



AMERICAN METEOROLOGICAL SOCIETY

Journal of Climate

EARLY ONLINE RELEASE

This is a preliminary PDF of the author-produced manuscript that has been peer-reviewed and accepted for publication. Since it is being posted so soon after acceptance, it has not yet been copyedited, formatted, or processed by AMS Publications. This preliminary version of the manuscript may be downloaded, distributed, and cited, but please be aware that there will be visual differences and possibly some content differences between this version and the final published version.

The DOI for this manuscript is doi: 10.1175/JCLI-D-15-0901.1

The final published version of this manuscript will replace the preliminary version at the above DOI once it is available.

If you would like to cite this EOR in a separate work, please use the following full citation:

Chen, C., M. Cane, A. Wittenberg, and D. Chen, 2016: ENSO in the CMIP5 simulations: lifecycles, diversity, and responses to climate change. *J. Climate*. doi:10.1175/JCLI-D-15-0901.1, in press.

© 2016 American Meteorological Society

ABSTRACT

6
7 Focusing on ENSO seasonal phase locking, diversity in peak location and propagation di-
8 rection, as well as the El Niño-La Niña asymmetry in amplitude, duration and transition, a
9 set of empirical probabilistic diagnostics (EPD) is introduced to investigate how the ENSO
10 behaviors reflected in SST may change in a warming climate.

11 EPD is first applied to estimate the natural variation of ENSO behaviors. In the obser-
12 vations El Niños and La Niñas mainly propagate westward and peak in boreal winter. El
13 Niños occur more at the eastern Pacific while La Niñas prefer the central Pacific. In a pre-
14 industrial control simulation of the GFDL CM2.1 model, the El Niño-La Niña asymmetry
15 is substantial. La Niña characteristics generally agree with observations but El Niños do
16 not, typically propagating eastward and showing no obvious seasonal phase locking. So an
17 alternative approach is using a stochastically forced simulation of a nonlinear data-driven
18 model, which exhibits reasonably realistic ENSO behaviors and natural variation ranges.

19 EPD is then applied to assess the potential changes of ENSO behaviors in the 21st century
20 using CMIP5 models. Other than the increasing SST climatology, projected changes in many
21 aspects of ENSO reflected in SST anomalies are heavily model-dependent and generally
22 within the range of natural variation. Shifts favoring eastward propagating El Niño and
23 La Niña are the most robust. Given various model biases for the 20th century and lack of
24 sufficient model agreements for the 21st century projection, whether the projected changes
25 for ENSO behaviors would actually take place remains largely uncertain.

1. Introduction

ENSO behaviors in observations and models have shown rich diversity and asymmetry. El Niños can peak at both the eastern Pacific (EP) and the central Pacific (CP) (e.g., Larkin and Harrison 2005; Ashok et al. 2007; Weng et al. 2007; Kao and Yu 2009; Kug et al. 2009; Taschetto and England 2009; Lee and McPhaden 2010; Newman et al. 2011; Takahashi et al. 2011; Karnamekas 2013; Capotondi et al. 2015; Fedorov et al. 2015; Chen et al. 2015a). Extreme El Niños propagate eastward while moderate El Niños and La Niñas tend to propagate westward (Fedorov and Philander 2001; McPhaden and Zhang 2009; Lengaigne and Vecchi 2010; Santoso et al. 2013; Kim and Cai 2014). Asymmetries between El Niño and La Niña have also been documented, e.g., El Niños often have larger amplitude than La Niñas, La Niñas are more durable than El Niños and La Niñas often tightly follow extreme El Niños but not vice versa (Kang and Kug 2002; Larkin et al. 2002; An and Jin 2004; Schopf and Burgman 2006; Ohba and Ueda 2009; Frauen and Dommenges 2010; Okumura et al. 2011; Choi et al. 2013; Dommenges et al. 2013). ENSO phase locked to the end of the calendar year has been found to be the outcome of several feedbacks and is subject to change when these competing influences change (e.g., Tziperman et al. 1995, 1997, 1998; Neelin et al. 2000; An and Wang 2001; Xiao and Mechoso 2009).

Detailed ENSO behaviors matter for ENSO teleconnection and impacts. For example, El Niños peaking at the central or eastern Pacific have been shown with varying impacts on the United States winter air temperature and precipitation (Yu et al. 2012). Detailed El Niño-La Niña transition (e.g., whether El Niños persist longer or rush to the La Niña phase), has been linked to varying likelihood of US regional tornado in the spring (Lee et al. 2016).

Usually individual ENSO behavior is investigated separately. To reach a large picture understanding, a comprehensive measure of various ENSO behaviors is required. In this study we introduce a set of empirical probabilistic diagnostics (EPD) to efficiently calculate the statistics for various ENSO behaviors, including ENSO seasonal phase locking, diversity in peak location and propagation direction, El Niño-La Niña (EN-LN) asymmetry in ampli-

53 tude, duration and transition. These diagnostics are first applied to the observed SST data
54 and the results agree with many previous understandings, which indicate this new diagnostic
55 framework is valid. One recent research focus is about ENSO in a changing climate (e.g.,
56 Collins et al. 2010; Cai et al. 2015). So after characterizing the ENSO behaviors in the past
57 ~ 150 years, we then assess how ENSO behaviors vary and change in the warming climate.

58 ENSO varies from century to century, not only in amplitude and frequency (Wittenberg
59 2009) but also in its diversity and asymmetry characteristics. For example, in the past 100-
60 year epoch, El Niños mainly peak at the eastern Pacific. In the recent decade El Niños more
61 prefer peaking at the central Pacific (Lee and McPhaden 2010). So the following questions
62 arise: whether in the warming 21st century, El Niños may switch to prefer peaking at the
63 central Pacific? If this change does happen, is it necessarily a result of the changing forcing,
64 or could it be merely a natural variation? Newman et al. (2011) and Yeh et al. (2011)
65 suggested that more occurrence of the central Pacific El Niño in the recent decade may be
66 a part of the natural variation.

67 Given only ~ 150 years of observation, how can we estimate the natural variation of each
68 ENSO behavior? All the aspects of ENSO behaviors are dynamically linked, so we apply
69 the following two approaches to ensure a coherency between the estimate of each ENSO
70 behavior. The first approach is using a long control simulation of a coupled GCM under
71 a constant forcing without a trend. Here a 4000-year pre-industrial simulation from the
72 GFDL CM2.1 coupled GCM (Delworth et al. 2006) is analyzed as one example. A second
73 approach is using a long stochastic forced simulation of a data-driven model. A 4000-year
74 simulation from an Empirical Model Reduction (EMR) (Kravtsov et al. 2005; Kondrashov
75 et al. 2005, 2015; Chen et al. 2016) is analyzed as one example. Compared to GFDL CM2.1,
76 EMR shows a slightly better overall performance for 9 aspects of ENSO behavior, therefore
77 we mainly use EMR to estimate how these ENSO behaviors may vary without a changing
78 forcing.

79 As to ENSO's response to the warming climate, it may not be merely reflected in its

80 amplitude but also in other ENSO characteristics. Here we mainly focus on SST climatology,
81 ENSO amplitude, annual cycle and 9 aspects from EPD. We analyze 37 Climate Model
82 Intercomparison Project phase 5 (CMIP5) models to investigate the following questions: If
83 the models suggest that ENSO behavior will change significantly in the 21st century (21C)?
84 Do models agree on the 21C change? Are the projected changes in ENSO behaviors more
85 response to the changing forcing or mainly part of the natural variation? Since EPD measures
86 various ENSO characteristics, we then have an opportunity to investigate which aspects are
87 most responsive to the trend forcing and which aspects vary most so that the forced changes
88 can not be easily distinguished from the range of the natural variation.

89 All estimates no matter for the natural variation or the forced change of ENSO behavior
90 have to be based on model simulations, so in this study, EPD is carried out with two
91 purposes. The first is to assess ENSO's variation and change and the second is to diagnose
92 model performance/biases on ENSO behaviors. When we analyze GFDL CM2.1 and EMR
93 models, we briefly investigate how the model biases on ENSO behavior may be related to
94 the model nonlinearity. When we assess the CMIP5 models, we investigate whether models
95 are able to represent realistic ENSO behavior in the 20C and thus reliable to project the
96 21C. We also briefly investigate how the models biases on ENSO behaviors may be related
97 to the model biases on the mean state.

98 **2. Data**

99 *a. Observation*

100 The 1870-present monthly HadISST v1.1 (Rayner et al. 2003) and 1850-present monthly
101 COBE v2 (Hirahara et al. 2014) datasets have relatively high spatial resolution ($1^\circ \times 1^\circ$) and
102 capture the diversity of ENSO behaviors. The results using COBE are overall consistent with
103 HadISST, so only the HadISST results, referred to as OBS, are shown hereafter. Tropical
104 Pacific (108E-72W, 30S-30N) SST anomalies (SSTA) are calculated by removing the monthly

105 climatology based on the commonly-used 1950-2010 period. A linear detrending is applied on
106 the SSTA at each grid point to remove the global warming trend. Then a 3-month running
107 average is applied to SSTA to smooth the temporal noise.

108 Leading modes of SST variability in the tropical Pacific region is depicted using Empirical
109 Orthogonal Function (EOF) analysis (Fig. 1). The leading EOF shows the classic El Niño
110 pattern, which is the dominant variability explaining 50% of the total variance in the tropical
111 Pacific. The second EOF shows a zonal dipole pattern with a positive loading in the western
112 Pacific and a negative loading in the eastern Pacific, which adds a central Pacific or eastern
113 Pacific “flavor” to the main El Niño pattern and explains 8% of the total variance. Following
114 the idea in the previous studies (e.g., Ashok et al. 2007; Takahashi et al. 2011), the first two
115 EOF modes and their Principal Components (PC) are used to categorize ENSO diversity in
116 the central or eastern Pacific. Takahashi et al (2011) showed that using PC1/PC2 as a basis
117 is equivalent to many other indices used to define EP/CP behavior. Details are given in the
118 method section.

119 The third EOF depicts an equatorial cooling and extra-equatorial warming that is similar
120 to the equatorial ocean dynamic thermostat pattern (Clement et al. 1996; Cane et al. 1997;
121 Solomon and Newman 2012). EOF3 explains 7% of the total variance. Its eigenvalue appears
122 not well separated from EOF2 (8%) in the observation. The sensitivity tests show that, EOF1
123 and EOF2 are robust modes in that their PCs are not as greatly influenced by the varying
124 climatology as EOF3/PC3, which has significant multidecadal variability. When we filter
125 out low frequencies (> 40 year periods) in the original data, the eigenvalue of EOF3 (5%) is
126 then well separated from EOF2 (9%).

127 *b. Coupled GCM: GFDL CM2.1*

128 Nature only provides one realization, thus the limited record of SST observations is an
129 obstacle to investigate the variation of ENSO behavior on a centurial scale. Therefore, long
130 simulations of coupled intermediate models like the Zebiak-Cane model (ZC; Zebiak and

131 Cane 1987) and fully coupled GCM with fixed external forcing are often used to investigate
132 the natural variation of ENSO (e.g., Cane et al. 1995; Wittenberg 2009; Yeh et al. 2011).

133 Here we analyze a 4000-year monthly control simulation from the GFDL CM2.1 coupled
134 GCM (Delworth et al. 2006) with the forcings, including solar irradiance, land cover and
135 atmospheric composition fixed at pre-industrial (1860) values. This simulation has been
136 analyzed in various ENSO studies (e.g., Wittenberg et al. 2006, Wittenberg 2009; Kug et al.
137 2010; Xie et al. 2010; Choi et al. 2013; Karamperidou et al. 2014; Wittenberg et al. 2014)
138 and is shown to have a reasonable ENSO performance though with a too strong amplitude
139 and too little seasonal synchronization. This simulation is referred to as “GCM” hereafter.

140 Tropical Pacific (108E-72W, 30S-30N) SSTA are calculated by removing the monthly
141 climatology based on the full length of the record. A linear detrending and a 3-month
142 running average are applied. EOF analysis is then performed. The leading three EOFs from
143 GCM explain 52%, 11% and 6% of the total variance in the tropical Pacific. GCM EOF
144 patterns are overall consistent with OBS (Fig. 1), although slightly shifted west and narrower
145 in the meridional direction as shown in Wittenberg et al. (2006).

146 *c. Data-driven modeling: EMR*

147 Since every coupled model has its own ENSO behavior that is to some extent biased
148 away from the current climate, the natural variation of ENSO estimated by an individual
149 GCM may be model-dependent. Long stochastically forced simulations from a data-driven
150 model offer an alternative approach. The model dynamics are fit from the observations thus
151 assuring that at least some of the statistics and features of the simulated ENSO resemble
152 the observed ENSO closely.

153 Here we apply the empirical model reduction (EMR) framework (e.g, Kravtsov et al.
154 2005; Kondrashov et al. 2005, 2015; Chen et al. 2016). It is a regression model with quadratic
155 nonlinearities constructed in a reduced EOF phase-space. It is fit from the observed SST
156 anomaly field and allows for ENSO nonlinearity, seasonality and memory effect for prior

157 times. Given that many ENSO behavior features are tightly linked to the nonlinearity in
158 the system (e.g., Choi et al. 2013; DiNezio and Deser 2014; Levine and Jin 2015; Chen et al.
159 2016), a nonlinear model setting is necessary. The real climate is subjected to a changing
160 forcing, so the detrended data are used to fit the model in order to produce a stationary
161 simulation. Detailed settings are given in the Appendix. A 4000-year stochastic-forced
162 EMR simulation is generated. For simplicity, this simulation is referred to as “EMR”. In
163 later sections, we will show that a EMR fit from the SSTA observation is well-behaved and
164 reproduces reasonably realistic ENSO statistics. EMR also has limitations, which will be
165 also discussed.

166 *d. Crude check on GFDL CM2.1 and EMR*

167 Before investigating the detailed ENSO behaviors, we make a crude check of ENSO
168 performance simulated by the EMR and GCM. First we check the main ENSO variability
169 represented in the tropical Pacific SST PC1 (\sim Niño-3.4). Both the EMR and GCM time
170 series appear reasonably realistic (not shown). Similar to the GFDL CM2.1 run (Wittenberg
171 2009), the long EMR time series has epochs with energetic ENSO events and epochs with
172 very weak anomalies.

173 Next we check the ENSO nonlinearity and diversity features represented in the skewed
174 probability density function (PDF) and curved 2-dimensional probability density function
175 (2dPDF) of two leading principal components (PC1-PC2). Kondrashov et al. (2005) and
176 Kravtsov et al. (2005) showed that quadratic nonlinearity is able to overall reproduce the
177 PDF and 2dPDF of a nonlinear system. Chen et al. (2016) further showed linear mod-
178 els generate an elliptic (symmetric) shape in PC1-PC2 rather than a curved (asymmetric)
179 shape. We follow the previous studies to coarsely check if these nonlinear features are re-
180 produced. Both EMR and GCM simulations show consistency with OBS in the sign of the
181 skewed distribution of PC1 and PC2, and resemble the curved 2dPDF seen in OBS (Fig.
182 2). Takahashi et al. (2011) showed that the intermediate coupled model ZC also shows this

183 curved feature in PC1-PC2.

184 Note that GFDL CM2.1 has a much stronger nonlinearity represented in a more skewed
185 distribution than OBS. Later analysis using empirical probabilistic diagnostics will show that
186 the strong nonlinearity in the GFDL CM2.1 may be one reason for the discrepancy between
187 model and OBS as to some aspects of ENSO behaviors.

188 *e. CMIP5 models*

189 Assessment of the projected climate change relies heavily on the state-of-the-art coupled
190 general circulation models (GCMs) (e.g., Capotondi et al. 2006; Guilyardi et al. 2003, 2009,
191 2012; Yu and Kim 2010; Stevenson et al. 2012; Ham and Kug 2012, 2014; Kim and Yu 2012;
192 Bellenger et al. 2013; Taschetto et al. 2014).

193 In this study, we assess the potential ENSO behavior changes in the 21st century using
194 37 CMIP5 models that participated in the Intergovernmental Panel on Climate Change
195 (IPCC) Fifth Assessment Report (AR5) (Table 1). Model descriptions and experiment
196 designs are given in Taylor et al. (2012). We analyzed three sets of simulation experiments:
197 (i) pre-industrial control simulation (PI). (ii) the historical simulations which are integrations
198 from around 1850 to at least 2005 using realistic natural and anthropogenic forcing. (iii)
199 representative concentrating pathway 8.5 (RCP8.5) simulations from the end of the historical
200 runs to 2100 when the radiative forcing reaches $8.5Wm^{-2}$. We concatenate on the historical
201 runs (1900-2005) and the RCP8.5 runs (2006-2099) and then divide them into the 20th
202 century runs (1900-1999, denoted as 20C) and the 21st century runs (2000-2099, 21C). In
203 the 21C runs, 2000-2005 part is from the historical run and the rest part is from the RCP8.5
204 run. All the PI runs are longer than 200 years. The length of the simulations vary for
205 individual models (listed in Table 1). For PI, 20C and 21C runs, monthly anomalies are
206 calculated by linearly detrending and subtracting the monthly climatology for the whole
207 period of each run.

3. Method: Empirical Probabilistic Diagnostics

The diagnostics are carried out in four steps. We first define the ENSO states. We then calculate the occurrence probability of ENSO states for each calendar month, and the transition probability between each ENSO state. After that, we derive a set of probability-based indices for ENSO seasonality (I_{season}), diversity ($I_{cp/ep}$, $I_{e/w}$) and EN-LN asymmetry (I_{amp} , I_{dur} , I_{tra}). These indices are further used to estimate the variation or change of ENSO behaviors.

a. Definition of ENSO states

We define a set of mutually exclusive ENSO states, so that we could categorize each monthly time step into one state and calculate the state transition probabilities. The states are determined using the full length of data in order to investigate ENSO behavior in shorter epochs without changing the definitions of states.

We start from the usual 3 states: El Niño (EN), Neutral (NEU) and La Niña (LN). The normalized tropical Pacific leading PC (PC1) is almost identical to Niño-3.4 with correlation $r = 0.97$, so we define the 3-state category as follows: EN is $PC1 > 0.7s.d.(PC1)$, where s.d. denotes one standard deviation. This threshold is generally consistent with the Niño 3.4 $> 0.5^{\circ}C$ criterion used by the NOAA Climate Prediction Center. Similarly, La Niña is $PC1 < -0.7s.d.(PC1)$; the remainder are defined as NEU.

When ENSO flavors are considered, the 3-state category is expanded to 5 states. In order to keep consistency with the 3-state category, we use the same PC1 threshold for El Niño/Neutral/ La Niña. Since EOF2 is a zonal dipole pattern, we use the positive or negative PC2 as the threshold to divide into EP/CP flavors (see Fig. 3). Note that in PC1-PC2 space, the upper right domain is for CPEN, while the lower left domain is for CPLN. In OBS, the normalized PC2 is highly correlated ($r = 0.88$) with the difference between normalized Niño-4 and Niño-3 indices, so the zero PC2 threshold is generally consistent with the EP/CP ENSO

233 categorization using the difference between normalized Niño-4 and Niño-3 indices (e.g., Kug
 234 et al. 2009). Typical patterns of the EP/CP flavors of El Niño and La Niña categorized
 235 in this diagnostics (Fig. 3b-e) are overall consistent with the patterns defined using Niño
 236 indices (Kao and Yu 2009; Kug et al. 2009) or C/E indices (Takahashi et al. 2011). Table
 237 1 of Yu et al. (2012) listed the major El Niño events from 1950 to 2010 with classification
 238 of EP/CP flavor based on three methods. Fig. 3f in this study shows a time series of the
 239 normalized PC1 from 1950 to 2010, with each ENSO state color-coded to individual month.
 240 The comparison indicates that our definition agrees well with the previous studies.

