Change*, **8**, e465, https://doi.org/10.1002/wcc.465.

Chen, C., M. A. Cane, A. T. Wittenberg, and D. Chen, 2017: ENSO in the CMIP5 simulations: Life cycles, diversity, and responses to climate change. *J. Climate*, **30**, 775–801, https://doi.org/10.1175/JCLI-D-15-0901.1.

Choi, K.-Y., G. A. Vecchi, and A. T. Wittenberg, 2015: Nonlinear zonal wind response to ENSO in the CMIP5 models: Roles of the zonal and meridional shift of the ITCZ/SPCZ and the simulated climatological precipitation. *J. Climate*, **28**, 8556–8573, https://doi.org/10.1175/JCLI-D-15-0211.1.

Chung, E.-S., A. Timmermann, B. J. Soden, K.-J. Ha, L. Shi, and V. O. John, 2019: Reconciling opposing Walker circulation trends in observations and model projections. *Nat. Climate Change*, **9**, 405–412, https://doi.org/10.1038/s41558-019-0446-4.

Collins, M., and Coauthors, 2010: The impact of global warming on the tropical Pacific Ocean and El Niño. *Nat. Geosci.*, **3**, 391–397, https://doi.org/10.1038/ngeo868.

Deser, C., L. Terray, and A. S. Phillips, 2016: Forced and internal components of winter air temperature trends over North American during the past 50 years: Mechanisms and implications. *J. Climate*, **29**, 2237–2258, https://doi.org/10.1175/JCLI-D-15-0304.1.

——, and Coauthors, 2020: Insights from Earth system model initial-condition large ensembles and future prospects. *Nat. Climate Change*, **10**, 277–286, https://doi.org/10.1038/s41558-020-0731-2

Di Lorenzo, E., and N. Mantua, 2016: Multi-year persistence of the 2014/15 North Pacific marine heatwave. *Nat. Climate Change*, **6**, 1042–1047, https://doi.org/10.1038/nclimate3082.

DiNezio, P., B. Kirtman, A. Clement, S.-K. Lee, G. Vecchi, and A. Wittenberg, 2012: Mean climate controls on the simulated response of ENSO to increasing greenhouse gases. *J. Climate*, **25**, 7399–7420, https://doi.org/10.1175/JCLI-D-11-00494.1.

Dunne, J. P., and Coauthors, 2012: GFDL's ESM2 global coupled climate–carbon Earth system models. Part I: Physical formulation and baseline simulation characteristics. *J. Climate*, **25**, 6646–6665, https://doi.org/10.1175/JCLI-D-11-00560.1.

Fasullo, J. T., B. L. Otto-Bliesner, and S. Stevenson, 2018: ENSO's changing influence on temperature, precipitation, and wildfire in a warming climate. *Geophys. Res. Lett.*, **45**, 9216–9225, https://doi.org/10.1029/2018GL079022.

Graham, F. S., A. T. Wittenberg, J. N. Brown, S. J. Marsland, and N. J. Holbrook, 2017: Understanding the double peaked El Niño in coupled GCMs. *Climate Dyn.*, **48**, 2045–2063, https://doi.org/10.1007/s00382-016-3189-1.

Guilyardi, E., P. Braconnot, F.-F. Jin, S.-T. Kim, M. Kolaskinski, T. Li, and A. Musat, 2009a: Atmosphere feedbacks during ENSO in a coupled GCM with a modified atmospheric convection scheme. *J. Climate*, **22**, 5698–5718, https://doi.org/10.1175/2009JCLI2815.1.

——, A. Wittenberg, A. Fedorov, M. Collins, C. Wang, A. Capotondi, G. J. van Oldenborgh, and T. Stockdale, 2009b: Understanding El Niño in ocean–atmosphere general circulation models: Progress and challenges. *Bull. Amer. Meteor. Soc.*, **90**, 325–340, https://doi.org/10.1175/2008BAMS2387.1.

