



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-12-00575.1

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, L., Y. Yu, and D. Sun, 2013: Cloud and Water Vapor Feedbacks to the El Niño Warming: Are They Still Biased in CMIP5 Models?. *J. Climate*. doi:10.1175/JCLI-D-12-00575.1, in press.



1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

Cloud and Water Vapor Feedbacks to the El Niño Warming: Are They Still Biased in CMIP5 Models?

Lin Chen^{1,2}, Yongqiang Yu¹, De-Zheng Sun³

1 LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences, China

2 University of Chinese Academy of Sciences, China

3 Cooperative Institute for Environmental Studies/University of Colorado

&NOAA/Earth System Research Laboratory, Boulder, Colorado, USA.

Submitted to J. Climate

Aug 2, 2013

Revised on Jan. 6

Corresponding author address: Dr. Yongqiang Yu, LASG, Institute of Atmospheric

Physics, Chinese Academy of Sciences, No. 40, Hua Yan Li, Beijing 100029, China.

Email: yyq@lasg.iap.ac.cn

18

Abstract

19 Previous evaluations of model simulations of the cloud and water vapor
20 feedbacks in response to El Niño warming have singled out two common biases in the
21 Phase 3 of Coupled Model Inter-comparison Project (CMIP3) models: an
22 underestimate of the negative feedback from the shortwave radiation forcing of clouds
23 (SWCRF) and an overestimate of the positive feedback from the greenhouse effect of
24 water vapor. Here we check whether these two biases are alleviated in the Phase 5 of
25 Coupled Model Inter-comparison Project (CMIP5) models. While encouraging
26 improvements are found, particularly in the simulation of the negative SWCRF
27 feedback, the biases in the simulation of these two feedbacks remain prevalent and
28 significant. It is shown that bias in the SWCRF feedback correlates well with biases in
29 the corresponding feedbacks from precipitation, large-scale circulation and longwave
30 radiation forcing of clouds (LWCRF). By dividing CMIP5 models into two
31 categories—high score models (HSM) and low score models (LSM)—based on their
32 individual skills of simulating the SWCRF feedback, we further find that
33 ocean-atmosphere coupling generally lowers the score of the simulated feedbacks of
34 water vapor and clouds, but the LSM is more affected by the coupling than the HSM.
35 We also find that the SWCRF feedback is simulated better in the models that have a
36 more realistic zonal extent of the equatorial cold tongue, suggesting that the
37 continuing existence of an excessive cold tongue is a key factor behind the persistence
38 of the feedback biases in models.

39 **1. Introduction**

40 Cloud and water vapor are major modulators in climate system and strongly
41 influence both the global circulation and energy balance through their radiative effects
42 (Manabe and Wetherald 1967; Hartmann and Short 1980; Harrison et al. 1990; Kiehl
43 and Trenberth 1997; Soden 1997; Houghton et al. 2001; Stephens 2005). The climate
44 sensitivity to the rise of the man-made greenhouse gases in the atmosphere depends
45 critically on the feedbacks from water vapor and clouds. However, the feedbacks from
46 cloud and water vapor are considered as the largest source of uncertainty in climate
47 predictions (Bony and Dufresne 2005; Randall et al. 2007). Therefore, assessing the
48 accuracy and narrowing the uncertainties of the cloud and water vapor feedbacks in
49 current leading climate models is of obvious importance. Indeed, persistent efforts
50 have been made in this regard (Cess et al. 1990; Sun and Held 1996; Bony and
51 Dufresne 2005; Stephens 2005; Bony et al. 2006; Lin 2007; Zhu et al. 2007; Soden et
52 al. 2008; Sun et al. 2009, hereafter referred to as Sun2009).

53 Work on the evaluation of the feedbacks from water vapor and clouds have
54 employed two main methodologies. The first one involves examining the response of
55 cloud and water vapor in different models to a prescribed or model predicted global
56 warming (Cess et al. 1990, 1996; Soden and Held 2006; among others). These studies
57 have revealed that the cloud feedbacks differ greatly among models while the globally
58 averaged feedback from water vapor in the models follows that of a constant relative

59 humidity model. As noted by Sun2009, however, these results have neither pointed
60 out which model has the right cloud feedback nor ruled out the possibility that all the
61 models have a similar bias in water vapor feedback. In addition, a uniform increase in
62 the SST is often used to examine the response of cloud and water vapor to global
63 warming in models. But there is uncertainty with the spatial pattern of global warming.
64 For one thing, the surface temperature is not likely to increase uniformly across the
65 globe (Xie et al. 2010).