241 *b. Seasonal occurrence probabilities*

242 Based on the above state definitions, the annual mean occurrence probability of each
 243 ENSO state is calculated, which gives a climatology probability distribution D_{cli} . For the
 244 3-state definition, $D_{cli3} = (P_{EN}, P_{NEU}, P_{LN})$ where $P_{EN} + P_{NEU} + P_{LN} = 1$. Similarly for the
 245 5-state definition, $D_{cli5} = (P_{EPEN}, P_{CPEN}, P_{NEU}, P_{EPLN}, P_{CPLN})$ where $P_{EPEN} + P_{CPEN} +$
 246 $P_{NEU} + P_{EPLN} + P_{CPLN} = 1$.

247 We then measure the occurrence probability of ENSO state in each calendar month,
 248 which depicts the seasonal phase locking features. Full year data is used thus the sample
 249 size for each calendar month is equal. Only the 5-state result is shown in Fig. 4; 3-state
 250 results could be obtained given that $P_{EN} = P_{EPEN} + P_{CPEN}$ and $P_{LN} = P_{EPLN} + P_{CPLN}$. In
 251 OBS (Fig. 4a), El Niños and La Niñas both have higher occurrence probabilities of peaking
 252 in winter, which agrees with the observed winter phase locking diagnosed in much prior
 253 work (e.g., Tziperman et al. 1995, 1997, 1998). In winter El Niños prefer to peak at the
 254 eastern Pacific while La Niñas prefer the central Pacific, in agreement with earlier studies
 255 on the EN-LN asymmetry (Kang and Kug 2002; Schopf and Burgman 2006; Frauen and
 256 Dommenget 2010; Dommenget et al. 2013). Turning now to the two simulations, EMR
 257 generally reproduces the observed winter phase locking for El Niños (P_{EN}) and La Niñas
 258 (P_{LN}), although Fig. 4b shows the CPEN seasonality is a couple of months off from OBS.

259 In GCM (Fig. 4c), La Niñas prefer winter, which agrees with OBS. But El Niños appear
 260 to have no seasonal preference (see $P_{EPEN} + P_{CPEN}$), which is largely due to the CPENs
 261 preferring to peak in winter while EPENs rather favoring summer.

262 *c. State transition probabilities*

263 We calculate the transition probability between each ENSO state by tracking their pre-
 264 cursors and successors. Given k El Niños for example, among the precursors τ months
 265 before there are m_1 El Niños, m_2 La Niñas and m_3 Neutral states ($m_1 + m_2 + m_3 = k$).
 266 Among successors τ months later, there are n_1 El Niños, n_2 La Niñas and n_3 Neutral states
 267 ($n_1 + n_2 + n_3 = k$). The transition probability from La Niña to El Niño at a τ month inter-
 268 val is calculated as the conditional probability $P_{LN(t-\tau)|EN(t)} = m_2/k$, where t is the time.
 269 The transition probability from El Niño to La Niña at a τ month interval is calculated as
 270 $P_{LN(t+\tau)|EN(t)} = n_2/k$. The self-transition probability of El Niño across a τ month interval
 271 is calculated as $P_{EN(t+\tau)|EN(t)} = n_1/k$ or $P_{EN(t-\tau)|EN(t)} = m_1/k$, which are equal by default.

272 The 3-state transition probabilities are shown in Fig. 5. Panel a shows the transition
 273 of El Niño in OBS. On the positive lead, EN persists ($P_{EN(t+\tau)|EN(t)}$) for several months
 274 and gradually migrates to La Niña ($P_{LN(t+\tau)|EN(t)}$) or neutral ($P_{NEU(t+\tau)|EN(t)}$) in about one
 275 year. At about three years or even longer interval $\tau \rightarrow \infty$, the transition converges toward
 276 the climatology distribution. In OBS, the asymmetry in transition between El Niño and La
 277 Niña is not substantial. In EMR, the transition characteristics agree with OBS. In GCM, the
 278 EN-LN transition asymmetry is much greater than OBS with a large discrepancy between
 279 $P_{LN(t-\tau)|EN(t)}$ and $P_{LN(t+\tau)|EN(t)}$ at $\tau \sim 1-2$ year.

280 Similarly we calculate the 5-state transition probabilities. Here we mainly focus on the fa-
 281 vored zonal propagation direction of El Niño and La Niña (Fig. 6). In OBS, $P_{EPEN(t-\tau)|CPEN(t)} >$
 282 $P_{EPEN(t+\tau)|CPEN(t)}$ across 0-6 months (Fig. 6a) indicates El Niños favor westward propaga-
 283 tion from EPEN to CPEN. $P_{CPLN(t-\tau)|EPLN(t)} < P_{CPLN(t+\tau)|EPLN(t)}$ across 0-6 months (Fig.
 284 6b) indicates La Niñas also favor westward propagation from EPLN to CPLN. In EMR, the

285 zonal transition of El Niños and La Niñas both agree with OBS (Fig. 6c and 6d). In GCM,
 286 La Niñas favor the westward propagation, similar to OBS (Fig. 6f). But El Niños differ
 287 from OBS by favoring the eastward propagation (Fig. 6e).

288 Probability measures show that individual ENSO behaviors have varying representative
 289 time scales. So we later define the indices based on the transition probabilities at their ac-
 290 cording time scales. In 3-state transition, we mainly focus on the persistence (self-transition)
 291 of El Niño and La Niña within 0-36 months and asymmetry in the EN-LN transition across
 292 0-18 months. In 5-state transition, we mainly focus on the zonal propagation of El Niño
 293 (asymmetry in EPEN-CPEN transition) and the zonal propagation of La Niña (asymmetry
 294 in EPLN-CPLN transition) across 0-6 months. Other transition probability results, e.g.,
 295 transition to/from NEU, self-transition of EP/CP ENSO states, and EN-LN transitions
 296 specifying EP/CP information like $P_{EPEN(t-\tau)|CPLN(t)}$, are not included in the following set
 297 of ENSO behavior indices but left for future study.

298 Here we use both forward and reverse conditional probability to characterize the transi-
 299 tion, which is based on the following consideration. In a stationary process, transition proba-
 300 bility across a long time interval converges to the climatology. Note that $\lim_{\tau \rightarrow \infty} P_{LN(t+\tau)|EN(t)} =$
 301 P_{LN} , while $\lim_{\tau \rightarrow \infty} P_{EN(t+\tau)|LN(t)} = P_{EN}$. To characterize the transition asymmetry, these two
 302 conditional probabilities need to be both normalized by their according occurrence proba-
 303 bility or one needs to be rescaled to match the other one. Here we use El Niño as the base
 304 state, the asymmetry is then calculated using the discrepancy $dP_{EN-LN} = P_{LN(t+\tau)|EN(t)} -$
 305 $P_{EN(t+\tau)|LN(t)}/P_{EN} \times P_{LN}$. Given Bayes' rule, we obtain $dP_{EN-LN} = P_{LN(t+\tau)|EN(t)} -$
 306 $(P_{LN(t)|EN(t+\tau)} \times P_{EN}/P_{LN})/P_{EN} \times P_{LN} = P_{LN(t+\tau)|EN(t)} - P_{LN(t-\tau)|EN(t)}$. So both the forward
 307 conditional probability $P_{LN(t+\tau)|EN(t)}$ and the reverse conditional probability $P_{LN(t-\tau)|EN(t)}$
 308 are shown in Fig. 5 to illustrate the transition asymmetry. If using La Niña as the base
 309 state instead, $dP_{EN-LN} = (P_{EN(t-\tau)|LN(t)} - P_{EN(t+\tau)|LN(t)})/P_{EN} \times P_{LN}$ gives the same result.
 310 Similarly, we characterize the zonal transition of El Niño based on CPEN and the discrep-
 311 ancy $dP_{CPEN-EPEN} = (P_{EPEN(t+\tau)|CPEN(t)} - P_{EPEN(t-\tau)|CPEN(t)})$. It could be also calcu-

312 lated based on EPEN, $dP_{CPEN-EPEN} = (P_{CPEN(t-\tau)|EPEN(t)} - P_{CPEN(t+\tau)|EPEN(t)})/P_{CPEN} \times$
313 P_{EPEN} . For the zonal transition of La Niña, we calculate the discrepancy $dP_{CPLN-EPLN} =$
314 $P_{CPLN(t-\tau)|EPLN(t)} - P_{CPLN(t+\tau)|EPLN(t)}$ based on EPLN. If based on CPLN instead, then
315 $dP_{CPLN-EPLN} = (P_{EPLN(t+\tau)|CPLN(t)} - P_{EPLN(t-\tau)|CPLN(t)})/P_{EPLN} \times P_{CPLN}$ gives the same
316 result.

317 *d. A set of indices for ENSO behaviors*

318 Here we define a set of indices to measure various aspects of ENSO behavior in a period
319 of 100 year. OBS is divided into 5 100-year overlapping epochs starting 10-years apart. The
320 4000-year simulations of EMR and GCM are divided into 391 overlapping 100-year epochs
321 also starting 10-years apart. We first summarize the ENSO behaviors in OBS and then
322 discuss EMR and GCM.

323 A seasonality index I_{sea} is defined to identify the favored peak season for a given epoch.
324 For El Niños, I_{sea} is the total occurrence of El Niño in the summer half year (March-August)
325 divided by the total occurrence of El Niño in the winter half year (Sep.-Feb.). $I_{sea} < 1$
326 (> 1) indicates El Niño preferentially peaks in winter (summer). The results (Fig. 7a and
327 7d) show that both El Niños and La Niñas in OBS prefer winter. In EMR, El Niños and
328 La Niñas also prefer the winter half year in all epochs. In GCM, La Niñas mainly prefer
329 winter, consistent with OBS and EMR, but for El Niños, I_{sea} is more centered around one,
330 suggesting majority of epochs do not have notable seasonal preference. This agrees with the
331 seasonal occurrence probability results in Fig. 4.

332 A diversity index $I_{cp/ep}$ is defined for El Niños to diagnose the dominant peak location in
333 a given epoch, calculated as the total occurrence of CPEN divided by the total occurrence
334 of EPEN. $I_{cp/ep} < 1$ (> 1) indicates El Niño prefers to peak at EP (CP). There is a similar
335 definition for La Niña. The results (Fig. 7b and 7e) show that, EPENs and CPLNs dominate
336 in OBS. In EMR, 91% of epochs are dominated by EPEN and 97% of epochs are dominated
337 by CPLN. The GCM overall agrees with the OBS and EMR, with most epochs favoring

338 EPENs and CPLNs.

339 Another diversity index $I_{e/w}$ is defined to diagnose the dominant zonal propagation in a
340 given epoch. For El Niños, $I_{e/w}$ is the average of $dP_{CPEN-EPEN}$ within a 6-month interval.
341 For La Niñas, $I_{e/w}$ is the average of $dP_{CPLN-EPLN}$ within a 6-month interval. $I_{e/w} < 0$
342 (> 0) indicates preferring westward (eastward) propagation. Figs. 7c and 7f show that
343 OBS has more westward moving El Niños and La Niñas. In EMR, El Niños favor westward
344 propagation in 74% of epochs while La Niñas favor westward propagation in all epochs. In
345 GCM, more than half of epochs favoring westward propagating La Niñas, which is generally
346 consistent with OBS and EMR. But all epochs favor eastward propagating El Niños, which
347 is not realistic. This model discrepancy agrees with the 5-state transition results (Fig. 6).

348 The asymmetry index I_{amp} , which diagnoses the relative amplitude of El Niño and La
349 Niña in a given epoch, is calculated as the mean of PC1 value for El Niño months in 100
350 years divided by the mean of PC1 value for La Niña months. $I_{amp} > 1$ indicates that the
351 overall amplitude of El Niño is larger than La Niña. Fig. 7g shows that El Niños have larger
352 amplitude than La Niñas in OBS. In EMR, 82% of epochs have $I_{amp} > 1$. GCM generally
353 agrees with OBS and EMR.

354 Another asymmetry index I_{dur} diagnoses the relative duration of El Niño and La Niña in a
355 given epoch. It is calculated as the mean of El Niño self-transition probabilities $P_{EN(t+\tau)|EN(t)}$
356 within a 36-month interval divided by the mean of La Niña self-transition probabilities
357 $P_{LN(t+\tau)|LN(t)}$ within a 36-month interval. $I_{dur} < 1$ indicates La Niña is more durable. Fig.
358 7h shows that La Niña is more durable in OBS. In EMR 61% of epochs have more durable
359 La Niñas. GCM agrees well with EMR and OBS.

360 A third asymmetry index I_{tra} , which diagnoses the transition asymmetry between El
361 Niño and La Niña, is calculated as average of dP_{EN-LN} within a 18-month interval. $I_{tra} > 0$
362 indicates La Niñas tightly follow El Niños more than vice versa. Fig. 7i shows that in OBS
363 the asymmetry in transition is not substantial. For the 20th century epoch, it has a slightly
364 positive asymmetry with $I_{tra} > 0$ (shown in Fig. 12d along with CMIP5 models). In EMR,

365 55% of epochs have $I_{tra} > 0$. Compared to OBS and EMR, GCM shows a much larger
366 asymmetry in transition, with all epoch favoring La Niña tightly following El Niño. This
367 discrepancy agrees with the 3-state transition in Fig.5. Our results is generally consistent
368 with Choi et al. (2013), in which the EN-LN transition is defined based on individual events
369 with the time range set to 12 months.

370 4. Natural variation of ENSO behaviors

371 In this section, we first summarize the performance of ENSO behaviors in two simulations.
372 GFDL CM2.1 serves as one example for the fully coupled GCMs and EMR is one example
373 for the data-driven models. We then discuss how the simulated ENSO behaviors depend on
374 the model nonlinearity. We last overview the natural variation of the ENSO behaviors.

375 a. GFDL CM2.1

376 Fig. 4-7 show that, GFDL CM2.1 overall agrees with OBS and EMR as to some aspects of
377 the ENSO behavior, e.g., the ratio of CP/EP ENSO (Fig. 7b,e), El Niño-La Niña asymmetry
378 in amplitude (Fig. 7g) and duration (Fig. 7h). It is mainly biased from the OBS in three
379 aspects, including the seasonality (Fig. 4; Fig. 7a, d), EN-LN transition (Fig. 5; Fig. 7i)
380 and the zonal propagation (Fig. 6; Fig. 7c,f).

381 We then briefly discuss these three biases. As to the seasonal phase locking, El Niños in
382 GFDL CM2.1 do not show notable seasonal preference (Fig. 4), which is mainly due to the
383 competing impacts of EPEN peaking in summer while CPEN peaking in winter. Overall,
384 SST anomaly peaks when the collective positive feedbacks are balanced by the negative
385 feedbacks (e.g., Tziperman et al. 1995, 1997, 1998; Neelin et al. 2000; An and Wang 2001;
386 Xiao and Mechoso 2009; Stein et al. 2010). In GFDL CM2.1, the biases in positive and
387 negative feedbacks may together alter the SSTA peak timing and location. Wittenberg et
388 al. (2006) once showed that GFDL CM2.1 simulated events tend to peak either in summer

389 or winter. This bias is likely tied to the semiannual cycle of the background convection and
390 currents, which is associated with double ITCZ and the seasonal reversal of the meridional
391 SST gradient and winds in the eastern Pacific.

392 As to the EN-LN transition and the zonal propagation, all epochs of GFDL CM2.1 show
393 a strong transition asymmetry with La Niñas tightly following El Niños (Fig. 7i) and all
394 epochs favor eastward propagating El Niños (Fig. 7c). GFDL CM2.1 is largely biased so
395 that its variation range does not even cover the observation. If only centennial-long rather
396 than 4000-yr GFDL CM2.1 simulations are available, the biases in these two aspects might
397 be the most distinguishable from OBS.

398 In this study we do not have a special category for the extreme El Niños since only a
399 few extreme El Niños occurred in OBS. The overall ENSO statistics in the observations are
400 dominated by the moderate events. GFDL CM2.1 has overly strong El Niño with many
401 extreme events (Wittenberg 2009; Takahashi and Dewitte 2015; Levine and Jin 2015). Ex-
402 treme El Niños usually peak in the eastern Pacific, while moderate El Niños peak either in
403 the central Pacific or in the eastern Pacific (Kug et al. 2009; Takahashi and Dewitte 2015).
404 Extreme El Niños tend to propagate eastward and moderate El Niños propagate westward
405 (Santoso et al. 2013; Kim and Cai 2014). Extreme El Niños are also associated with a
406 large EN-LN asymmetry in transition (Choi et al. 2013). Thus the statistics in GFDL
407 CM2.1 is largely shifted by extreme events to favor more eastward propagating El Niños and
408 much strong asymmetry in EN-LN transition. These behavior biases could be further traced
409 back to model’s overly strong nonlinearity, which is also manifested in, e.g., largely skewed
410 probability density function of SSTA in Fig. 2.

411 We also notice that GFDL CM2.1 does better with SSTAs associated with La Niña than
412 El Niño, especially for seasonality and zonal propagation. Since the coupling only becomes
413 nonlinear above a certain temperature threshold (Takahashi and Dewitte 2015; Levine and
414 Jin 2015), El Niños (as a warming anomaly on top of the mean temperature) are more
415 sensitive to the extent of nonlinearity in the system than La Niña (a cooling anomaly).

416 This may be one reason why La Niñas do not show as greater a diversity as El Niños in
417 Kug and Ham (2011). A good performance on El Niño demands that the strength of the
418 model’s nonlinearity resemble that in the real climate. On the contrary, La Niñas may still
419 be simulated realistically even for a model with a too strong nonlinearity.

420 *b. EMR*

421 Fig. 4-7 show that, EMR overall agrees with OBS as to most aspects of ENSO behaviors,
422 e.g., seasonal phase locking (Fig. 4; Fig. 7a, d), the ratio of CP/EP ENSO (Fig. 7b,e), El
423 Niño-La Niña asymmetry in amplitude (Fig. 7g), duration (Fig. 7h) and transition (Fig. 5;
424 Fig. 7i). It shows slight biases for zonal propagation (Fig. 6; Fig. 7c,f).

425 The EMR is built to capture the transition from one month to the next, which includes
426 some nonlinear dynamics, memory effects from a single prior time step, and annual periodic
427 terms. On one hand, the results that EMR overall agrees with OBS as to the nonlinear
428 measure 2dPDF in Fig. 2h and for the seasonal phase locking in Fig. 4 are expected. On the
429 other hand, EMR does not explicitly build in different peak locations, different propagation
430 directions, or the EN-LN asymmetry in amplitude, duration and transition. Its ability to
431 capture these aspects is an implicit and non-obvious consequence of the model construction.
432 Its extended behaviors in the long runs are at least as plausible as the GCM. Moreover, as
433 a low-order empirical model, an EMR simulation is computationally efficient.

434 *c. ENSO behavior dependence on nonlinearity*

435 GFDL CM2.1’s biases on ENSO behavior suggest that simulated ENSO behavior is very
436 sensitive to the model nonlinearity, especially the asymmetry between El Niño and La Niña.
437 In this section, we fit a linear EMR model (EMR-L) to the observation (details in Appendix)
438 and generate a 4000-year simulation. In total we compare the ENSO behaviors in four
439 systems with varying levels of nonlinearity: EMR-L, OBS, EMR, and GFDL CM2.1. The

440 results are shown in Fig. 8.

441 We first discuss the ENSO behavior simulated in the linear system, and then compare
442 with other systems. In the linear model simulation, 2dPDF in PC1-PC2 space is centered
443 at zero without a curved shape (Fig. 8b), which is as expected. There is also no notable
444 EN-LN asymmetry in amplitude (Fig. 8e), duration (Fig. 8f) and transition (Fig. 8g). Both
445 El Niño and La Niña prefer winter peaking (Fig. 8h). There is no EP/CP preference for
446 peaking location (Fig. 8i). Both El Niño and La Niña favor west propagating (Fig. 8j). The
447 distribution of El Niño and La Niña are almost identical. Overall, ENSO behaviors in the
448 linear system lack of the EN-LN asymmetries, which indicates the nonlinearity is necessary
449 to create the EN-LN asymmetry.