——, W. Cai, M. Collins, A. Fedorov, F.-F. Jin, A. Kumar, D.-Z. Sun, and A. Wittenberg, 2012: New strategies for evaluating ENSO processes in climate models. *Bull. Amer. Meteor. Soc.*, **93**, 235–238, https://doi.org/10.1175/BAMS-D-11-00106.1.

He, J., N. C. Johnson, G. A. Vecchi, B. Kirtman, A. T. Wittenberg, and S. Sturm, 2018: Precipitation sensitivity to local variations in tropical sea surface temperature. *J. Climate*, **31**, 9225–9238, https://doi.org/10.1175/JCLI-D-18-0262.1.

Held, I., and B. Soden, 2006: Robust responses of the hydrological cycle to global warming. *J. Climate*, **19**, 5686–5699, https://doi.org/10.1175/JCLI3990.1.

Hoegh-Guldberg, O., 1999: Climate change, coral bleaching and the future of the world's coral reefs. *Mar. Freshwater Res.*, **50**, 839–866, https://doi.org/10.1071/MF99078.

Huang, B., and Coauthors, 2016: Further exploring and quantifying uncertainties for Extended Reconstructed Sea Surface Temperature (ERSST) version 4 (v4). *J. Climate*, **29**, 3119–3142, https://doi.org/10.1175/JCLI-D-15-0430.1.

Huang, P., 2016: Time-varying response of ENSO-induced tropical Pacific rainfall to global warming in CMIP5 models. Part I: Multimodel ensemble results. *J. Climate*, **29**, 5763–5778, https://doi.org/10.1175/JCLI-D-16-0058.1.

——, 2017: Time-varying response of ENSO-induced tropical Pacific rainfall to global warming in CMIP5 models. Part II: Intermodel uncertainty. *J. Climate*, **30**, 595–608, https://doi.org/10.1175/JCLI-D-16-0373.1.

——, and S. Xie, 2015: Mechanisms of change in ENSO-induced tropical Pacific rainfall variability in a warming climate. *Nature Geosci.*, **8**, 922–926, https://doi.org/10.1038/ngeo2571.

Johnson, N. C., and S.-P. Xie, 2010: Changes in the sea surface temperature threshold for tropical convection. *Nature Geosci.*, **3**, 842–845, https://doi.org/10.1038/ngeo1008.

Kay, J. E., and Coauthors, 2015: The Community Earth System Model (CESM) large ensemble project: A community resource for studying climate change in the presence of internal climate variability. *Bull. Amer. Meteor. Soc.*, **96**, 1333–1349, https://doi.org/10.1175/BAMS-D-13-00255.1.

Knutti, R., D. Masson, and A. Gettelman, 2013: Climate model genealogy: Generation CMIP5 and how we got there. *Geophys. Res. Lett.*, **40**, 1194–1199, https://doi.org/10.1002/grl.50256.

Power, S., F. Delage, C. Chung, G. Kociuba, and K. Keay, 2013: Robust twenty-first-century projections of El Niño and related precipitation variability. *Nature*, **502**, 541–545, https://doi.org/10.1038/nature12580.

Ramanathan, V., and W. Collins, 1991: Thermodynamic regulation of ocean warming by cirrus clouds deduced from observations of the 1987 El Niño. *Nature*, **351**, 27–32, https://doi.org/10.1038/351027a0.

Rodgers, K. B., J. Lin, and T. L. Frolicher, 2015: Emergence of multiple ocean ecosystem drivers in a large ensemble suite with an Earth system model. *Biogeosciences*, **12**, 3301–3320, https://doi.org/10.5194/bg-12-3301-2015.

Ropelewski, C. F., and M. S. Halpert, 1986: North American Precipitation and Temperature patterns associated with the El Niño/Southern Oscillation (ENSO). *Mon. Wea. Rev.*, **114**, 2352–2362, https://doi.org/10.1175/1520-0493(1986)114<2352:NAPATP>2.0.CO;2.