66 Another methodology used to examine the cloud and water vapor feedbacks in
67 climate models involves examining the response of cloud and water vapor to SST
68 changes on the time scales of El Niño-Southern Oscillation (ENSO) (Sun et al. 2003,
69 2006, 2009; Lloyd et al. 2009, 2011, 2012; among others). By comparing the response
70 of cloud and water vapor to the ENSO forcing in nature with that in Atmospheric
71 Model Inter-comparison Project (AMIP) simulations by some leading climate models,
72 Sun et al. (2006) revealed two common biases in the AMIP runs of models: 1) an
73 underestimate of the strength of the negative shortwave radiation forcing of clouds
74 (SWCRF) feedback and 2) an overestimate of the positive feedback from the
75 greenhouse effect of water vapor. Extending the same analysis to the fully coupled
76 simulations of these models as well as other coupled models in CMIP3, Sun2009
77 found that these two biases persist. Studies by Lloyd et al. (2009, 2011, 2012) using
78 CMIP3 outputs show that bias in the SWCRF feedback is the main source of model
79 uncertainty and the biases in the cloud response to dynamical changes dominate the

80 modeled SWCRF feedback. The main purpose of present study is to update the results
81 of Sun2009 using CMIP5 models, and specifically explore whether the newer models
82 have improvements in simulating the two aforementioned feedbacks and the
83 corresponding origins for the improvements.

84 The remainder of the paper is organized as follows. After describing the data and
85 methods used in this study in Section 2, we show the changes and improvements from
86 CMIP3 to CMIP5 in Section 3, particularly in reference to the key results from
87 Sun2009. We then show the correlation among biases in different feedbacks and
88 explore the causes for the biases in section 4. Finally, a summary and discussion is
89 given in Section 5.

90 **2. Data and Methodology**

91 *a. Model, Observational and Reanalysis Datasets*

92 The model data are derived from CMIP5 multimodel dataset (Talor et al. 2012).
93 Outputs from 18 CMIP5 models participating in the Intergovernmental Panel for
94 Climate Change Fifth Assessment Report (IPCC AR5) are used in this study. Their
95 acronyms and the names of the institution that have provided them to IPCC AR5 are
96 listed in Table 1. The reason we chose these 18 models for our initial focus is that they
97 are the models whose outputs including both historical runs and AMIP runs had been
98 submitted to the Program for Climate Model Diagnosis and Intercomparison (PCMDI)

99 at the time of our analysis. The historical (AMIP) runs of these models cover the
100 period 1850-2005 (1979-2005).

101 The radiative fluxes supplied by the International Satellite Cloud Climatology
102 Project (ISCCP; Zhang et al. 2004) and the precipitation derived from the Global
103 Precipitation Climatology Project (GPCP; Huffman et al. 2009) data are used in this
104 study. In addition, the datasets including 1) Hadley Centre Sea Ice and Sea Surface
105 Temperature dataset (HadISST; Rayner et al. 2003) and 2) the US National Centers
106 for Environmental Prediction-Department of Energy reanalysis data (NCEP-DOE,
107 Kanamitsu et al. 2002) are also employed. As the surface heat fluxes of NCEP2 are
108 subject to larger errors in the eastern Pacific (Cronin et al. 2006), to provide a lower
109 bound to the uncertainty, the turbulent heat fluxes from OAFflux (Yu et al. 2007) are
110 employed to calculate the observed feedbacks associated with the surface heat flux. In
111 the present analysis, all the monthly datasets are analyzed during the period from
112 1984 to 2005 and have been interpolated to a common horizontal grid of $2.5^\circ \times 2.5^\circ$.