450 Using the linear EMR as a reference, we then discuss the ENSO behaviors in the nonlinear
451 EMR. It reproduces the curved 2dPDF in PC1-PC2. An EN-LN asymmetry appears in
452 amplitude and duration, though it is still not notable for transition. As to the peak season,
453 El Niño and La Niña starts to show a small discrepancy (Fig. 8k). As to the peak location,
454 El Niño prefers the eastern Pacific and La Niña prefers the central Pacific (Fig. 8l), which
455 asymmetry agrees with the observation. As to the zonal propagation, El Niños start to shift
456 a bit toward favoring eastward while La Niñas shifting toward favoring westward (Fig. 8m).

457 Then we discuss the ENSO behaviors in the strong nonlinear GFDL CM2.1. It reproduces
458 the 2dPDF with a larger curvature. It shows a larger EN-LN asymmetry in amplitude,
459 duration and transition. As to the peak season, El Niño and La Niña show a larger difference
460 (Fig. 8n). Distribution of El Niño shifts toward favoring summer peaking. As to the peak
461 location, El Niño prefers the eastern Pacific and La Niña prefers the central Pacific (Fig.
462 8o). As to the zonal propagation, La Niñas still favor westward propagation while El Niños
463 favor eastward propagation.

464 We previously discussed that GFDL CM2.1 shows a better performance for La Niña com-
465 pared to El Niño. Here the comparison between linear and nonlinear system also suggests
466 that, when the system become more nonlinear, La Niñas better preserve its usual character-

467 istics, while El Niños are very sensitive so the characteristics may change quickly and exhibit
468 a larger diversity.

469 Model nonlinearity is influenced by many different physical processes. DiNezio and Deser
470 (2014) addressed the nonlinear controls on the persistence of La Niña using 1300-year sim-
471 ulation of the CCSM4 model. The authors fit a nonlinear delayed oscillation model which
472 illustrates the nonlinearity in the delayed thermocline feedback plays a role for the persistence
473 of La Niña. Takahashi and Dewitte (2015) have shown that, moderate and strong nonlinear
474 ENSO regimes exist. They found that extreme El Niño events simulated by GFDL CM2.1
475 has consistent temporal evolutions as the observed strong El Niños. So a GFDL CM2.1
476 simulation is analyzed to reconstruct a robust evolution profile for SST, wind stress and
477 the thermocline tilting. The authors also showed that the existence of these regimes is very
478 likely due to the nonlinearity in the Bjerknes feedback.

479 In summary, comparison between EMR-L, EMR and GFDL CM2.1 shows that, a non-
480 linear model is necessary to reproduce the comprehensive ENSO behaviors including the
481 EN-LN asymmetry. For a given nonlinear model, it is also important that this model has a
482 proper extent of model nonlinearity that resembles the reality. The results also show that
483 the nonlinearity mainly influences the ENSO behavior by controlling the extent of the EN-
484 LN asymmetry. El Niños are more sensitive to the system nonlinearity and exhibit a larger
485 diversity than the La Niñas.

486 *d. Overview of ENSO behavior variations*

487 GFDL CM2.1 may provide an reasonable estimate for variation of EP/CP ENSO flavors,
488 but may not be suitable for the zonal propagation and the EN-LN transition. EMR is
489 overall better for 9 aspects. Note that EMR's nonlinearity may be also slightly larger than
490 OBS, which might be further adjusted in the future study. Before a better model becomes
491 available, the current EMR model is still useful to provide a relatively realistic estimate for
492 ENSO behaviors.

493 We now summarize the natural variation of ENSO behavior given by EMR. Under the
494 white noise forcing without a trend, epochs with ENSO peaking in summer and epochs
495 favoring eastward propagating La Niñas are still not likely, but the following scenarios may
496 occur with a certain likelihood, e.g., epochs with El Niño preferring central Pacific (9%),
497 epochs with La Niña favoring eastern Pacific (3%), epochs preferring eastward propagating
498 El Niños (26%), epochs with La Niña having larger amplitude than El Niño (18%), epochs
499 with more persist El Niño than La Niña (39%) and epochs with a quicker transition from La
500 Niña to El Niño (45%). Under stochastic noise, epochs with characteristics different from
501 the past 100-year OBS could occur. One needs to be cautious when attributing the unusual
502 ENSO variations for a certain period as a response to the changing external forcing.

503 *e. Uncertainty in estimates of the true distribution*

504 We have calculated a distribution of indices from overlapping 100-year epochs of a long
505 (4000-year) model simulation. Note that model-reconstructed distribution is not constrained
506 to be centered at the OBS samples. Usually such a match is built-in by resampling the
507 data or sampling from an assumed distribution (e.g. Gaussian) with specified sample mean
508 and standard deviation. Both the GCM and EMR distributions are based on 391 samples
509 while the OBS distribution is based on only 5 samples. The OBS sample distribution, as
510 an approximation to the unknown true distribution, is used as the metric to examine the
511 modeled distributions. Given only a few available OBS samples, its distribution may not
512 reflect the statistics of the true distribution (Wittenberg, 2009). The shortness of the OBS
513 record also makes it difficult to establish that a model distribution is significantly different
514 from OBS.

515 We apply the Kolmogorov-Smirnov (KS) test to estimate whether or not two distributions
516 are alike. KS tests show that among the 9 aspects in Fig. 7, the EMR and GCM distributions
517 are significantly different for 8 of them with $p < 10^{-3}$, while for I-dur $p = 0.02$. The GCM
518 distributions of I-cp/ep-LN, I-amp, I-dur and I-season-LN are not significantly different

519 from OBS distribution while the GCM distributions of I-season-EN, I-cp/ep-EN, I-e/w-EN,
520 I-e/w-LN and I-tra differ from OBS at greater than the 1% level ($p < 0.01$). With the same
521 significance level, EMR distribution differs from the OBS distributions for I-season-EN, I-
522 season-LN, I-e/w-EN and I-e/w-LN.

523 Besides the uncertainty coming from the shortage of data to fit and constrain robust
524 model coefficients for the data-driven models, another aspect of uncertainty comes from the
525 model construction itself. We showed in Chen et al. (2016) that there are many different ways
526 to formulate an EMR or other low order models. It is difficult to construct one particular
527 model that captures every conceivable feature. Though the current version of EMR including
528 nonlinearity, memory effect and seasonality is the overall best choice in the study, it fails to
529 match OBS's distribution closely for a few aspects. We use EMR as an example to illustrate
530 both using a data-driven model and a GCM to estimate the natural variation of the ENSO
531 characteristics. Both models are flawed and the bias correction is a potential topic for future
532 study.

533 **5. ENSO behavior change in the 21st century**

534 In this section, we analyze whether the ENSO behaviors may exhibit notable changes from
535 20C to 21C using CMIP5 projections under the RCP8.5 scenario. Besides the aforementioned
536 ENSO behaviors, we also estimate the annual mean SST, annual cycle, the standard deviation
537 and skewness of SSTA in Niño-3.4 region.

538 When we analyze the CMIP5 models, we notice that the PC-based definition for ENSO
539 diversity works well for the models with correct representations of EOF1 and EOF2, but
540 this definition may not be optimal for the models with poor performance on ENSO diversity.
541 So we will use a similar but Niño-based definition for CMIP5 model evaluation. Normalized
542 Niño-3.4 replaces normalized PC1. Normalized Niño-4 minus normalized Niño-3 replaces
543 normalized PC2. All are normalized to have standard deviation equal to one.

544 We will first evaluate whether 37 CMIP5 models could reproduce realistic statistics (as
545 compared with the observation in the 20th century). The results from each model are sorted
546 in an ascending order for individual aspects of ENSO behavior. Table 1 gives the model
547 rank for each aspect. A multi-model mean (MMM) for 37 models is calculated to represent
548 an overall performance of CMIP5 models. After that we compare the number of models
549 projecting an increase or a decrease to assess whether the projected changes for the 21st
550 century are supported with sufficient model agreement.

551 We next identify if an apparent change is a significant response to the changing external
552 forcing. For each individual model, the change from 20C to 21C is viewed as a significant
553 change if satisfying one of the following: 1) $I_{21C} > I_{20C}$ (increase) and $I_{21C} > 97.5\text{th percentile}$
554 of PI run, 2) $I_{21C} < I_{20C}$ (decrease) and $I_{21C} < 2.5\text{th percentile}$ of PI run, where $I_{20C}(I_{21C})$
555 is the given index calculated for 20C(21C).

556 Note that there is a large spread of the natural variation given by each model's PI
557 runs, so an additional estimate of the natural variation is provided as reference using the
558 aforementioned 4000-year stochastic-forced EMR simulation fit from OBS.

559 *a. SST climatology and anomaly*

560 We first investigate whether these CMIP5 models reproduce a realistic tropical Pacific
561 climatology. The time series of Niño-3.4 SST from 1900 to 2100 are shown in Fig. 9a. Niño-
562 3.4 SST averaged in the 20th century (20C) and 21st century (21C) are shown in Fig. 9b.
563 The averaged Niño-3.4 SST in sliding 100-year epochs for the pre-industrial control run (PI)
564 is also provided as reference. There is a considerable spread compared to observations in
565 20th century runs, and the MMM of 37 models slightly underestimates the 20C mean SST.
566 All 37 models project a warming future for the RCP8.5 scenario with a $\sim 2^\circ\text{C}$ temperature
567 increase. The changes are significant in all models.

568 The Niño-3.4 SST anomalies in 20C and 21C are obtained by linearly detrending and
569 removing monthly climatology in each 100-year segment (Fig. 9c) and the standard deviation

570 (s.d.) of Niño-3.4 are shown in Fig. 9d. The natural variation range is provided by PI control
571 runs of each model and the 4000-year stochastic-forced EMR simulation fit from OBS. The
572 MMM for 20C overestimates the ENSO amplitude. There is large spread ranging from half
573 the amplitude of OBS to nearly twice, consistent with the findings in Bellenger et al. (2013).
574 For the 21C, MMM of 37 models projects an increase in Niño-3.4 s.d.. 20/17 models show
575 a decrease/increase, among which 6/13 models are significant.

576 *b. SST annual cycle and seasonality*

577 The annual cycle of 20C Niño-3.4 SST of all CMIP5 models are presented in Fig. 10a.
578 The structure of the annual cycle is measured using an index I_{seacli} defined as the averaged
579 SST during winter half year (Sept.-Feb.) minus that during summer half year (Mar.-Aug.).
580 Most models produce a reasonable annual cycle but some models show a semiannual cycle.
581 MMM for 20C indicates a weaker annual cycle than observation. For 21C, MMM does not
582 project an apparent change. 18/19 models project a decrease/increase, among which 15/10
583 models are significant.

584 The seasonal phase locking for El Niño and La Niña are shown using occurrence prob-
585 ability for each calendar month (Fig. 10c and 10e). Most models produce a winter phase
586 locking as in 20C OBS, consistent with Taschetto et al. (2014). But some models show no
587 preferred peak season or peak in summer. Previous studies (e.g., Guilyardi et al. 2003; Ham
588 and Kug 2014) suggested that the biased models in seasonality tend to also have biases in
589 climatology and oceanic mean state. The seasonal locking for El Niño and La Niña measured
590 by I_{sea} are shown in Fig. 10d and 10f. Though there is large spread among models, MMM
591 for 20C appears to overall match the observation. For 21C, MMM does not project much
592 change for El Niño and only a slight change for La Niña. For El Niños, 21/16 models project
593 a decrease/increase, among which 6/5 models are significant. For La Ninas, 22/15 models
594 project a decrease/increase, among which 15/2 models are significant.

595 *c. ENSO diversity in peak location*

596 For 20C simulation, 37 models show a large spread for the ratio between CP/EP ENSO
597 $I_{cp/ep}$ (Fig. 11a and 11c). More than half of the models resemble the observation that El
598 Niños favor the eastern Pacific though MMM slightly overestimates the value of $I_{cp/ep}$. More
599 than half of the models resemble the observation to favor La Niñas peaking at the central
600 Pacific though MMM slightly underestimates the value of $I_{cp/ep}$.

601 For 21C projection, 20/17 models project a decrease/increase for El Niños, among which
602 1/2 models are significant. For La Ninas, 15/22 models project a decrease/increase, among
603 which 4/5 models are significant. MMM results show no notable change. Yeh et al. (2009)
604 analyzed 12 CMIP3 models in which an increased frequency of CPEN compared to EPEN
605 is suggested to be related to the flattening of the thermocline in the equatorial Pacific. Kim
606 and Yu (2012) analyzed 16 CMIP5 models under the RCP4.5 scenario which suggest an
607 increased ratio of CP to EP El Niño. Taschetto et al. (2014) used 27 CMIP5 models under
608 the RCP8.5 scenario which suggest no notable enhancement of the ratio of CP/EP ENSO.
609 Here we analyzed 37 CMIP5 models in RCP8.5 scenario and measure the projected change
610 from 20C to 21C using the probability shift of relative occurrence ($I_{cp/ep}$). The discrepancy
611 among above studies suggests the projections heavily depend on the selected models and the
612 uncertainty is large given varying performance of these models for the historical period.

613 *d. ENSO diversity in propagation direction*

614 For 20C simulation, more than half of the models favor westward propagating El Niños as
615 the observation though MMM overestimates the value of $I_{e/w}$ (Fig. 11b). Almost all models
616 favor westward propagating La Niñas and MMM overall matches with the observations (Fig.
617 11d). For El Niños, 13/24 models project a decrease/increase, among which 2/14 models
618 are significant. For La Ninas, 12/25 models project a decrease/increase, among which 6/15
619 models are significant. As suggested by Santoso et al. (2013), the projected weakening of the

620 westward mean equatorial currents may explain the projected shift toward more eastward
621 propagating El Niños and La Niñas in a warmer world.

622 *e. ENSO asymmetry in amplitude, duration and transition*

623 The skewness of Niño-3.4 SST anomaly (Fig. 12a) coarsely measures the El Niño-La Niña
624 asymmetry in amplitude, so it is discussed here together along with I_{amp} . For 20C, more than
625 half of the 37 models show a positive sign of skewness in agreement with OBS, though MMM
626 underestimates the value of skewness. For 21C, MMM shows a slight decrease of skewness.
627 24/13 models project a decrease/increase, among which 8/2 models are significant.

628 The amplitude asymmetry I_{amp} (Fig. 12b) shows that more than half the models agree
629 with OBS in 20C with larger amplitude in El Niño than in La Niña, though MMM un-
630 derestimate the asymmetry. For 21C, MMM projects a decrease in the asymmetry. 21/16
631 models project a decrease/increase, among which 7/3 models are significant. The duration
632 asymmetry I_{dur} (Fig. 11c) shows that only half the models agree with 20C OBS in showing
633 a more persistent La Niña. MMM also underestimates the asymmetry. For 21C, MMM
634 does not show much change. 16/21 models project a decrease/increase, among which 5/2
635 models are significant. As to the transition I_{tra} , most models show much larger transition
636 asymmetry than OBS in 20C and MMM also overestimate the asymmetry. For 21C, MMM
637 does not show much change. 19/18 models project a decrease/increase, among which 3/5
638 models are significant. For all three aspects of EN-LN asymmetry, the projected changes are
639 largely within the natural variation range based upon the control run.

640 Previous studies (e.g., Zhang and Sun 2014; Cai et al. 2015a) pointed out most CMIP5
641 models underestimate the ENSO asymmetry. Here the results in our diagnostics show that,
642 among these three aspects, the EN-LN asymmetry in amplitude and duration are underes-
643 timated by most models, and the EN-LN asymmetry in transition is usually overestimated.

6. Discussion

a. Which aspect may show a robust change in 21C?

Now we would like to summarize, among all the ENSO behaviors we have analyzed in this study, which aspect may show a robust change in 21C. Increasing SST climatology, as the result of changing external forcing, is consistently predicted by all models. Other than that, changes in many SSTA aspects are not that robust, and the difference between 20C and 21C are largely within the range of the natural variation.

We first estimate which aspects are most responsive to the trend forcing, in which case the changes could be readily detected from the natural variation in the model projection. For each aspect, we count the number of models (denoted as N_c) out of 37 models showing a significant change (no matter positive or negative changes). The results are as follows: SST climatology ($N_c = 37$), annual cycle I_{seacli} ($N_c = 25$), standard deviation of Niño-3.4 ($N_c = 19$), diversity in zonal propagation for El Niño $I_{e/w-EN}$ ($N_c = 16$) and La Niña $I_{e/w-LN}$ ($N_c = 21$), seasonality of El Niño I_{sea-EN} ($N_c = 11$) and La Niña I_{sea-LN} ($N_c = 12$), the EN-LN asymmetry in I_{amp} ($N_c = 10$), I_{dur} ($N_c = 7$), I_{tra} ($N_c = 8$), diversity in peak location for El Niño $I_{cp/ep-EN}$ ($N_c = 3$) and La Niña $I_{cp/ep-LN}$ ($N_c = 9$).

Among these 12 aspects, SST climatology gives $N_c = 37$, suggesting that SST mean state quickly adjusts to the increasing greenhouse gas emissions so that all 37 models can readily capture this change. However, diversity in peak location $I_{cp/ep}$ only gives $N_c = 3$ for El Niño and $N_c = 9$ for La Niña. It suggests that El Niño peaking at eastern Pacific or central Pacific often varies even under a constant forcing. Therefore the large range of natural variation makes it difficult to detect the forced response.

Although the change of the annual cycle is the very responsive ($N_c = 25$), 37 models still do not give a clear direction of change (18/19 models show a decrease/increase, among which 15/10 models show a significant decrease/increase). Therefore, it is difficult to argue whether the annual cycle would become weaker in the 21C.

670 The change of the standard deviation of Niño-3.4 is also very responsive ($N_c = 19$).
671 20/17 models project a decrease/increase, among which 6/13 models project a significant
672 decrease/increase. Chen et al. (2015b) investigated the physical mechanisms for four individ-
673 ual models showing either increasing or decreasing ENSO amplitude in the warming climate.
674 The authors found that models with a stronger (weaker) equatorial thermocline response to
675 the zonal wind anomaly tend to project a strengthened (weakened) ENSO amplitude.

676 A shift toward eastward propagation of El Niños and La Niñas are the responsive aspect
677 with the most robust change supported by MMM and nearly 2/3rd of the models. Santoso
678 et al. (2013) have shown that the westward mean current is one main reason for ENSO's
679 westward propagation. Therefore a shift toward favoring the eastward propagating El Niño
680 may be the response to the weakened westward mean current as projected for 21C.

681 By basing our metrics on SSTA we follow the practice in the vast majority of prior ENSO
682 literature. However, based on observations (Karl et al. 1995), CMIP modeling results and
683 theory (Allen and Ingram 2002), it is expected that with global warming the atmospheric
684 water vapor content will increase along with the intensifying convective events. Thus we
685 expect that the rainfall associated with ENSO events will increase, as has been found by
686 Power et al. (2013), Cai et al (2014) and Cai et al. (2015b).

687 *b. Projection using a subset of models*

688 Studies on CMIP models often identify subsets of good models based on various metrics
689 (e.g., Gleckler et al. 2008; Kim and Yu 2012; Bellenger et al. 2013). In this study, we
690 would like to estimate if a subset of good models may give a more reliable projection for 21C
691 with a better model agreement. Here we identify good models based on 13 aspects of ENSO
692 behavior. Table 1 summarizes 37 CMIP5 model performances based on these 13 metrics.
693 For each aspect, 10 models with the smallest errors in 20C are tagged with an asterisk. We
694 then use the total number of asterisk to identify the overall best 10 models. After that we
695 summarize the model performance in 20C and the projection for 21C using the MMM of all

696 37 models, the overall best 10 models (b10) and the best 10 models for individual aspect
697 (c10) (Fig.13). Skewness is not shown in this summary since it is closely correlated with
698 EN-LN asymmetry in amplitude.