Seager, R., M. Cane, N. Henderson, D.-E. Lee, R. Abernathey, and H. Zhang, 2019: Strengthening tropical Pacific zonal sea surface temperature gradient consistent with rising greenhouse gases. *Nat. Climate Change*, **9**, 517–522, https://doi.org/10.1038/s41558-019-0505-x.

Stevenson, S., 2012: Significant changes to ENSO strength and impacts in the twenty-first century: Results from CMIP5. *Geophys. Res. Lett.*, **39**, L17703, https://doi.org/10.1029/2012GL052759

——, B. Fox-Kemper, M. Jochum, B. Rajagopalan, and S. Yeager, 2010: ENSO model validation using wavelet probability analysis. *J. Climate*, **23**, 5540–5547, https://doi.org/10.1175/2010JCLI3609.1.

——, ——, ——, R. Neale, C. Deser, and G. Meehl, 2012: Will there be a significant change to El Niño in the twenty-first century? *J. Climate*, **25**, 2129–2145, https://doi.org/10.1175/JCLI-D-11-00252.1.

——, J. T. Fasullo, B. L. Otto-Bliesner, R. A. Tomas, and C. Gao, 2017: Role of eruption season in reconciling model and proxy responses to tropical volcanism. *Proc. Natl. Acad. Sci. USA*, **114**, 1822–1826, https://doi.org/10.1073/pnas.1612505114.

Stuecker, M. F., and Coauthors, 2020: Strong remote control of future equatorial warming by off-equatorial forcing. *Nat. Climate Change*, **10**, 124–129, https://doi.org/10.1038/s41558-019-0667-6.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2009: A summary of the CMIP5 experimental design. CLIVAR Rep., 32 pp., www.clivar.org/organization/wgcm/references/Taylor_CMIP5.pdf.

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

Vecchi, G., and B. Soden, 2007: Global warming and the weakening of the tropical circulation. *J. Climate*, **20**, 4316–4340, https://doi.org/10.1175/JCLI4258.1.

——, and A. Wittenberg, 2010: El Niño and our future climate: Where do we stand? *Wiley Interdiscip. Rev.: Climate Change*, **1**, 260–270, https://doi.org/10.1002/wcc.33.

Watanabe, M., and A. T. Wittenberg, 2012: A method for disentangling El Niño–mean state interaction. *Geophys. Res. Lett.*, **39**, L14702, https://doi.org/10.1029/2012GL052013.

——, J.-S. Kug, F.-F. Jin, M. Collins, M. Ohba, and A. T. Wittenberg, 2012: Uncertainty in the ENSO amplitude change from the past to the future. *Geophys. Res. Lett.*, **39**, L20703, https://doi.org/10.1029/2012GL053305.

Wittenberg, A. T., 2009: Are historical records sufficient to constrain ENSO simulations? *Geophys. Res. Lett.*, **36**, L12702, https://doi.org/10.1029/2009GL038710.

——, A. Rosati, T. L. Delworth, G. A. Vecchi, and F. Zeng, 2014: ENSO modulation: Is it decadally predictable? *J. Climate*, **27**, 2667–2681, https://doi.org/10.1175/JCLI-D-13-00577.1.

Xie, S.-P., C. Deser, G. Vecchi, J. Ma, H. Teng, and A. Wittenberg, 2010: Global warming pattern formation: Sea surface temperature and rainfall. *J. Climate*, **23**, 966–986, https://doi.org/10.1175/2009JCLI3329.1.

Ying, J., and P. Huang, 2016: Cloud–radiation feedback as a leading source of uncertainty in the tropical Pacific SST warming pattern in CMIP5 models. *J. Climate*, **29**, 3867–3881, https://doi.org/10.1175/JCLI-D-15-0796.1.

Zheng, X.-T., S.-P. Xie, L.-H. Lv, and Z.-Q. Zhou, 2016: Intermodel uncertainty in ENSO amplitude change tied to Pacific Ocean warming pattern. *J. Climate*, **29**, 7265–7279, https://doi.org/10.1175/JCLI-D-16-0039.1.