113 *b. Analysis method*

114 In order to have a straight comparison with the results shown in Sun2009 and
115 showcase the improvements from CMIP3 to CMIP5 in the key feedbacks underscored
116 in Sun2009, the primary methodology used here is the same as that employed in
117 Sun2009: we obtain SSTA averaged over the equatorial Pacific (5°S - 5°N ,
118 150°E - 110°W) by removing the monthly annual cycle of SST, and then calculate the

119 values of feedbacks during ENSO through a linear regression using the interannual
120 variations of the SSTA and the corresponding variables. To obtain a more complete
121 characterization of the model deficiencies and identify common factors responsible
122 for the discrepancies in physical processes, we further introduce a skill score as an
123 additional measure of the model deficiencies. In reference to the previous skill score
124 formula for model performance (Taylor 2001; Hirota et al. 2011), the skill score in the
125 present study is given by the following formula:

$$126 \quad Skill\ Score = \frac{(1+R)^2}{\left(SDR + \frac{1}{SDR}\right)^2} \quad (1)$$

127 where R is a pattern correlation between the observation and the models, and SDR is
128 the ratio of spatial standard deviations of the models against that of the observation. It
129 is clear from equation (1) that both the spatial distribution and magnitude have been
130 considered in this expression while comparing model results with observations.

131 **3. Changes and Improvements from CMIP3 to CMIP5**

132 *a. Comparison of SWCRF feedback between CMIP3 and CMIP5*

133 The feedbacks from cloud and water vapor to El Niño warming in CMIP5
134 models are presented in Table 2, that is the same as the Table 1 and Table 2 in
135 Sun2009 except that the results are from 18 CMIP5 models, while the results in
136 Sun2009 are based on 12 CMIP3 models. The ensemble model mean values for high

137 score models and low score models are also listed in Table 2 for later discussion.

138 As shown in column 1 of Table 2, nearly all the CMIP5 models still overestimate
139 the positive feedback from the greenhouse effect of water vapor ($\frac{\partial(G_w)}{\partial T}$). The
140 magnitude of the bias is similar to that was found in CMIP3 models. Although there
141 are variations among models, CMIP5 models on average do not show significant
142 improvements in their representation of the water vapor feedback. In contrast, the
143 underestimate of the strength of the negative SWCRF feedback ($\frac{\partial(C_s)}{\partial T}$), emphasized
144 in Sun2009, is notably alleviated in some of CMIP5 models. A striking improvement
145 is noted in the NCAR and IAP models. Their CMIP5 version (CCSM4 and
146 FGOALS-g2) now have a negative SWCRF feedback ($-12.73 \text{ W m}^{-2} \text{ K}^{-1}$ and -14.46
147 $\text{W m}^{-2} \text{ K}^{-1}$ in the historical runs) that are very close to the observed value (-13.33 W
148 $\text{m}^{-2} \text{ K}^{-1}$), recalling that the values are respectively $-3.69 \text{ W m}^{-2} \text{ K}^{-1}$ and -2.25 W m^{-2}
149 K^{-1} in the coupled runs of CMIP3 version (CCSM3 and FGOALS-g1.0) (see Table 2).
150 In addition, the SWCRF feedback in the historical runs of CCSM4, FGOALS-g2,
151 GISS-E2-R and NorESM1-M, agree well with observations over the equatorial Pacific.
152 This is also an encouraging improvement over CMIP3 models, in which nearly all of
153 them have notable biases in simulating the SWCRF feedback (see the italicization in
154 Table 2). However, the negative SWCRF feedback is still underestimated in the
155 historical runs of more than half of CMIP5 models examined here. The numbers in the
156 parenthesis in column 4 of Table 2 are the SWCRF feedback ($\frac{\partial(C_s)}{\partial T}$) estimated from
157 the corresponding AMIP runs. It is evident that most CMIP5 models perform better in

158 simulating the SWCRF feedback in the AMIP runs than those in the historical runs.

159 As listed in the second column of Table 2, the positive feedback of longwave
160 radiation forcing of clouds (LWCRF) is also underestimated in most models. But the
161 magnitude of the biases from LWCRF does not appear to always compensate those in
162 the feedback from water vapor and SWCRF, leading to diverse results among models
163 in the feedback from the total greenhouse effect of water vapor and clouds ($G_a + C_l$)
164 (column 3, Table 2) and the net clouds feedback ($C_l + C_s$) (column 5, Table 2). The
165 feedback from the atmospheric transport ($\frac{\partial(D_a)}{\partial T}$), the net atmospheric feedback
166 ($\frac{\partial(F_a)}{\partial T}$), and the feedback from the net surface heat flux into the ocean ($\frac{\partial(F_s)}{\partial T}$) are
167 also listed in Table 2. In general, they show a slight improvement from the present
168 CMIP5 models compared to the CMIP3 models (Sun2009).