699 We note that the MMM using the subset of good models are closer to the 20C observation
700 for each aspect, but the future projection based on these models still lack of consensus for
701 most aspects. It is possible that (i) the external forcing for 21C may be implemented slightly
702 differently in each individual model (ii) model dynamics are slightly different in each model
703 so that they may drive different responses even under the same forcing. Therefore one needs
704 to be cautious for the future projection even using a subset of “good” models.

705 *c. ENSO behavior biases in CMIP5 models*

706 Compared to the 20C observation, 37-model mean results (a37 in Fig. 13) show the
707 CMIP5 model biases are reflected in many aspects, e.g., underestimated mean SST, overly
708 weak annual cycle, overly strong SST variability, more CPEN and EPLN, more eastward
709 propagating EN, underestimated EN-LN asymmetry in amplitude and duration and excessive
710 asymmetry in transition.

711 Do the biases in ENSO behavior have some relation with the biases in the simulated mean
712 state of Niño-3.4? Fig. 14 displays the results for PI, 20C and 21C. Note that the scatter
713 plot shows the overall spread when 37 earth-like systems respond given the same forcing. It
714 may not be interpreted as a physical relation, since it does not come from one consistent
715 system under a series of different mean states. For example, Fig. 13b shows that, among all
716 37 models, models which simulate a relatively warmer mean state tend to also simulate a
717 weak annual cycle with a small summer-winter difference in SST. But this apparent relation
718 may not suggest that the annual cycle will become weaker in the warming climate. Previous
719 section already showed that the change of the annual cycle from 20C to 21C varies in each
720 individual model and the MMM shows no significant change.

7. Conclusion

We introduced a set of empirical probabilistic diagnostics for ENSO behaviors, including variations in peak season, location and propagation direction as well as El Niño-La Niña asymmetries. The diagnostics applied to SST observations show that, El Niños and La Niñas are phase-locked to boreal winter. They both favor westward propagation. El Niños mainly occur at the eastern Pacific and La Niñas prefer the central Pacific. These results agree with current understanding and thus provide support for the validity of our new diagnostics.

The diagnostics were applied to evaluate ENSO behaviors in two example simulations. The first is a 4000-year pre-industrial control simulation of the GFDL CM2.1 coupled GCM. The strong nonlinearity of this model is indicated by an exaggerated El Niño-La Niña asymmetry. Although modeled La Niñas generally behave like the observations, El Niños behave quite differently. El Niño’s winter phase-locking feature is largely missed since EPENs prefer peaking in summer while CPENs prefer winter. Eastward propagating El Niños are dominant. The overall statistics is largely dominated by extreme El Niños.

The diagnostics were also applied to a 4000-year stochastic-forced simulation of a non-linear empirical model reduction (EMR) fit using SST observations. This simulation is reasonably realistic in broad aspects of ENSO behavior and thus may be considered as an extension to observations to help us assess the range of ENSO variation. Most epochs in a 4000-year simulation agree well with observations. But epochs with more CP El Niños or epochs with more eastward El Niños do exist when stochastic noise is the only forcing. No forcing trend such as that due to greenhouse gases is required.

The diagnostics were then applied to assess the potential change of ENSO behaviors in a warming climate using 37 CMIP5 models that participated in IPCC AR5. Evaluation of model performance used 20th century runs (20C, historical, 1900-1999) show that each model has pros and cons for varying aspects of ENSO behavior.

As to the projected changes from the 20th century to 21st century (21C, RCP8.5 scenario, 2000-2099), except for a consensus in tropical Pacific SST increase due to the forcing, changes

748 in other aspects are all model dependent. Except for the warming climatology, many 21C
749 changes are within the bounds of the natural variation range produced by the pre-industrial
750 control runs (PI) control run. Overall the multi-model mean (MMM) suggests that changes
751 in many ENSO statistics measured in SSTA may not be significant, e.g., diversity peaking
752 in eastern Pacific /central Pacific and El Niño-La Niña asymmetries. Although a few models
753 do show significant changes, the degree of model agreement on the projected change is low
754 for all aspects. A shift favoring eastward propagating El Niño and La Niña shows slightly
755 more robustness.

756 Projections for the future based on CMIP models often involve considerable uncertainty
757 (Vecchi and Wittenberg 2010). Changes in ENSO are difficult to detect given large natu-
758 ral variability present in each model (e.g., Wittenberg 2009) as well as the lack of model
759 agreement (e.g., Guilyardi 2006; Collins et al. 2010; Stevenson 2012; Taschetto et al. 2014).
760 In this study, various model biases for 20C ENSO behaviors leave little to confidently pre-
761 dict the future of ENSO. Whether the projected changes could actually take place in the
762 future remains largely uncertain. The ENSO behavior diagnostics introduced in this study
763 and data-driven models (e.g., EMR) fit from the observation may be useful along with the
764 development of CMIP models.

765 *Acknowledgments.*

766 We thank Dong Eun Lee, Naomi Henderson and David Chapman for inspiring discussions
767 and Haibo Liu for archiving the CMIP5 data at LDEO. We also thank three anonymous re-
768 viewers for their helpful comments. This study is supported by the Office of Naval Research
769 under the research grant MURI (N00014-12-1-0911). DC also acknowledge grants from the
770 National Basic Research Program (2013CB430302), the National Natural Science Founda-
771 tion of China (41321004, 91128204) and the IPOVAR Project. We acknowledge the World
772 Climate Research Programme’s Working Group on Coupled Modelling, which is responsi-
773 ble for CMIP, and we thank the climate modeling groups (listed in Table 1 of this paper)

774 for producing and making available their model output. For CMIP the U.S. Department
 775 of Energy’s Program for Climate Model Diagnosis and Intercomparison provides coordinat-
 776 ing support and led development of software infrastructure in partnership with the Global
 777 Organization for Earth System Science Portals.

778 APPENDIX

780 EMR methodology

781 Empirical Model Reduction (EMR) is an empirical modeling framework, allowing nonlin-
 782 earity, seasonality and memory effect (Kravtsov et al. 2005; Kondrashov et al. 2005, 2015).
 783 The operational version of EMR (labelled as UCLA-TCD) participates in IRI’s ENSO pre-
 784 diction plume, and it is very competitive among both dynamical and statistical models
 785 (Barnston et al. 2012; [http://iri.columbia.edu/our-expertise/climate/forecasts/](http://iri.columbia.edu/our-expertise/climate/forecasts/enso/current/)
 786 [enso/current/](http://iri.columbia.edu/our-expertise/climate/forecasts/enso/current/)).

787 The setting of EMR used in this study is as follows. The state vector $\mathbf{x} = \{x_i\}$ is
 788 the leading 3 normalized PCs of detrended tropical Pacific SSTA. Quadratic nonlinearity is
 789 included in the main level:

$$dx_i = (\mathbf{x}^T \mathbf{A}_i \mathbf{x} + \mathbf{b}_i^1 \mathbf{x} + c_i^1) dt + dr_i^1; \quad i = 1, \dots, 3. \quad (\text{A1})$$

790 The model coefficients in matrices \mathbf{A}_i , the vectors \mathbf{b}_i of matrix \mathbf{B} , the components c_i of vector
 791 \mathbf{c} and the components r_i of the residual \mathbf{r} are determined by multiple polynomial regression.

792 Seasonality is included by adding additional coefficients into the main level of the model:

$$\mathbf{B} = \mathbf{B}_n + \mathbf{B}_s \sin(2\pi t/T) + \mathbf{B}_c \cos(2\pi t/T) \quad (\text{A2})$$

$$\mathbf{c} = \mathbf{c}_n + \mathbf{c}_s \sin(2\pi t/T) + \mathbf{c}_c \cos(2\pi t/T) \quad (\text{A3})$$

793 where the matrix \mathbf{B}_n and vector \mathbf{c}_n are the original annually averaged (non-seasonal) terms
 794 as in Eq. (A1), matrices \mathbf{B}_s and \mathbf{B}_c add an multiplicative seasonality, vectors \mathbf{c}_s and \mathbf{c}_c
 795 add an additive seasonality. The period $T = 12$ months to account for an annual cycle of
 796 seasonality. All these coefficients are determined simultaneously with the other coefficients
 797 in the main level.

798 The ENSO memory effect is embedded in a 2 time-level model construction. An addi-
 799 tional level is added by fitting the temporal increment of the residual at the main level $d\mathbf{r}^1$
 800 using a linear function of an extended state vector $[\mathbf{x}, \mathbf{r}^1]$.

$$dr_i^1 = \mathbf{b}_i^2[\mathbf{x}, \mathbf{r}^1]dt + dr_i^2; \quad i = 1, \dots, 3 \quad (\text{A4})$$

801 where \mathbf{b}_i^2 and r_i^2 for the second level (i.e., one timestep back) are determined after the main
 802 level. Results from Kondrashov et al. (2005) and Chen et al. (2016) indicate that 2 time
 803 levels are sufficient to embed a memory effect for ENSO simulation. A stochastic simulation
 804 is forced using a spatially coherent multivariate white noise given by the residual covariance
 805 matrix estimated along with the model fitting. See more details in Kondrashov et al. (2005).

806 In this study, EMR uses the same setting as the model 2L+S+NL in Chen et al. (2016),
 807 which is a nonlinear model with two time levels and an annual seasonality. To investigate
 808 the dependence of ENSO behaviors on the system nonlinearity, we also construct a linear
 809 model denoted as “EMR-L” for a comparison. It uses the same setting as the model 2L+S in
 810 Chen et al. (2016), which has two time levels and an annual seasonality. The only difference
 811 between EMR and EMR-L is with or without nonlinearity.

REFERENCES

- 814 Allen, M. R. and W. J. Ingram, 2002: Constraints on future changes in climate and the
815 hydrologic cycle. *Nature*, **419**, 224–232.
- 816 An, S. I. and F.-F. Jin, 2004: Nonlinearity and asymmetry of ENSO. *J. Climate*, **17**, 2399–
817 2412.
- 818 An, S.-I. and B. Wang, 2001: Mechanisms of locking of the El Niño and La Niña mature
819 phases to boreal winter. *J. Climate*, **14**, 2164–2176.
- 820 Ashok, K., S. K. Behera, S.A. Rao, H. Weng, and T. Yamagata, 2007: El Niño Modoki and
821 its possible teleconnection. *J. Geophys. Res.*, **112**, C11 007.
- 822 Barnston, A. G., M. K. Tippett, M. L. L’Heureux, S. Li, and D. G. DeWitt, 2012: Skill of
823 real-time seasonal ENSO model predictions during 2002–11: Is our capability increasing?
824 *Bull. Amer. Meteor. Soc.*, **93**, 631–651.
- 825 Bellenger, H., E. Guilyardi, J. Leloup, M. Lengaigne, and J. Vialard, 2013: ENSO represen-
826 tation in climate models: from CMIP3 to CMIP5. *Climate Dyn.*, **42**, 1999–2018.
- 827 Cai, W., et al., 2014: Increasing frequency of extreme El Niño events due to greenhouse
828 warming. *Nat. Climate Chang.*, **4**, 111–116.
- 829 Cai, W., et al., 2015a: ENSO and greenhouse warming. *Nat. Climate Chang.*, **5**, 849–859.
- 830 Cai, W., et al., 2015b: Increased frequency of extreme LaNiña events under greenhouse
831 warming. *Nat. Climate Chang.*, **5**, 132–137.
- 832 Cane, M. A., A. C. Clement, A. Kaplan, Y. Kushnir, D. Pozdnyakov, R. Seager, S. E.
833 Zebiak, and R. Murtugudde, 1997: Twentieth-century sea surface temperature trends.
834 *Science*, **275**, 957–960.

- 835 Cane, M. A., S. E. Zebiak, and Y. Xue, 1995: Model studies of the long term behavior of
836 ENSO. *Nat. Climate Var. Decad. time scales*, Natl. Acad. Press, 442–457.
- 837 Capotondi, A., A. Wittenberg, and S. Masina, 2006: Spatial and temporal structure of
838 Tropical Pacific interannual variability in 20th century coupled simulations. *Ocean Model.*,
839 **15**, 274–298.
- 840 Capotondi, A., et al., 2015: Understanding ENSO diversity. *Bull. Amer. Meteor. Soc.*, **96**,
841 921–938.
- 842 Chen, C., M. A. Cane, N. Henderson, D. E. Lee, D. Chapman, D. Kondrashov, and M. D.
843 Chekroun, 2016: Diversity, Nonlinearity, Seasonality, and Memory Effect in ENSO Sim-
844 ulation and Prediction Using Empirical Model Reduction. *J. Climate*, **29**, 1809–1830,
845 doi:10.1175/JCLI-D-15-0372.1.
- 846 Chen, D., et al., 2015a: Strong influence of westerly wind bursts on El Niño diversity. *Nat.*
847 *Geosci.*, **8**, 339–345.
- 848 Chen, L., T. Li, and Y. Yu, 2015b: Causes of Strengthening and Weakening of ENSO
849 Amplitude under Global Warming in Four CMIP5 Models. *J. Climate*, **28**, 3250–3274,
850 doi:10.1175/JCLI-D-14-00439.1.
- 851 Choi, K.-Y., G. A. Vecchi, and A. T. Wittenberg, 2013: ENSO transition, duration, and
852 amplitude asymmetries: role of the nonlinear wind stress coupling in a conceptual model.
853 *J. Climate*, **26**, 9462–9476.
- 854 Clement, A. C., R. Seager, M. A. Cane, and S. E. Zebiak, 1996: An ocean dynamical
855 thermostat. *J. Climate*, **9**, 2190–2196.
- 856 Collins, M., et al., 2010: The impact of global warming on the tropical Pacific Ocean and
857 El Niño. *Nat. Geosci.*, **3**, 391–397.

- 858 Delworth, T. L., et al., 2006: GFDL’s CM2 global coupled climate models. part I: formulation
859 and simulation characteristics. *J. Climate*, **19**, 643–674.
- 860 DiNezio, P. N. and C. Deser, 2014: Nonlinear Controls on the Persistence of La Niña. *J.*
861 *Climate*, **27**, 7335–7355, doi:10.1175/JCLI-D-14-00033.1.
- 862 Dommenges, D., T. Bayr, and C. Frauen, 2013: Analysis of the non-linearity in the pattern
863 and time evolution of El Niño southern oscillation. *Climate Dyn.*, **40**, 2825–2847.
- 864 Fedorov, A. V. and S. G. Philander, 2001: A stability analysis of tropical ocean-atmosphere
865 interactions: Bridging measurements and theory for El Niño. *J. Climate*, **14**, 3086–3101.
- 866 Fedorov, A. V., S. Hu, M. Lengaigne, and E. Guilyardi, 2015: The impact of westerly wind
867 bursts and ocean initial state on the development, and diversity of El Niño events. *Climate*
868 *Dyn.*, **44**, 1381–1401.
- 869 Frauen, C. and D. Dommenges, 2010: El Niño and La Niña amplitude asymmetry caused
870 by atmospheric feedbacks. *Geophys. Res. Lett.*, **37**, L18 801.
- 871 Gleckler, P. J., K. E. Taylor, and C. Doutriaux, 2008: Performance metrics for climate
872 models. *J. Geophys. Res. Atmos.*, **113**, 1–20.
- 873 Guilyardi, E., 2006: El Niño-mean state-seasonal cycle interactions in a multi-model ensem-
874 ble. *Climate Dyn.*, **26**, 329–348.
- 875 Guilyardi, E., W. Cai, M. Collins, A. Fedorov, F.-F. Jin, A. Kumar, D.-Z. Sun, and A. Wit-
876 tenberg, 2012: New strategies for evaluating ENSO processes in climate models. *Bull.*
877 *Amer. Meteor. Soc.*, **93**, 235–238.
- 878 Guilyardi, E., P. Delecluse, S. Gualdi, and A. Navarra, 2003: Mechanisms for ENSO phase
879 change in a coupled GCM. *J. Climate*, **16**, 1141–1158.

- 880 Guilyardi, E., A. Wittenberg, A. Fedorov, M. Collins, C. Wang, A. Capotondi, G. J. van Old-
881 enborgh, and T. Stockdale, 2009: Understanding El Niño in Ocean–Atmosphere General
882 Circulation Models: Progress and Challenges. *Bull. Amer. Meteor. Soc.*, **90**, 325–340.
- 883 Ham, Y.-G. and J.-S. Kug, 2012: How well do current climate models simulate two types of
884 El Niño? *Climate Dyn.*, **39**, 383–398.
- 885 Ham, Y.-G. and J.-S. Kug, 2014: ENSO phase-locking to the boreal winter in CMIP3 and
886 CMIP5 models. *Climate Dyn.*, **43**, 305–318.
- 887 Hirahara, S., M. Ishii, and Y. Fukuda, 2014: Centennial-scale sea surface temperature anal-
888 ysis and its uncertainty. *J. Climate*, **27**, 57–75.
- 889 Kang, I.-S. and J.-S. Kug, 2002: El Niño and La Niña sea surface temperature anomalies:
890 asymmetry characteristics associated with their wind stress anomalies. *J. Geophys. Res.*,
891 **107**, 4372.
- 892 Kao, H.-Y. and J.-Y. Yu, 2009: Contrasting eastern-Pacific and central-Pacific types of
893 ENSO. *J. Climate*, **22**, 615–632.
- 894 Karamperidou, C., M. A. Cane, U. Lall, and A. T. Wittenberg, 2014: Intrinsic modulation
895 of ENSO predictability viewed through a local Lyapunov lens. *Climate Dyn.*, **42**, 253–270.
- 896 Karl, T. R., R. W. Knight, and N. Plummer, 1995: Trends in high-frequency climate vari-
897 ability in the twentieth century. *Nature*, **377**, 217–220.
- 898 Karnauskas, K. B., 2013: Can we distinguish canonical El Niño from Modoki? *Geophys.*
899 *Res. Lett.*, **40**, 5246–5251.
- 900 Kim, S. T. and J.-Y. Yu, 2012: The two types of ENSO in CMIP5 models. *Geophys. Res.*
901 *Lett.*, **39**, L11 704.
- 902 Kim, W. and W. Cai, 2014: The importance of the eastward zonal current for generating
903 extreme El Niño. *Climate Dyn.*, **42**, 3005–3014.

- 904 Kondrashov, D., M. D. Chekroun, and M. Ghil, 2015: Data-driven non-Markovian closure
905 models. *Phys. D Nonlinear Phenom.*, **297**, 33–55.
- 906 Kondrashov, D., S. Kravtsov, A. W. Robertson, and M. Ghil, 2005: A hierarchy of data-
907 based ENSO models. *J. Climate*, **18**, 4425–4444.
- 908 Kravtsov, S., D. Kondrashov, and M. Ghil, 2005: Multilevel regression modeling of nonlinear
909 processes: derivation and applications to climatic variability. *J. Climate*, **18**, 4404–4424.
- 910 Kug, J.-S., J. Choi, S.-I. An, F.-F. Jin, and A. T. Wittenberg, 2010: Warm pool and cold
911 tongue El Niño events as simulated by the GFDL 2.1 coupled GCM. *J. Climate*, **23**,
912 1226–1239.
- 913 Kug, J.-S. and Y.-G. Ham, 2011: Are there two types of La Nina? *Geophys. Res. Lett.*, **38**,
914 L16 704.
- 915 Kug, J.-S., F.-F. Jin, and S.-I. An, 2009: Two Types of El Niño events: Cold Tongue El
916 Niño and Warm Pool El Niño. *J. Climate*, **22**, 1499–1515.
- 917 Larkin, N. and D. Harrison, 2002: ENSO warm (El Niño) and cold (La Niña) event life
918 cycles: Ocean surface anomaly patterns, their symmetries, asymmetries, and implications.
919 *J. Climate*, **15**, 1118–1140.
- 920 Larkin, N. K. and D. E. Harrison, 2005: On the definition of El Niño and associated seasonal
921 average U.S. weather anomalies. *Geophys. Res. Lett.*, **32**, L13 705.
- 922 Lee, S.-K., A. T. Wittenberg, D. B. Enfield, S. J. Weaver, C. Wang, and R. Atlas, 2016:
923 US regional tornado outbreaks and their links to spring ENSO phases and North Atlantic
924 SST variability. *Environ. Res. Lett.*, **11**, 044 008, doi:10.1088/1748-9326/11/4/044008.
- 925 Lee, T. and M. J. McPhaden, 2010: Increasing intensity of El Niño in the central-equatorial
926 Pacific. *Geophys. Res. Lett.*, **37**, L14 603.

- 927 Lengaigne, M. and G. A. Vecchi, 2010: Contrasting the termination of moderate and extreme
928 El Niño events in coupled general circulation models. *Climate Dyn.*, **35**, 299–313.
- 929 Levine, A. F. Z. and F. F. Jin, 2015: A simple approach to quantifying the noise-ENSO
930 interaction. Part I: deducing the state-dependency of the windstress forcing using monthly
931 mean data. *Climate Dyn.*, doi:10.1007/s00382-015-2748-1.
- 932 McPhaden, M. J. and X. Zhang, 2009: Asymmetry in zonal phase propagation of ENSO sea
933 surface temperature anomalies. *Geophys. Res. Lett.*, **36**, L13 703.
- 934 Neelin, J. D., F.-F. Jin, and H.-H. Syu, 2000: Variations in ENSO phase locking. *J. Climate*,
935 **13**, 2570–2590.
- 936 Newman, M., S.-I. Shin, and M. Alexander, 2011: Natural variation in ENSO flavors. *Geo-*
937 *phys. Res. Lett.*, **38**, L14 705.
- 938 Ohba, M. and H. Ueda, 2009: Role of nonlinear atmospheric response to SST on the asym-
939 metric transition process of ENSO. *J. Climate*, **22**, 177–192.
- 940 Okumura, Y. M., M. Ohba, C. Deser, and H. Ueda, 2011: A proposed mechanism for the
941 asymmetric duration of El Niño and La Niña. *J. Climate*, **24**, 3822–3829.
- 942 Power, S., F. Delage, C. Chung, G. Kociuba, and K. Keay, 2013: Robust twenty-first-century
943 projections of El Niño and related precipitation variability. *Nature*, **502**, 541–545.
- 944 Rayner, N. A., D. E. Parker, E. B. Horton, C. Folland, L. V. Alexander, D. P. Rowell, E. C.
945 Kent, and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice, and night
946 marine air temperature since the late nineteenth century. *J. Geophys. Res.*, **108**, 4407.
- 947 Santoso, A., S. McGregor, F.-F. Jin, W. Cai, M. H. England, S.-I. An, M. J. McPhaden,
948 and E. Guilyardi, 2013: Late-twentieth-century emergence of the El Niño propagation
949 asymmetry and future projections. *Nature*, **504**, 126–30.

- 950 Schopf, P. and R. Burgman, 2006: A simple mechanism for ENSO residuals and asymmetry.
951 *J. Climate*, **19**, 3167–3179.
- 952 Solomon, A. and M. Newman, 2012: Reconciling disparate twentieth-century Indo-Pacific
953 ocean temperature trends in the instrumental record. *Nat. Climate Chang.*, **2**, 691–699.
- 954 Stein, K., N. Schneider, A. Timmermann, and F.-F. Jin, 2010: Seasonal Synchronization
955 of ENSO Events in a Linear Stochastic Model. *J. Climate*, **23**, 5629–5643, doi:10.1175/
956 2010JCLI3292.1.
- 957 Stevenson, S., B. Fox-Kemper, M. Jochum, R. Neale, C. Deser, and G. Meehl, 2012: Will
958 there be a significant change to El Niño in the twenty-first century? *J. Climate*, **25**,
959 2129–2145.
- 960 Stevenson, S. L., 2012: Significant changes to ENSO strength and impacts in the twenty-first
961 century: Results from CMIP5. *Geophys. Res. Lett.*, **39**, L17 703.
- 962 Takahashi, K. and B. Dewitte, 2015: Strong and moderate nonlinear El Niño regimes. *Cli-*
963 *mate Dyn.*, in press, doi:10.1007/s00382-015-2665-3.
- 964 Takahashi, K., A. Montecinos, K. Goubanova, and B. Dewitte, 2011: ENSO regimes: Rein-
965 terpreting the canonical and Modoki El Niño. *Geophys. Res. Lett.*, **38**, L10 704.
- 966 Taschetto, A. S. and M. H. England, 2009: El Niño Modoki impacts on Australian rainfall.
967 *J. Climate*, **22**, 3167–3174.
- 968 Taschetto, A. S., A. S. Gupta, N. C. Jourdain, A. Santoso, C. C. Ummenhofer, and M. H.
969 England, 2014: Cold tongue and warm pool ENSO events in CMIP5: mean state and
970 future projections. *J. Climate*, **27**, 2861–2885.
- 971 Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the
972 experiment design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498.

- 973 Tziperman, E., M. A. Cane, and S. E. Zebiak, 1995: Irregularity and locking to the seasonal
974 cycle in an ENSO prediction model as explained by the quasi-periodicity route to chaos.
975 *J. Atmos. Sci.*, **52**, 293–306.
- 976 Tziperman, E., M. A. Cane, S. E. Zebiak, Y. Xue, and B. Blumenthal, 1998: Locking of
977 El Niño’s peak time to the end of the calendar year in the delayed oscillator picture of
978 ENSO. *J. Climate*, **11**, 2191–2199.
- 979 Tziperman, E., S. E. Zebiak, and M. A. Cane, 1997: Mechanisms of seasonal - ENSO
980 interaction. *J. Atmos. Sci.*, **54**, 61–71.
- 981 Vecchi, G. A. and A. T. Wittenberg, 2010: El Niño and our future climate: where do we
982 stand? *Wiley Interdiscip. Rev. Chang.*, **1**, 260–270.
- 983 Weng, H., K. Ashok, S. K. Behera, S. a. Rao, and T. Yamagata, 2007: Impacts of recent
984 El Niño Modoki on dry/wet conditions in the Pacific rim during boreal summer. *Climate*
985 *Dyn.*, **29**, 113–129.
- 986 Wittenberg, A. T., 2009: Are historical records sufficient to constrain ENSO simulations?
987 *Geophys. Res. Lett.*, **36**, L12702.
- 988 Wittenberg, A. T., A. Rosati, T. L. Delworth, G. A. Vecchi, and F. Zeng, 2014: ENSO
989 modulation: Is it decadal predictability? *J. Climate*, **27**, 2667–2681.
- 990 Wittenberg, A. T., A. Rosati, N.-C. Lau, and J. J. Ploshay, 2006: GFDL’s CM2 Global
991 Coupled Climate Models. Part III: Tropical Pacific Climate and ENSO. *J. Climate*, **19**,
992 698–722.
- 993 Xiao, H. and C. R. Mechoso, 2009: Seasonal Cycle–El Niño Relationship: Validation of
994 Hypotheses. *J. Atmos. Sci.*, **66**, 1633–1653.
- 995 Xie, S.-P., C. Deser, G. A. Vecchi, J. Ma, H. Teng, and A. T. Wittenberg, 2010: Global

996 Warming Pattern Formation: Sea Surface Temperature and Rainfall. *J. Climate*, **23**, 966–
997 986.

998 Yeh, S.-W., B. P. Kirtman, J.-S. Kug, W. Park, and M. Latif, 2011: Natural variability of
999 the central Pacific El Niño event on multi-centennial timescales. *Geophys. Res. Lett.*, **38**,
1000 L02704.

1001 Yeh, S.-W., J.-S. Kug, B. Dewitte, M.-H. Kwon, B. P. Kirtman, and F.-F. Jin, 2009: El
1002 Niño in a changing climate. *Nature*, **461**, 511–4.

1003 Yu, J.-Y. and S. T. Kim, 2010: Identification of Central-Pacific and Eastern-Pacific types of
1004 ENSO in CMIP3 models. *Geophys. Res. Lett.*, **37**, L15705.

1005 Yu, J.-Y., Y. Zou, S. T. Kim, and T. Lee, 2012: The changing impact of El Niño on US
1006 winter temperatures. *Geophys. Res. Lett.*, **39**, L15702., doi:10.1029/2012GL052483.

1007 Zebiak, S. E. and M. A. Cane, 1987: A Model El Niño-Southern Oscillation. *Mon. Wea.*
1008 *Rev.*, **115**, 2262–2278.

1009 Zhang, T. and D. Z. Sun, 2014: ENSO asymmetry in CMIP5 models. *J. Climate*, **27**, 4070–
1010 4093.

1011 List of Tables

1012 1 List of 37 CMIP5 models analyzed in this study. Due to the lack of availability
1013 in certain models for temperature of ocean surface “tos”, we instead analyzed
1014 monthly surface temperature “ts” in each model’s r1i1p1 run. 1st column is
1015 the official model name. 2nd column is the length of pre-industrial control run
1016 (year). 3rd to 15th columns are the model rank as shown in each individual
1017 figure and panel, e.g., f9b indicating Fig. 9 panel b. According ENSO aspect
1018 is labelled in 2nd and 3rd rows. Note that for individual ENSO behavior
1019 models with * are 10 models with smallest error between each model’s 20th
1020 century run (orange o) and the 20th century observation value (black line in
1021 the panel). The last column is the total number of * for each model. There
1022 are best 10 models with 5 and above *, which are indicated with + at the end
1023 the model name in 1st column. Note that these relative better models are
1024 only restricted to ENSO behavior aspects analyzed in this study, therefore,
1025 it is not generally applicable to model performances on other phenomena.
1026 Model center information and experiment designs see Taylor et al. (2012)
1027 and CMIP5 website (<http://cmip-pcmdi.llnl.gov/cmip5/>)

43

TABLE 1. List of 37 CMIP5 models analyzed in this study. Due to the lack of availability in certain models for temperature of ocean surface “tos”, we instead analyzed monthly surface temperature “ts” in each model’s r1i1p1 run. 1st column is the official model name. 2nd column is the length of pre-industrial control run (year). 3rd to 15th columns are the model rank as shown in each individual figure and panel, e.g., f9b indicating Fig. 9 panel b. According ENSO aspect is labelled in 2nd and 3rd rows. Note that for individual ENSO behavior models with * are 10 models with smallest error between each model’s 20th century run (orange o) and the 20th century observation value (black line in the panel). The last column is the total number of * for each model. There are best 10 models with 5 and above *, which are indicated with + at the end the model name in 1st column. Note that these relative better models are only restricted to ENSO behavior aspects analyzed in this study, therefore, it is not generally applicable to model performances on other phenomena. Model center information and experiment designs see Taylor et al. (2012) and CMIP5 website (<http://cmip-pcmdi.llnl.gov/cmip5/>)

name	PI	f9b	f10b	f9d	f12a	f10d	f10f	f11a	f11c	f11b	f11d	f12b	f12c	f12d	num
		cli	secli	s.d.	s.k.	I_{sea} EN	I_{sea} LN	$I_{cp/ep}$ EN	$I_{cp/ep}$ LN	$I_{e/w}$ EN	$I_{e/w}$ LN	I_{amp}	I_{dur}	I_{tra}	
ACCESS1.0+	250	23*	24*	8	13	5	3	4*	7	17	22*	12	33	11*	5
ACCESS1.3+	500	26*	8	10	22	34	35	13*	6	6*	13	26*	12	7*	5
BCC-CSM1.1	500	22*	5	12	14	2	2	20	3	19	21*	13	13	3*	3
BCC-CSM1.1(m)	400	29*	9	32	29*	9*	5	18	24	29	6	23	3*	27	4
BNU-ESM+	559	25*	11	36	9	11*	18*	5*	30*	35	32	14	5*	25	6
CanESM2	996	17	26*	28	17	14*	17*	19	5	5*	7	19	29	31	4
CCSM4	501	27*	23*	31	34	23	19*	26	37	27	1	35*	11	35	4
CESM1(BGC)	500	28*	14	24	20	29	26	6*	34*	23	11	24	22	30	3
CESM1(CAM5)	319	12	37	29	35	15*	7	36	27*	20	10	31*	17	34	3
CMCC-CESM	277	24*	22*	35	33	33	34	16	23	3	8	32*	21	23	3
CMCC-CM	330	32	27*	9	28*	24	23	8*	25	30	35	28*	19	24	4
CMCC-CMS+	500	30*	29*	23*	16	28	27	14	31*	11*	15*	16	15	17	6
CNRM-CM5	850	19	28*	27	26*	1	1	22	15	34	30	25	18	26	2
CSIRO-Mk3.6.0	500	1	32	22*	4	31	36	34	1	7*	19*	2	31	14	3
FGOALS-g2	700	21	16	18*	7	8	9	12*	13	33	27	7	27	9*	3
FIO-ESM	800	34	6	33	3	18*	14*	3	18	37	31	3	36	2	2
GFDL-CM3	500	10	4	30	12	19	25	28	11	14	4	8	25	29	0
GFDL-ESM2G+	500	5	30*	14*	23*	32	10*	35	26*	4*	5	15	34	13	6
GFDL-ESM2M	500	15	13	37	36	37	32	30	35*	36	20*	36	4*	37	3
GISS-E2-H	240	33	7	3	8	10*	12*	23	2	16	29	11	23	22	2
GISS-E2-H-CC	251	35	12	16*	2	3	13*	27	19	25	14	1	37	20	2
GISS-E2-R	300	36	21*	7	18	6	6	2	22	28	33	21	6*	10*	3
GISS-E2-R-CC	251	37	20	4	1	4	11*	7*	16	31	36	4	24	18	2
HadGEM2-CC	240	8	17	19*	5	20	20	15	14	26	28	6	20	32	1
HadGEM2-ES+	239	9	18	17*	6	12*	8	29	29*	22	17*	5	26	4*	5
INM-CM4	500	13	3	6	24*	27	29	21	4	18	25	20	16	21	1
IPSL-CM5A-LR	1000	7	19	15*	15	30	30	17	21	2	12	18	1*	6*	3
IPSL-CM5A-MR	300	16	10	20*	25*	21	31	25	10	1	9	30*	2*	15	4
IPSL-CM5B-LR+	300	31*	2	11	21	16*	28	1	28*	13*	23*	29*	9*	1	7
MIROC-ESM	531	2	36	1	32*	35	24	32	9	21	37	34*	32	12*	3
MIROC-ESM-CHEM	255	3	35	2	30*	25	21	24	17	15	34	33*	14	5*	3
MIROC5	200	11	31	34	37	22	16*	37	36	32	16*	37	35	36	2
MPI-ESM-LR+	1000	4	34	21*	31*	26	33	11*	20	8*	18*	27*	28	28	6
MPI-ESM-MR	1000	6	33	13	19	36	37	33	12	24	26	17	10*	16	1
MRI-CGCM3+	500	20	1	5	10	17*	22	10*	8	9*	24*	9	7*	8*	6
NorESM1-M+	501	18	15	25	11	13*	15*	9*	33*	12*	3	10	8*	33	6
NorESM1-ME	252	14	25*	26	27*	7	4	31	32*	10*	2	22	30	19	4

1028 List of Figures

- 1029 1 The first three normalized EOF patterns of Tropical Pacific SSTA in OBS
1030 (1870-present monthly HadISST v1.1) (a, c, e) and GCM (4000-year monthly
1031 GFDL CM2.1 pre-industrial control run) (b, d, f). Positive/negative values
1032 are shown in solid/dashed contours. The zero value is highlighted in the thick
1033 solid contour. 50
- 1034 2 Simulation evaluation: (a, d, g) for OBS, (b, e, h) for 4000 year EMR and (c,
1035 f, i) for 4000-year GCM. (a, b, c) for Probability density function (PDF) of
1036 PC1. (d, e, f) for PDF of PC2. (g, h, i) for decimal logarithm of the bivariate
1037 probability density function (2dPDF) in PC1-PC2 space. In PDF panels,
1038 OBS is divided into 5 overlapping 100-year epochs with 10 years apart. EMR
1039 and GCM are both divided into 40 non-overlapping 100-year epochs. PDF
1040 and according skewness are calculated in epochs and the average is shown. 51
- 1041 3 Definition of 5 ENSO states (a) (Eastern/Central Pacific El Niño/La Niña
1042 and neutral patterns, denoted as EPEN/ CPEN/ EPLN/ CPLN/ NEU) are
1043 shown using smoothed monthly PCs from 1870-present HadiSST v1.1. ± 0.7
1044 s.d. (PC1), where s.d. is one standard deviation, is used to distinguish EN/LN
1045 from NEU. Zero line of PC2 is used to distinguish EP/CP states. Each state
1046 is assigned a color code for further analysis. (b-e) the SSTA patterns averaged
1047 for eastern/central Pacific ENSO states. (f) normalized PC1 (1950-2010) with
1048 each ENSO states color-coded. 52

1049 4 Monthly occurrence probabilities for 5 ENSO state in OBS (a), EMR (b) and
1050 GFDL CM2.1 GCM (c) are shown. Stacked bars along the vertical coordinate
1051 are the occurrence probabilities of each color-coded state. The horizontal co-
1052 ordinate is the calendar month from Jan to Dec. Full year data is used thus
1053 the sample size for each calendar month is equal. In panel a, higher prob-
1054 ability of EN/LN in winter months than summer months indicates observed
1055 ENSO's winter phase locking. Higher probability of EPEN over CPEN in-
1056 dicates El Niño favor peaking in EP, and higher probability of CPLN over
1057 EPLN indicates LN favor peaking in CP. 53

1058 5 State transition probabilities for EN (1st row), NEU (2nd row) and LN (3rd
1059 row) in OBS (first column), EMR (2nd column) and GFDL CM2.1 GCM (3rd
1060 column) are shown. The horizontal coordinate represents the transition from
1061 the past (-3 years) to the future (+3 years) in monthly intervals, with zero in-
1062 dicating the current state. Taking GCM EN transition (panel c) for example,
1063 bars along the vertical coordinate at +1 year (+12 months) represent: the self-
1064 transition probability $P_{EN(t+\tau)|EN(t)}$ (upper bar), the $P_{NEU(t+\tau)|EN(t)}$ (middle
1065 bar) and the opposite-sign transition $P_{LN(t+\tau)|EN(t)}$ (lower bar). The decay-
1066 ing of $P_{EN(t+\tau)|EN(t)}$ as function of lead time indicates EN's duration. The
1067 discrepancy between $P_{LN(t-\tau)|EN(t)}$ (lead<0 side) and $P_{LN(t+\tau)|EN(t)}$ (lead>0
1068 side) indicates the EN-LN asymmetry in transition. The transition probabil-
1069 ities generally converge to the climatology, i.e., the nonseasonal occurrence
1070 probability of each state (dotted line in each panel). 54