169 *b. Spatial structure of the feedbacks in El Niño Warming*

170 Fig. 1 and Fig. 2 further show the spatial structure of the response of SWCRF to
171 El Niño warming in the AMIP runs and the historical runs of CMIP5 models,
172 respectively. The spatial distribution and magnitude of the response of SWCRF is
173 generally captured well in most AMIP runs (Fig. 1). In reference to the results of
174 Sun2009, the AMIP results show a slight improvement from CMIP3 to CMIP5.
175 Comparisons with the observations also reveal that the simulation of SWCRF
176 feedback in CMIP5 models performs better in the AMIP runs than those in the
177 historical runs. Relative to the AMIP runs, the corresponding historical runs are found

178 to have a larger spread among models and a greater difference between models and
179 observations (Fig. 2). The most notable discrepancies with observations are the
180 location and magnitude of the maximum response of SWCRF. First, with a few
181 exceptions (i.e., ACCESS1-0, CCSM4, BCC-CSM1-1, FGOALS-g2, GISS-E2-R and
182 NorESM1-M), the maximum response of SWCRF has an obvious westward extension
183 in other models. Second, the maximum response of SWCRF is significantly
184 underestimated in most CMIP5 models except for CCSM4, CNRM-CM5,
185 FGOALS-g2, GISS-E2-R and NorESM1-M.

186 Compared to their corresponding AMIP runs, 9 out of the 18 CMIP5 models
187 examined here (i.e., CanESM2, CSIRO-Mk3-6-0, INMCM4, IPSL-CM5A-LR,
188 FGOALS-s2.0, MIROC5, the two MPI models and MRI-CGCM3) show that the
189 maximum response of SWCRF shifts westward in the historical runs. In addition, the
190 magnitude of maximum response of the SWCRF shown in the historical runs (Fig. 2)
191 has a general weakening in the response of SWCRF compared to their respective
192 AMIP results that show a better agreement with the observations (Fig. 1). This
193 common bias exists in more than half of the 18 CMIP5 models examined herein.

194 Overall, in terms of both the structure and magnitude, the AMIP simulations of
195 response of SWCRF to the El Niño warming in CMIP5 models show a slight
196 improvement over CMIP3 models. In addition, the performance in the historical runs
197 is significantly improved in about one third of CMIP5 models relative to CMIP3
198 results (Sun2009). Nonetheless, a large spread still exists among models, and the

199 biases in the representation of the cloud and water vapor feedbacks in the ENSO cycle
200 are generally amplified in the historical runs compared to the corresponding AMIP
201 runs.

202 **4. An analysis of the origin of the biases**

203 *a. Relationship of biases among different feedbacks*

204 For further evaluating model performance and revealing the origin of biases
205 analyzed above, we have compared the skill scores of the feedbacks of SWCRF,
206 LWCRF, precipitation, vertical velocity at 500 hPa (ω_{500}) (Fig. 3). As is shown in Fig.
207 3a-c, there is a positive correlation between the inter-model variations in the simulated
208 SWCRF feedback and the inter-model variations in other simulated feedbacks—the
209 feedback of precipitation, the LWCRF feedback and the feedback of the large scale
210 circulation (e.g., ω_{500}). It is suggested that models with higher skill in simulating the
211 response of SWCRF to El Niño warming generally perform better in other physical
212 feedback processes (Fig. 3a, 3b and 3c). The apparent correlations among model
213 biases imply that all the biases might be resulted from the same fundamental origins.

214 Hence, according to the skill in simulating the SWCRF feedback in the historical
215 runs, we divide the 18 CMIP5 models into 2 categories: high score models (HSM)
216 and low score models (LSM). The HSM, identified with the higher skill scores above
217 0.5 (see the models above the dashed line in Fig. 3d), includes 9 models: CCSM4,
218 CNRM-CM5, NorESM1-M, FGOALS-g2, ACCESS1-0, BCC-CM1-1, BNU-ESM,

219 GISS-E2-R and HadGEM2-ES models (sorted by the score in the simulation of
220 SWCRF feedback), and the LSM is identified with the lower skill scores below 0.5
221 (see the models below the dashed line in Fig. 3d). It is found that the ensemble of the
222 HSM definitely performs better than the ensemble of the LSM in the historical runs
223 over the tropical region (Fig. 2t and Fig. 2u). However, the difference between the
224 HSM and the LSM in simulating the SWCRF feedback is smaller in the AMIP runs
225 (see Fig. 1t, Fig. 1u, Fig. 2t and Fig. 2u). It suggests that a better representation of the
226 SWCRF feedback in the AMIP runs cannot guarantee a better simulation in its
227 historical counterpart (as shown in Fig. 3d). As shown in the last two rows of Table 2,
228 the values of most feedback coefficients listed in Table 2 generally agree better with
229 observations in the HSM than those in the LSM, even though only variations in the
230 equatorial Pacific (5°S - 5°N , 150°E - 110°W) are used to obtain these values in Table 2.

231 *b. Origins of the biases*

232 Based on the results analyzed above, one may question what causes the biases of
233 the cloud and water vapor feedbacks in models and significant differences between
234 AMIP runs and historical runs.

235 When it comes to the improvement from CMIP3 models to the CMIP5 models in
236 the historical runs, it is tempting to attribute the biases of the feedbacks to the biases
237 of the simulation of ENSO. However, it is noted that a better simulation of ENSO
238 variability is generally not accompanied by a more realistic SWCRF feedback to
239 ENSO. Fig. 4 displays the relationship between the skill of reproducing ENSO

240 variability as measured by the ratio of the standard deviation (STD) of the simulated
241 Niño3 SSTA over the STD of the observed Niño3 SSTA, and the skill in the
242 corresponding SWCRF feedbacks in the historical runs. The red open circles and the
243 blue open squares in the figure indicate the results from the HSM and the LSM
244 respectively. The figure shows clearly that the higher skill score of SWCRF feedback
245 is not necessarily accompanied by a better simulation of ENSO. For example, the
246 CCSM4 model, which has overestimated the amplitude of ENSO, shows a well
247 simulated SWCRF feedback in its historical runs. But the IPSL-CM5A-LR model, in
248 which the ENSO amplitude is close to the observation, does not simulate the SWCRF
249 feedback well. Therefore, the bias in the simulation of ENSO is unlikely a major
250 factor behind the biases of feedbacks in the coupled GCMs.

251 The leading candidate for the feedback biases in the historical runs may be the
252 excessive cold tongue in models (Sun et al. 2006, 2009). To test this hypothesis, we
253 have examined the climatological mean tropical Pacific SST in the historical runs of
254 the HSM and the LSM (Fig. 5b and Fig. 5c). Although the spatial pattern of the mean
255 SST is similar to the observations, the cold bias in the equatorial Pacific appears in
256 both the HSM and the LSM except the eastern equatorial Pacific, and the warm bias is
257 prevalent adjacent to the Peru coast, which have been pointed out earlier and continue
258 to exist in current coupled GCMs (Mechoso et al. 1995; Davey et al. 2002; Latif et al.
259 2001; Sun et al. 2006). Note that the bias of the climatological mean SST is generally
260 smaller in the HSM than that in the LSM, especially over the central equatorial Pacific.

261 As we know, the deep convection tends to occur where SST is higher than a threshold
262 around 28°C (Graham and Barnett 1987) and the dependence of precipitation on SST
263 is highly nonlinear (see Fig. 5a in Sun2009). Therefore the cold bias of SST in central
264 equatorial Pacific may lead to a decrease in the precipitation over the central
265 equatorial Pacific and shift the deep convection westward. By comparing the response
266 of precipitation to El Niño warming in both AMIP and historical runs of two
267 categories of models (figures not shown), it is found that the response of precipitation
268 in the LSM historical runs is significantly weaker than that in observations and has a
269 westward shift, whereas the response of precipitation in the HSM historical runs is
270 slightly weaker and the location matches better with the observations. However, the
271 precipitation feedback in the corresponding AMIP runs of both the HSM and the LSM
272 are well simulated (the skill scores are 0.92 and 0.93). Recall that the AMIP runs are
273 forced by the same observed SST, it thus suggests that the mean SST bias in the
274 historical runs especially the excessive cold tongue may contribute to the
275 underestimate of the SWCRF feedback and the westward shift of the maximum
276 response of SWCRF.