- 1071 6 Zonal propagation asymmetry in OBS(a,b), EMR(c,d) and GCM(e,f). (a,c,e):
1072 transition probabilities conditioned on CPEN state $P_{EPEN(t-\tau)|CPEN(t)}$ (neg-
1073 ative lead) and $P_{EPEN(t+\tau)|CPEN(t)}$ (positive lead), which generally converge
1074 to the occurrence probability P_{EPEN} (dotted line) toward a large lead time.
1075 (b,d,f): transition probabilities conditioned on EPLN state $P_{CPLN(t-\tau)|EPLN(t)}$
1076 (negative lead) and $P_{CPLN(t+\tau)|EPLN(t)}$ (positive lead), which generally con-
1077 verge to the occurrence probability P_{CPLN} (dotted line). 55
- 1078 7 Variation of ENSO behaviors in 4000-year EMR and GFDL CM2.1 GCM
1079 simulation. Each index (see text for definitions) is calculated in 100-year
1080 overlapping epochs 10 years apart. The probability density function (PDF) is
1081 shown in the blue curve. Index values in epochs of OBS are shown in magenta
1082 lines (five in total). ENSO diversity indices, including I_{sea} , $I_{cp/ep}$, $I_{e/w}$, are
1083 shown in the top rows for El Niño and La Niña. EN-LN asymmetry indices,
1084 including I_{amp} , I_{dur} , I_{tra} are shown in the bottom row. For each panel, the
1085 percentage of epochs in the EMR satisfying the specified index range is shown.
1086 Taking panel b for example, OBS have more EPEN than CPEN ($I_{cp/ep} < 1$).
1087 Among 391 100-year epochs in EMR, 91% of epochs have more EPEN than
1088 CPEN. 56
- 1089 8 ENSO behaviors in OBS, a linear EMR (EMR-L), a nonlinear EMR and
1090 GFDL CM2.1. (a-d) are decimal logarithm of the bivariate probability density
1091 function (2dPDF) in PC1-PC2 space. (e-p) distribution is calculated in 100-
1092 year overlapping epochs 10 years apart. (e-g) are I_{amp} , I_{dur} , I_{tra} , in which
1093 observation is in magenta, EMR-L in solid black curve, EMR in dash-dot
1094 curve, GFDL CM2.1 in dashed curve. (h, k, n) are I_{sea} , (i, l, o) are $I_{cp/ep}$, (j,
1095 m, p) are $I_{e/w}$. El Niño in red and La Niña in blue. 57

1096 9 Niño-3.4 SST climatology and anomaly in the 20th century (20C, historical
1097 run, 1900-1999) and the 21st century (21C, RCP8.5 scenario, 2000-2099) in
1098 37 CMIP5 models that participated in IPCC AR5. (a) The 20-year running
1099 average Niño-3.4 SST with OBS in black, CMIP5 models in gray. (b) The
1100 mean SST in the 20C and 21C for each CMIP5 model. The models are sorted
1101 according to the 100-year averaged Niño-3.4 SST in the 20C runs. The black
1102 vertical line marks the 20C OBS value. Multi-model mean (MMM) is shown
1103 at the top, 20C in orange and 21C in blue. Pre-industrial control simulations
1104 of each model are divided into 100-year sliding epochs to calculate the 100-
1105 year averaged SST and the 2.5-97.5 percentile of the distribution are shown as
1106 gray horizontal lines. The number of models with decreased/ increased change
1107 is indicated in a number with $< / >$. Number in the brackets is the count
1108 for significant changes out of the range given by the control run. (c) Niño 3.4
1109 SST anomaly time series from 1900 to 2099. (d) The standard deviation of
1110 SSTA in pre-industrial, 20C and 21C runs. In addition here the 21C results
1111 with an increased change are shown filled and those with decreased values
1112 are unfilled. Meanwhile, the 2.5, 50, 97.5 percentile range estimated from the
1113 distribution in the EMR simulation is shown in brown line at the bottom.

58

- 1114 10 Annual cycle and ENSO seasonality change from 20th century (20C) to 21st
1115 century (21C) in 37 CMIP5 models: (a) The 20C seasonal cycle of Niño-3.4
1116 SST, with OBS in black, CMIP5 models in gray. The horizontal axis shows the
1117 calendar month. (b) The Niño-3.4 SST difference between the March-August
1118 (summer half year) average and the September-February (winter half year)
1119 average. The models are sorted according to the 20C value of this difference.
1120 (c) 20C occurrence probability of El Niño P_{EN} , (e) 20C occurrence probability
1121 of La Niña P_{LN} . (d) seasonality index of El Niño ($I_{season-EN}$) defined using
1122 summer half year averaged P_{EN} divided by winter half year averaged P_{EN} ,
1123 thus $I_{season-EN} < 1$ indicates El Niño in a given 100-year epoch prefers winter
1124 phase locking. (f) $I_{season-LN}$, defined the same way, but for La Niña event. 59
- 1125 11 ENSO peaking location and propagation direction in the 20th century (20C)
1126 and 21st century (21C) in 37 CMIP5 models. (a) The location diversity index
1127 $I_{cp/ep}$ for EN, defined as P_{CPEN} divided by P_{EPEN} . (c) $I_{cp/ep}$ for LN. $I_{cp/ep} > 1$
1128 indicates El Niños or La Niñas preferentially peak in CP. (b) The propagation
1129 diversity index $I_{e/w}$ for EN. (d) $I_{e/w}$ for La Niña. $I_{e/w} > 0$ indicates El Niños
1130 or La Niñas prefer eastward propagation. 60
- 1131 12 EN-LN asymmetry in the 20th century (20C) and 21st century (21C) in
1132 37 CMIP5 models. (a) the skewness of Niño-3.4 SSTA. (b) The amplitude
1133 asymmetry index (I_{amp}) defined as EN amplitude divided by LN amplitude.
1134 $I_{amp} > 1$ indicates El Niños have larger amplitude than La Niñas in a given
1135 100-year epoch. (c) the duration asymmetry index (I_{dur}). $I_{dur} < 1$ indicates
1136 La Niñas are more persistent than El Niños. (d) the transition asymmetry
1137 index (I_{tra}). $I_{tra} > 0$ indicates El Niños are quickly followed by La Niñas but
1138 not vice versa. 61

1139 13 Summary of ENSO behaviors in the 20th century and the 21st century using
1140 all 37 CMIP5 models (a37, blue), overall best 10 models (b10, yellow) and
1141 best 10 models for individual aspect of behavior (c10, red). The 20th cen-
1142 tury observation (obs, black) results are shown as a reference. A 4000-year
1143 stochastically forced simulation of EMR model fit from the observation pro-
1144 vides the natural variation range (the 2.5-97.5 percentile range is shown). In
1145 each panel, a pair of numbers indicate the degree of model agreement. The
1146 left one is the number of models showing a decrease from the 20th century
1147 to the 21st century while the right one is the number of models showing an
1148 increase. 62

1149 14 ENSO behaviors biases in CMIP5 models: (a) shows the Niño-3.4 SST clima-
1150 tology in each model for the pre-industrial period (PI) (gray circle), the 20th
1151 century (20C) (orange circle) to the 21st century (21C) (blue triangle). 20C
1152 observation is shown in asterisk. (b-1) shows each individual ENSO behavior
1153 varies at different mean states. Change from 20C MMM to 21C MMM is
1154 shown in black line. 2.5-97.5 percentile of natural variation range by EMR is
1155 shown in dashed line. 63

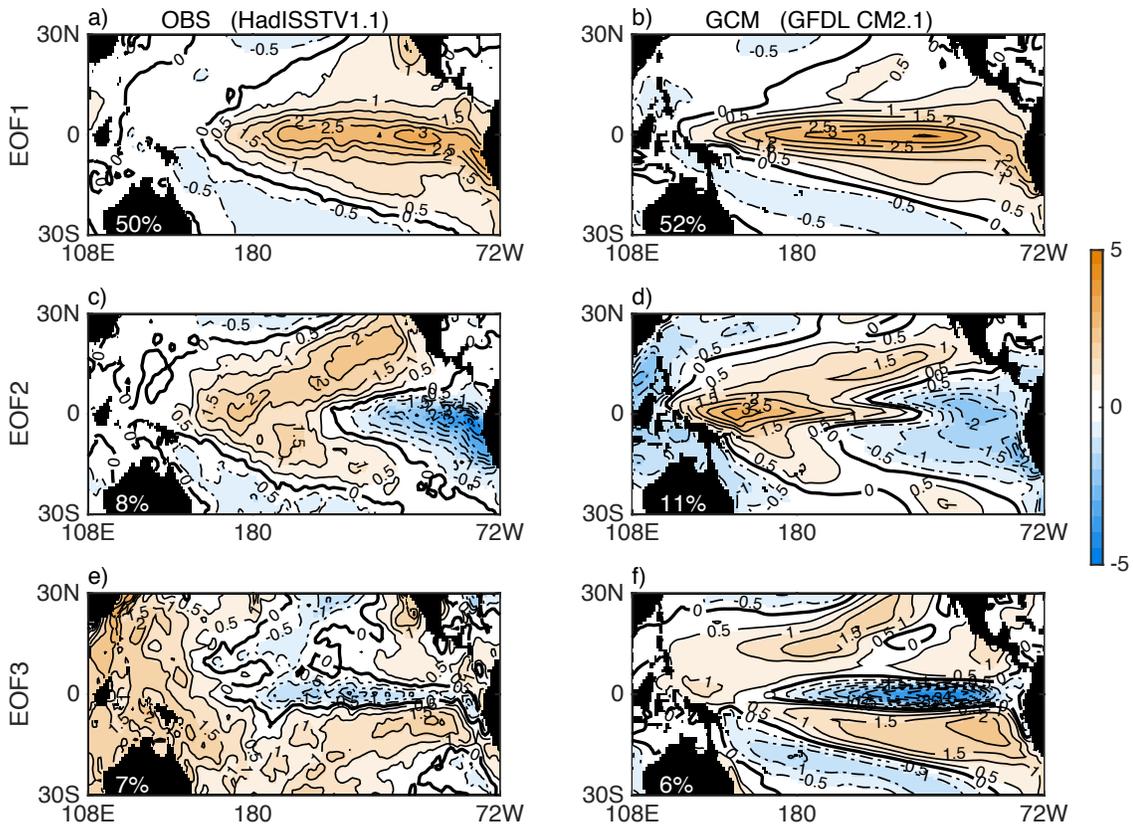


FIG. 1. The first three normalized EOF patterns of Tropical Pacific SSTA in OBS (1870-present monthly HadISST v1.1) (a, c, e) and GCM (4000-year monthly GFDL CM2.1 pre-industrial control run) (b, d, f). Positive/negative values are shown in solid/dashed contours. The zero value is highlighted in the thick solid contour.

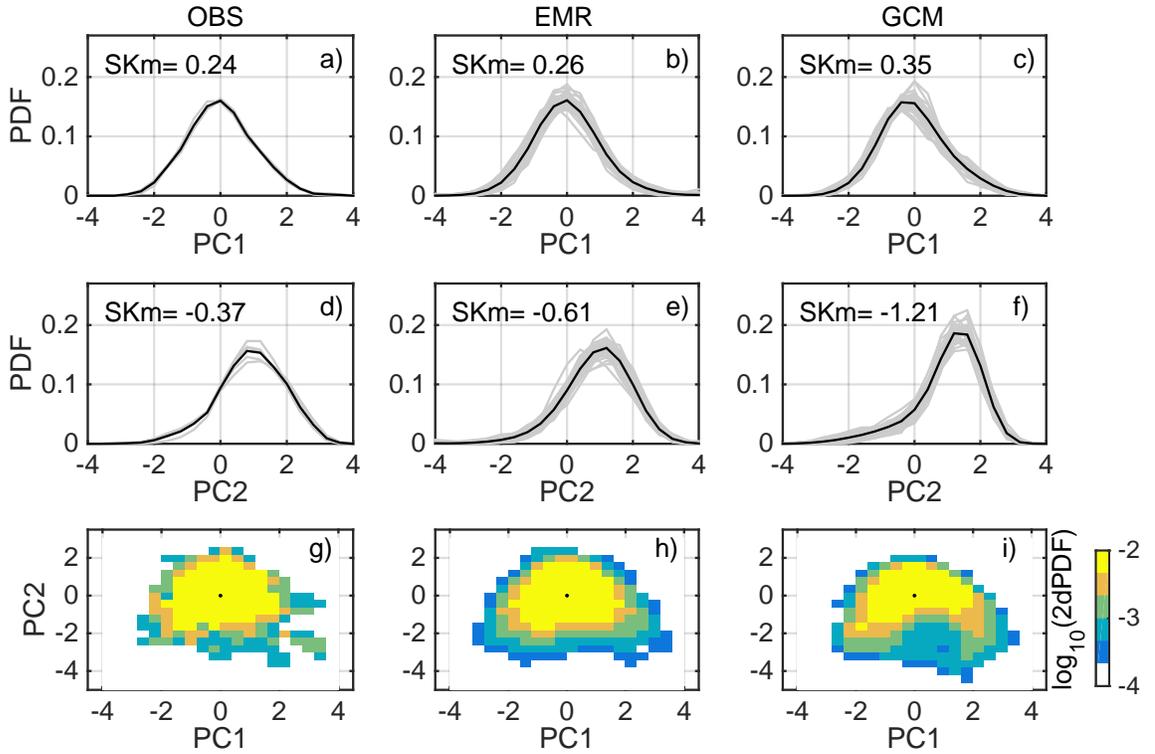


FIG. 2. Simulation evaluation: (a, d, g) for OBS, (b, e, h) for 4000 year EMR and (c, f, i) for 4000-year GCM. (a, b, c) for Probability density function (PDF) of PC1. (d, e, f) for PDF of PC2. (g, h, i) for decimal logarithm of the bivariate probability density function (2dPDF) in PC1-PC2 space. In PDF panels, OBS is divided into 5 overlapping 100-year epochs with 10 years apart. EMR and GCM are both divided into 40 non-overlapping 100-year epochs. PDF and according skewness are calculated in epochs and the average is shown.

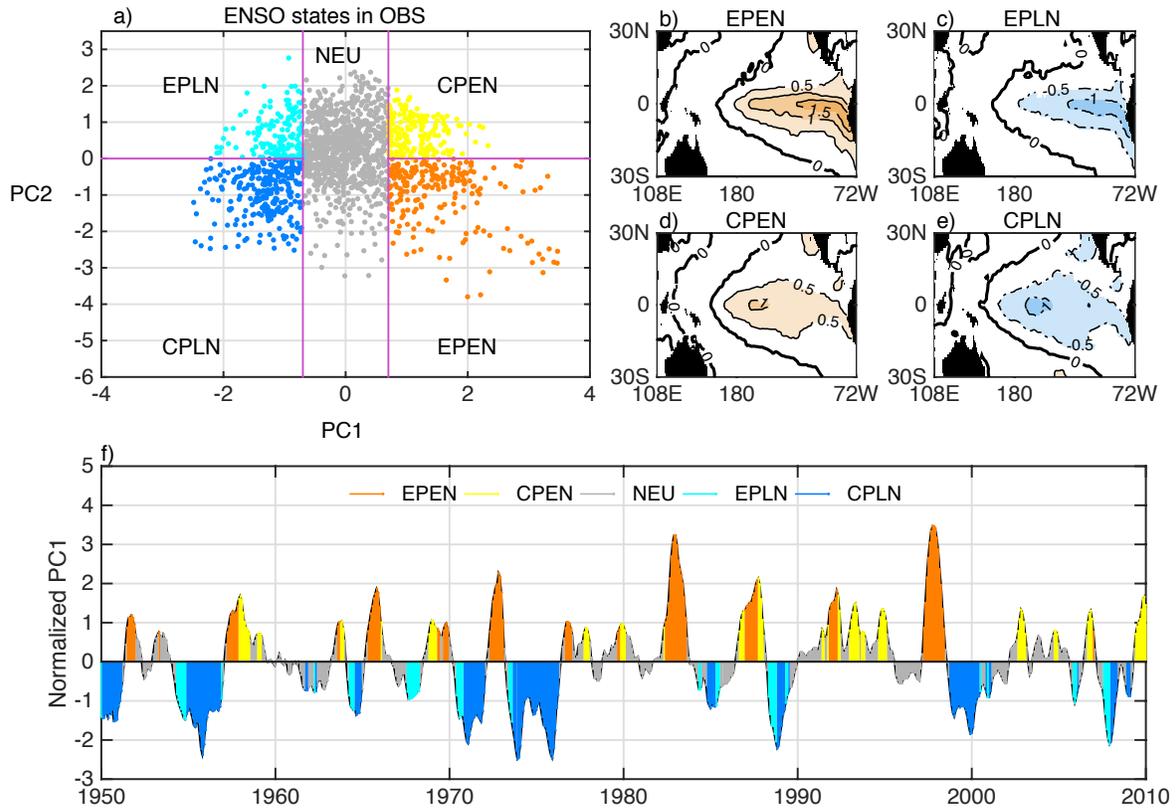


FIG. 3. Definition of 5 ENSO states (a) (Eastern/Central Pacific El Niño/La Niña and neutral patterns, denoted as EPEN/ CPEN/ EPLN/ CPLN/ NEU) are shown using smoothed monthly PCs from 1870-present HadISST v1.1. ± 0.7 s.d. (PC1), where s.d. is one standard deviation, is used to distinguish EN/LN from NEU. Zero line of PC2 is used to distinguish EP/CP states. Each state is assigned a color code for further analysis. (b-e) the SSTA patterns averaged for eastern/central Pacific ENSO states. (f) normalized PC1 (1950-2010) with each ENSO states color-coded.

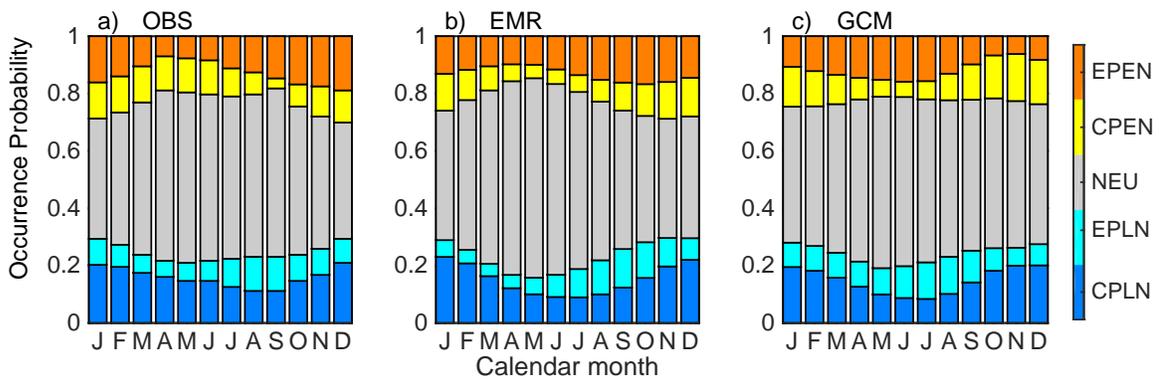


FIG. 4. Monthly occurrence probabilities for 5 ENSO state in OBS (a), EMR (b) and GFDL CM2.1 GCM (c) are shown. Stacked bars along the vertical coordinate are the occurrence probabilities of each color-coded state. The horizontal coordinate is the calendar month from Jan to Dec. Full year data is used thus the sample size for each calendar month is equal. In panel a, higher probability of EN/LN in winter months than summer months indicates observed ENSO’s winter phase locking. Higher probability of EPEN over CPEN indicates El Niño favor peaking in EP, and higher probability of CPLN over EPLN indicates LN favor peaking in CP.

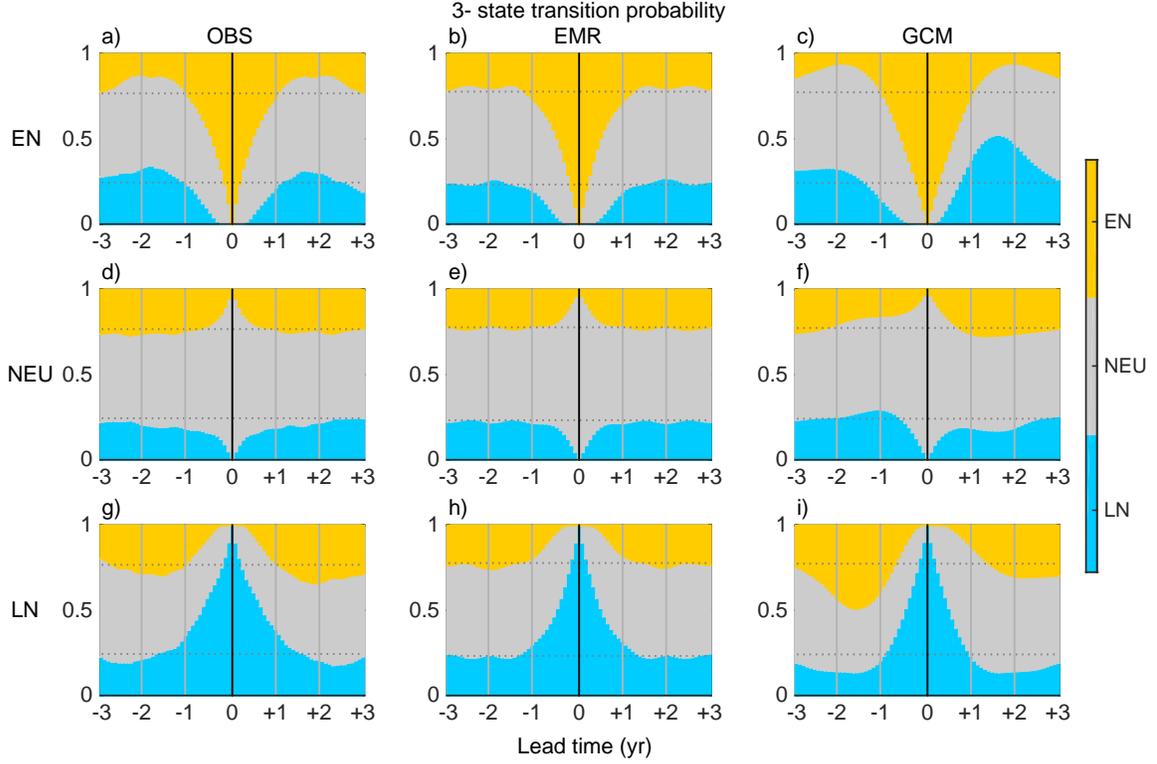


FIG. 5. State transition probabilities for EN (1st row), NEU (2nd row) and LN (3rd row) in OBS (first column), EMR (2nd column) and GFDL CM2.1 GCM (3rd column) are shown. The horizontal coordinate represents the transition from the past (-3 years) to the future (+3 years) in monthly intervals, with zero indicating the current state. Taking GCM EN transition (panel c) for example, bars along the vertical coordinate at +1 year (+12 months) represent: the self-transition probability $P_{EN(t+\tau)|EN(t)}$ (upper bar), the $P_{NEU(t+\tau)|EN(t)}$ (middle bar) and the opposite-sign transition $P_{LN(t+\tau)|EN(t)}$ (lower bar). The decaying of $P_{EN(t+\tau)|EN(t)}$ as function of lead time indicates EN's duration. The discrepancy between $P_{LN(t-\tau)|EN(t)}$ (lead<0 side) and $P_{LN(t+\tau)|EN(t)}$ (lead>0 side) indicates the EN-LN asymmetry in transition. The transition probabilities generally converge to the climatology, i.e., the nonseasonal occurrence probability of each state (dotted line in each panel).

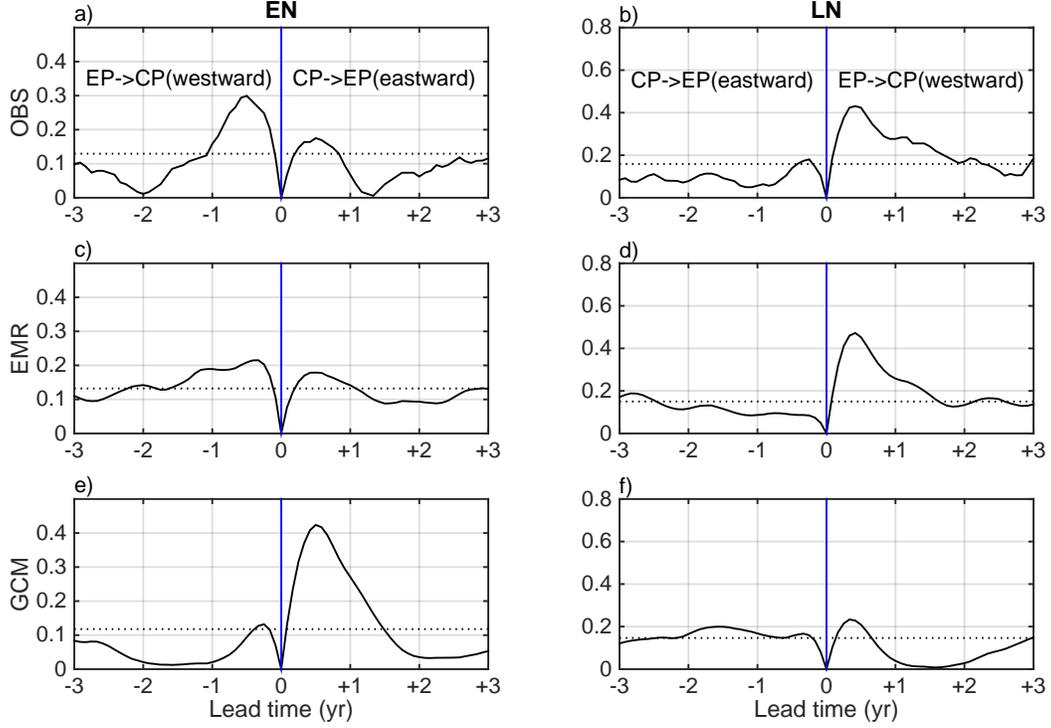


FIG. 6. Zonal propagation asymmetry in OBS(a,b), EMR(c,d) and GCM(e,f). (a,c,e): transition probabilities conditioned on CPEN state $P_{EPEN(t-\tau)|CPEN(t)}$ (negative lead) and $P_{EPEN(t+\tau)|CPEN(t)}$ (positive lead), which generally converge to the occurrence probability P_{EPEN} (dotted line) toward a large lead time. (b,d,f): transition probabilities conditioned on EPLN state $P_{CPLN(t-\tau)|EPLN(t)}$ (negative lead) and $P_{CPLN(t+\tau)|EPLN(t)}$ (positive lead), which generally converge to the occurrence probability P_{CPLN} (dotted line).

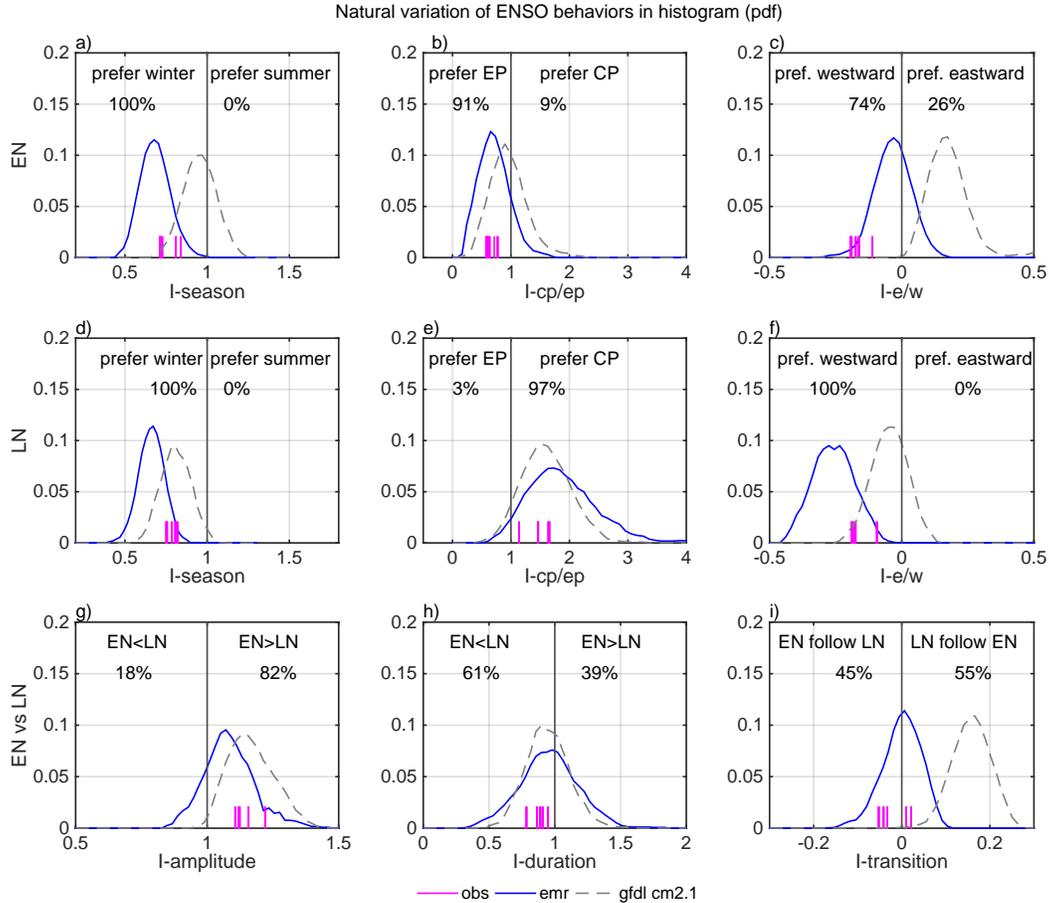


FIG. 7. Variation of ENSO behaviors in 4000-year EMR and GFDL CM2.1 GCM simulation. Each index (see text for definitions) is calculated in 100-year overlapping epochs 10 years apart. The probability density function (PDF) is shown in the blue curve. Index values in epochs of OBS are shown in magenta lines (five in total). ENSO diversity indices, including I_{sea} , $I_{cp/ep}$, $I_{e/w}$, are shown in the top rows for El Niño and La Niña. EN-LN asymmetry indices, including I_{amp} , I_{dur} , I_{tra} are shown in the bottom row. For each panel, the percentage of epochs in the EMR satisfying the specified index range is shown. Taking panel b for example, OBS have more EPEN than CPEN ($I_{cp/ep} < 1$). Among 391 100-year epochs in EMR, 91% of epochs have more EPEN than CPEN.

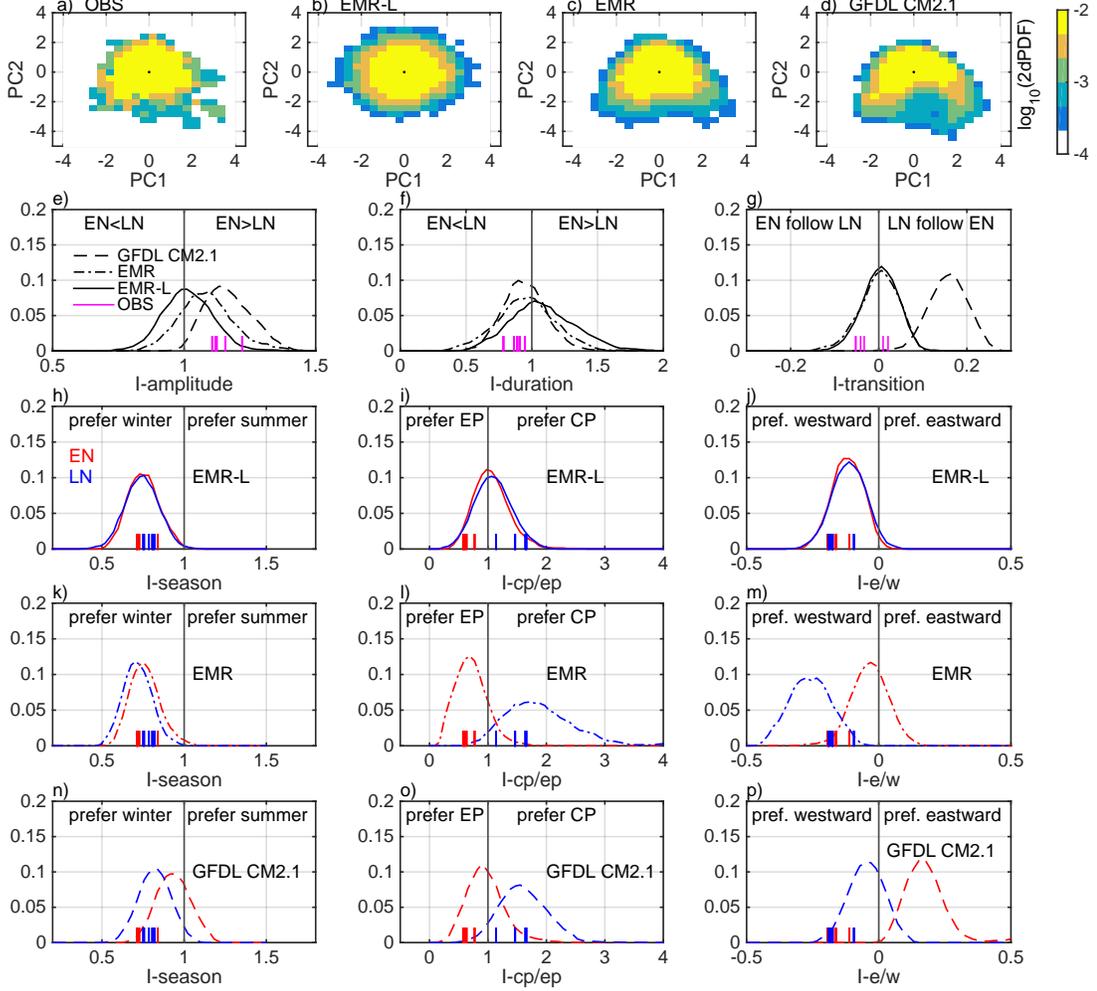


FIG. 8. ENSO behaviors in OBS, a linear EMR (EMR-L), a nonlinear EMR and GFDL CM2.1. (a-d) are decimal logarithm of the bivariate probability density function (2dPDF) in PC1-PC2 space. (e-p) distribution is calculated in 100-year overlapping epochs 10 years apart. (e-g) are I_{amp} , I_{dur} , I_{tra} , in which observation is in magenta, EMR-L in solid black curve, EMR in dash-dot curve, GFDL CM2.1 in dashed curve. (h, k, n) are I_{sea} , (i, l, o) are $I_{cp/ep}$, (j, m, p) are $I_{e/w}$. El Niño in red and La Niña in blue.

Tropical Pacific SST (N3.4) climatology and anomaly change from 20C to 21C in CMIP5

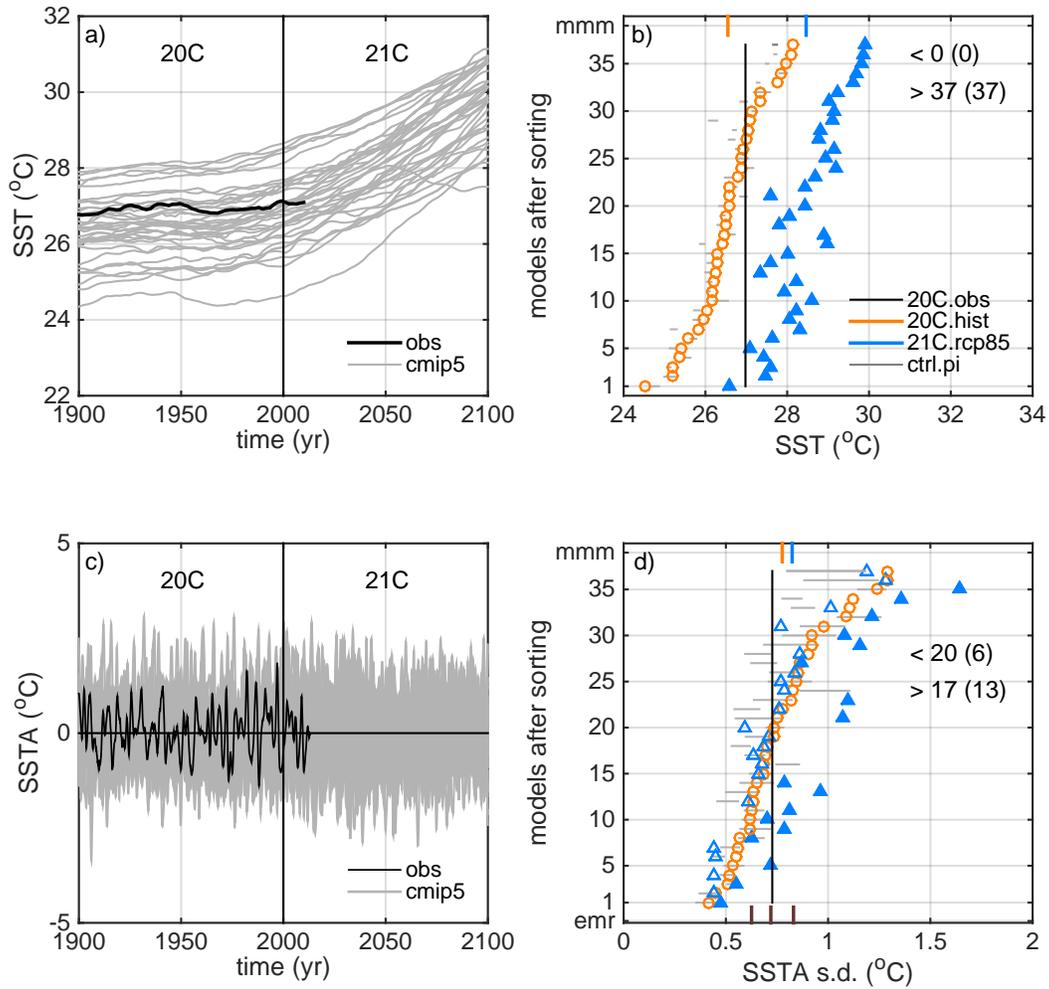


FIG. 9. Niño-3.4 SST climatology and anomaly in the 20th century (20C, historical run, 1900-1999) and the 21st century (21C, RCP8.5 scenario, 2000-2099) in 37 CMIP5 models that participated in IPCC AR5. (a) The 20-year running average Niño-3.4 SST with OBS in black, CMIP5 models in gray. (b) The mean SST in the 20C and 21C for each CMIP5 model. The models are sorted according to the 100-year averaged Niño-3.4 SST in the 20C runs. The black vertical line marks the 20C OBS value. Multi-model mean (MMM) is shown at the top, 20C in orange and 21C in blue. Pre-industrial control simulations of each model are divided into 100-year sliding epochs to calculate the 100-year averaged SST and the 2.5-97.5 percentile of the distribution are shown as gray horizontal lines. The number of models with decreased/ increased change is indicated in a number with < / >. Number in the brackets is the count for significant changes out of the range given by the control run. (c) Niño 3.4 SST anomaly time series from 1900 to 2099. (d) The standard deviation of SSTA in pre-industrial, 20C and 21C runs. In addition here the 21C results with an increased change are shown filled and those with decreased values are unfilled. Meanwhile, the 2.5, 50, 97.5 percentile range estimated from the distribution in the EMR simulation is shown in brown line at the bottom.

SST (N3.4) annual cycle and anomaly seasonality change from 20C to 21C in CMIP5

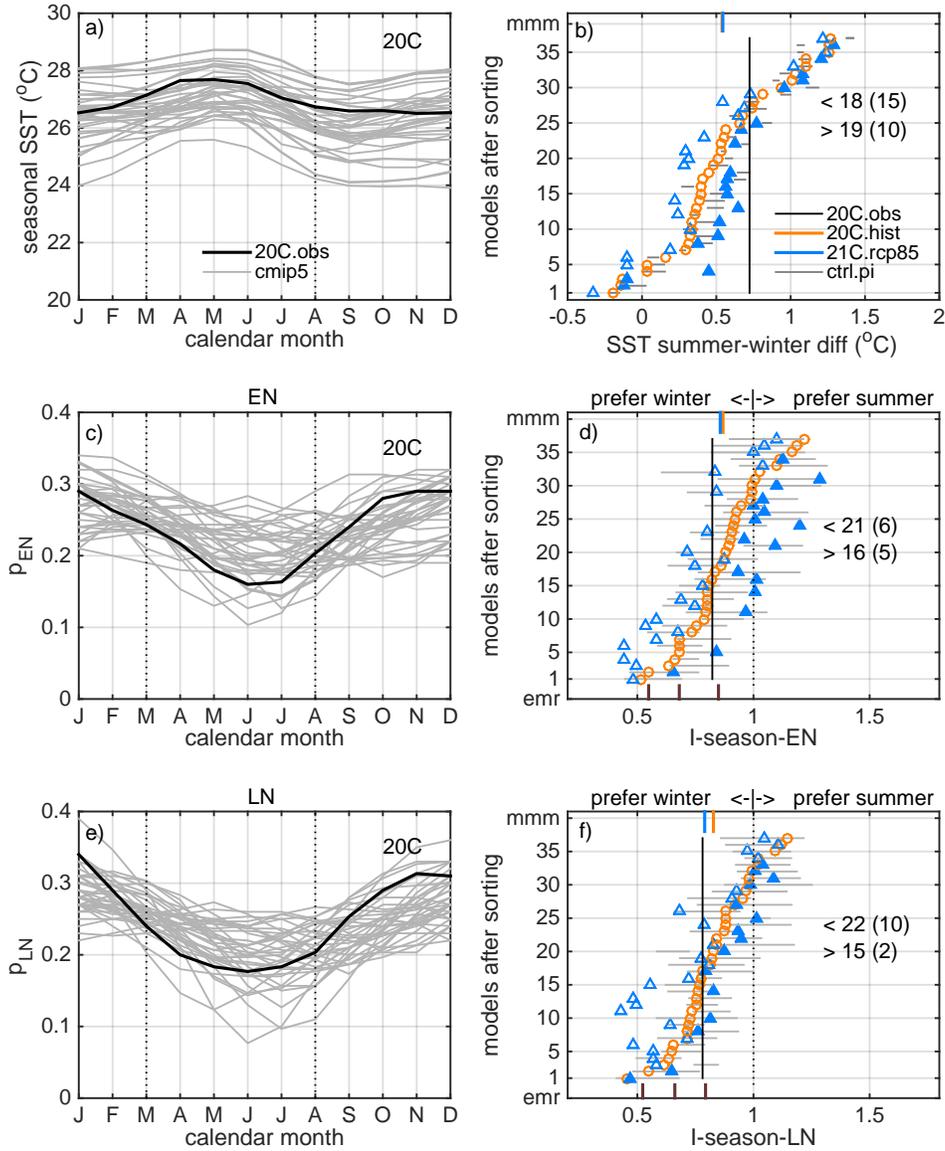


FIG. 10. Annual cycle and ENSO seasonality change from 20th century (20C) to 21st century (21C) in 37 CMIP5 models: (a) The 20C seasonal cycle of Niño-3.4 SST, with OBS in black, CMIP5 models in gray. The horizontal axis shows the calendar month. (b) The Niño-3.4 SST difference between the March-August (summer half year) average and the September-February (winter half year) average. The models are sorted according to the 20C value of this difference. (c) 20C occurrence probability of El Niño P_{EN} , (e) 20C occurrence probability of La Niña P_{LN} . (d) seasonality index of El Niño ($I_{season-EN}$) defined using summer half year averaged P_{EN} divided by winter half year averaged P_{EN} , thus $I_{season-EN} < 1$ indicates El Niño in a given 100-year epoch prefers winter phase locking. (f) $I_{season-LN}$, defined the same way, but for La Niña event.

ENSO location and propagation direction change from 20C to 21C in CMIP5

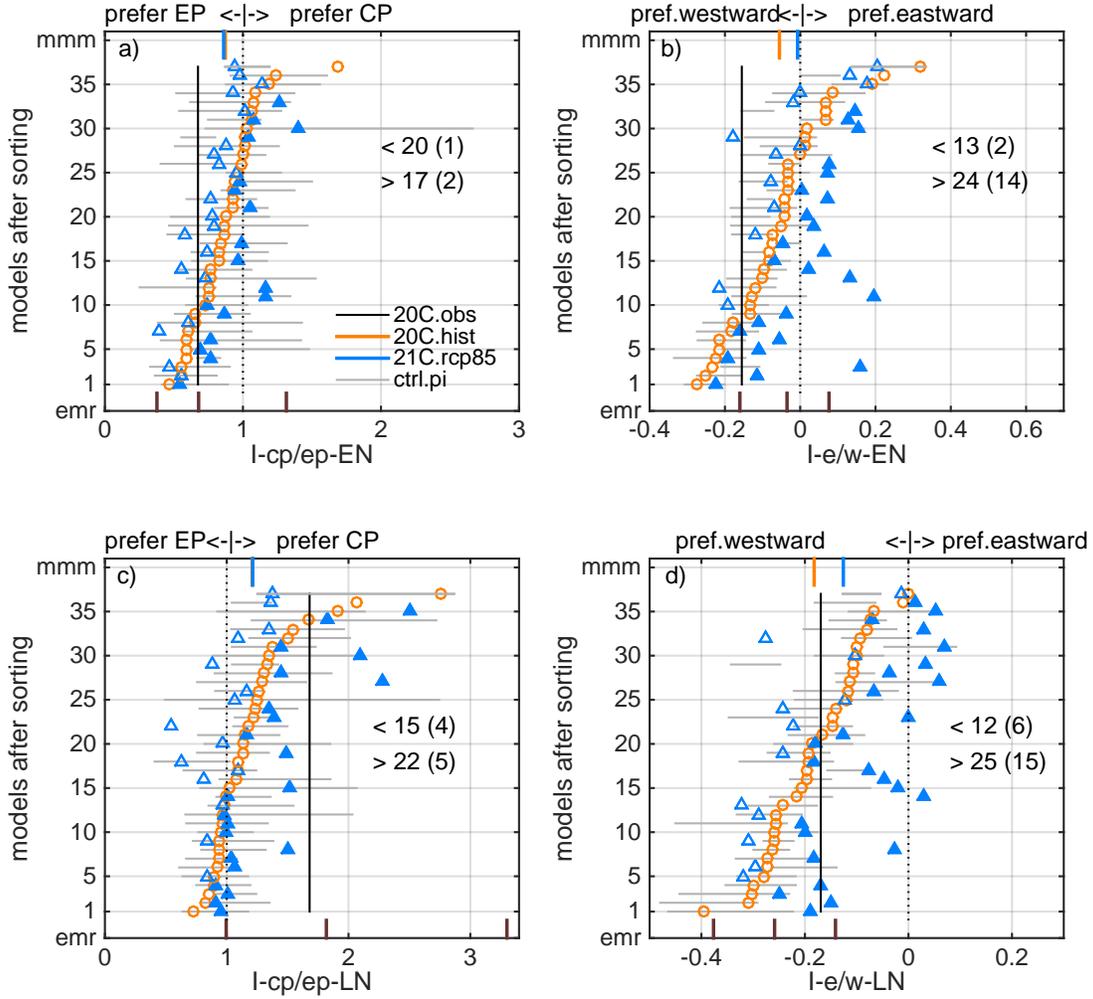


FIG. 11. ENSO peaking location and propagation direction in the 20th century (20C) and 21st century (21C) in 37 CMIP5 models. (a) The location diversity index $I_{cp/ep}$ for EN, defined as P_{CPEN} divided by P_{EPEN} . (c) $I_{cp/ep}$ for LN. $I_{cp/ep} > 1$ indicates El Niños or La Niñas preferentially peak in CP. (b) The propagation diversity index $I_{e/w}$ for EN. (d) $I_{e/w}$ for La Niña. $I_{e/w} > 0$ indicates El Niños or La Niñas prefer eastward propagation.

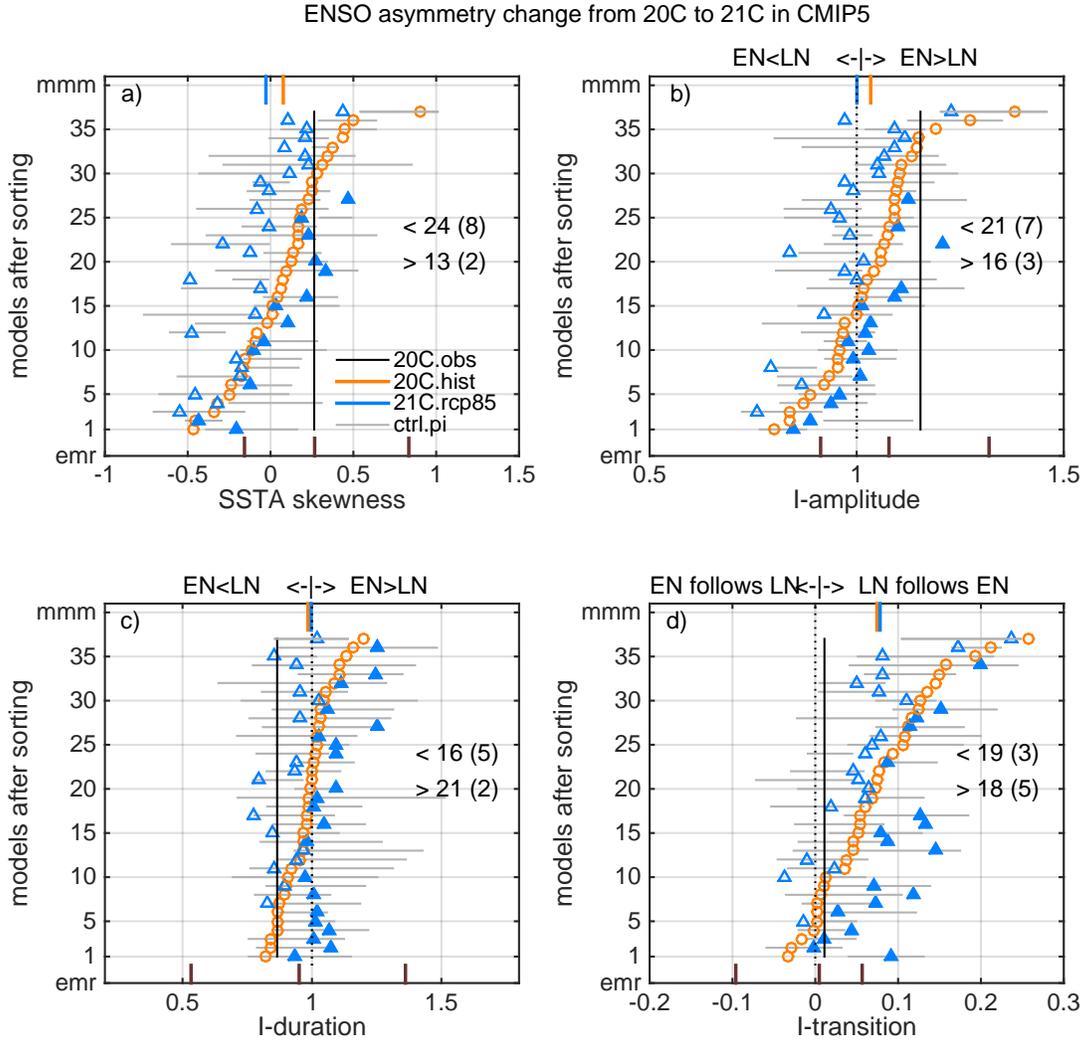


FIG. 12. EN-LN asymmetry in the 20th century (20C) and 21st century (21C) in 37 CMIP5 models. (a) the skewness of Niño-3.4 SSTA. (b) The amplitude asymmetry index (I_{amp}) defined as EN amplitude divided by LN amplitude. $I_{amp} > 1$ indicates El Niños have larger amplitude than La Niñas in a given 100-year epoch. (c) the duration asymmetry index (I_{dur}). $I_{dur} < 1$ indicates La Niñas are more persistent than El Niños. (d) the transition asymmetry index (I_{tra}). $I_{tra} > 0$ indicates El Niños are quickly followed by La Niñas but not vice versa.

20C and 21C ENSO behaviors in CMIP5 models

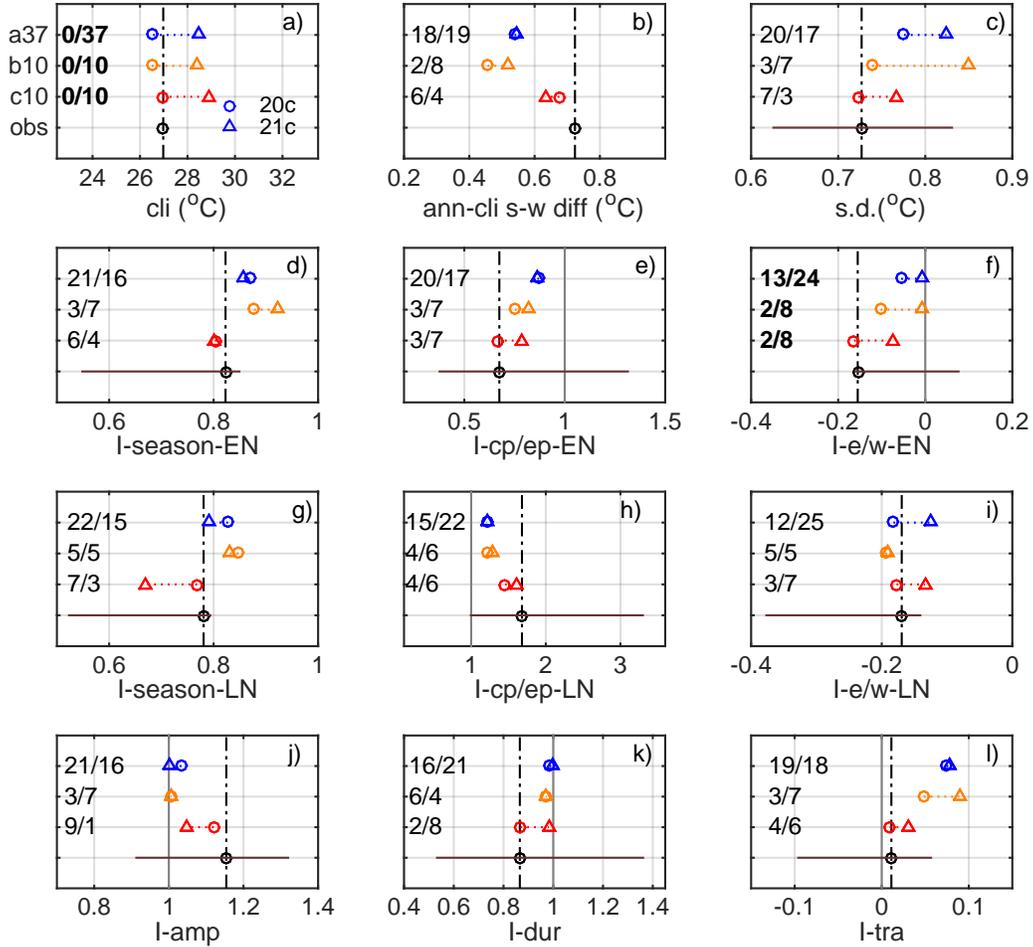


FIG. 13. Summary of ENSO behaviors in the 20th century and the 21st century using all 37 CMIP5 models (a37, blue), overall best 10 models (b10, yellow) and best 10 models for individual aspect of behavior (c10, red). The 20th century observation (obs, black) results are shown as a reference. A 4000-year stochastically forced simulation of EMR model fit from the observation provides the natural variation range (the 2.5-97.5 percentile range is shown). In each panel, a pair of numbers indicate the degree of model agreement. The left one is the number of models showing a decrease from the 20th century to the 21st century while the right one is the number of models showing an increase.

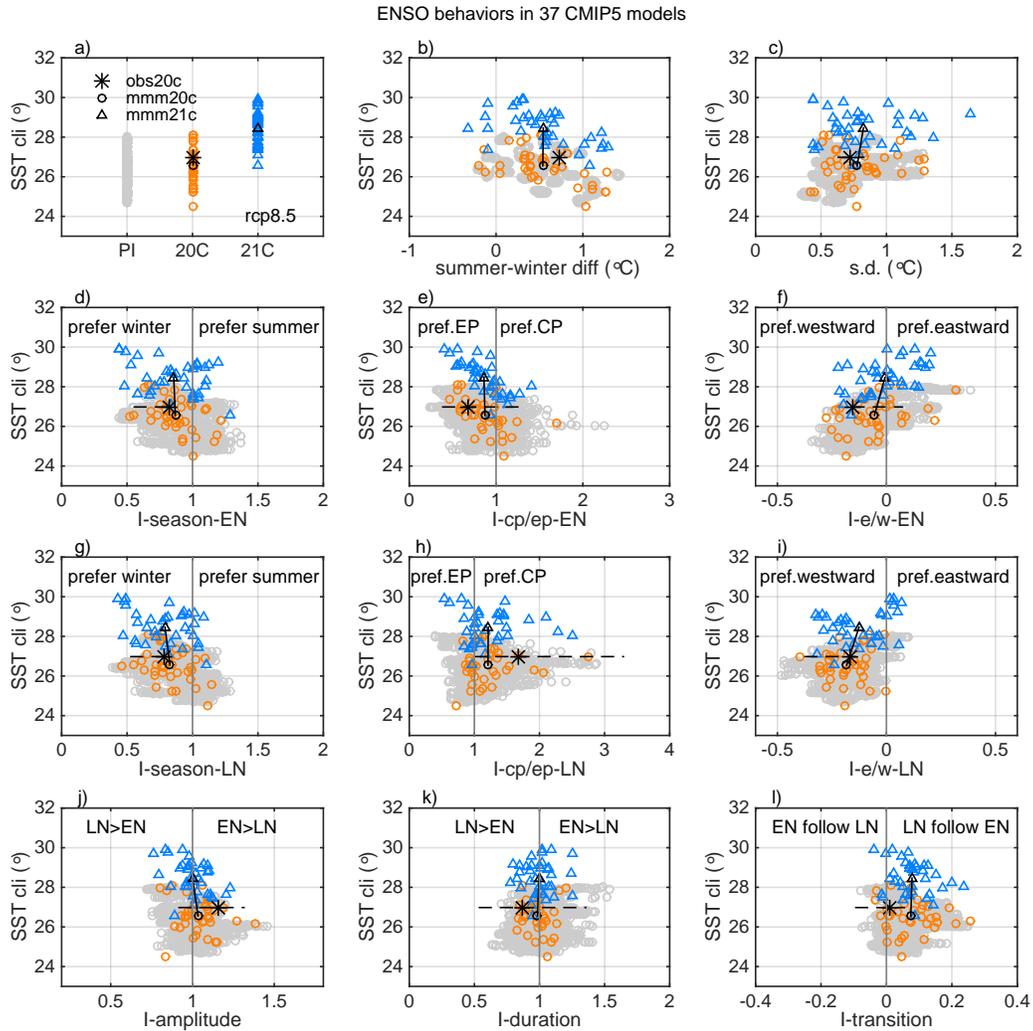


FIG. 14. ENSO behaviors biases in CMIP5 models: (a) shows the Niño-3.4 SST climatology in each model for the pre-industrial period (PI) (gray circle), the 20th century (20C) (orange circle) to the 21st century (21C) (blue triangle). 20C observation is shown in asterisk. (b-l) shows each individual ENSO behavior varies at different mean states. Change from 20C MMM to 21C MMM is shown in black line. 2.5-97.5 percentile of natural variation range by EMR is shown in dashed line.