277 To further test this hypothesis, we have examined the relationship between the
278 cold-tongue biases and the skill scores of the SWCRF feedback in all the historical
279 runs (Fig. 6). In Fig. 6, we use the mean SST bias over the equatorial Pacific (5°S-5°N,
280 150°E-110°W) as a measure of the cold-tongue bias. A colder SST bias in this region
281 indicates a more excessive cold tongue. Fig. 6 shows that the models with smaller

282 cold-tongue biases are generally accompanied by the higher scores of the SWCRF
283 feedback (i.e. the better simulation of the SWCRF feedback), implying that reducing
284 the cold bias of the mean SST may be beneficial for the better simulation of the
285 SWCRF feedback. Also note that most of the CMIP5 coupled models exhibit cold
286 SST biases. Consistent with Fig.5, the models in the group of HSM generally exhibit
287 smaller cold SST biases or even warm biases compared to the models in the LSM.

288 It should be noted, however, the excessive cold tongue in the models may in turn
289 be affected by the biases in the net atmospheric feedbacks. According to Sun et al.
290 (2003) (see their Fig. 6), the prevalence of the excessive cold tongue is possibly
291 related to the incorrect representation of the net atmospheric feedbacks (i.e.,
292 $\partial(F_a)/\partial T = \partial(G_a)/\partial T + \partial(C_l)/\partial T + \partial(C_s)/\partial T + \partial(D_a)/\partial T$). Fig 7 shows the observed
293 and model-simulated $\partial(F_a)/\partial T$ over the tropical Pacific. It is found that in the
294 historical runs, $\partial(F_a)/\partial T$ from both the HSM and the LSM are much weaker than
295 the observation (see Fig. 7b and Fig. 7c), and the latter is even weaker (note that this
296 feedback in the LSM is barely negative over the central equatorial Pacific). Also as
297 shown in the last second column of Table 2, $\partial(F_a)/\partial T$ over the equatorial Pacific
298 (5°S - 5°N , 150°E - 110°W) in the HSM historical runs and the LSM historical runs are
299 $-7.23 \pm 0.78 \text{ W m}^{-2} \text{ K}^{-1}$ and $-0.31 \pm 0.65 \text{ W m}^{-2} \text{ K}^{-1}$, respectively. The latter feedback
300 is much weaker than the former one. Therefore, the cold SST bias is less regulated
301 by the atmospheric process in the coupled models, leading to the formation of an
302 excessive cold tongue in these models. The excessive cold tongue then further

303 weakens the net atmospheric feedback. Such a vicious cycle may be a root cause for
304 the difficulty to make the equatorial Pacific cold tongue simulated correctly by
305 coupled GCMs.

306 Indeed, we find in our analysis that the air-sea interaction or coupling between
307 the atmosphere and ocean amplifies the errors in the atmospheric feedbacks. By
308 comparing the HSM and the LSM in the quantification of the skill scores over the
309 tropical Pacific (see Fig. 8), it is found that the simulated feedbacks in the AMIP
310 experiments are generally better than the historical experiments. Although the
311 representations of the feedbacks in the HSM AMIP runs are very close to those in the
312 LSM AMIP runs, they are very different in their corresponding historical runs. The
313 simulated feedbacks in the HSM historical runs are much better than those in the LSM
314 historical runs. Apparently, biases of the cloud and water vapor feedbacks in the LSM
315 AMIP runs are amplified more easily in their historical counterparts than those in the
316 HSM AMIP runs. This also implies that the improvements in the AMIP runs can have
317 an amplified gain in their corresponding coupled runs.

318 The persistence of the overestimated water vapor feedback in CMIP5 models is
319 linked with the bias of water vapor. As the greenhouse gases feedback is more
320 proportional to its percentage change (Shine and Sinha 1991; Schneider et al. 1999;
321 Zhang and Sun 2008), we have calculated the response of the specific humidity
322 change percentage (anomalies/mean) to El Niño warming in both the AMIP runs and
323 the historical runs of CMIP5 models in the same manner as Zhang and Sun (2006,

324 2008). We found the excessive response of water vapor in the upper troposphere in all
325 models (figures not shown).

326 **5. Summary and Discussion**

327 Previous studies have revealed two common biases in the simulation of the
328 response of the cloud and water vapor to the El Niño warming in CMIP3 models: an
329 underestimate of the negative cloud albedo feedback and an overestimate of the
330 positive water vapor feedback. To examine whether these two biases still exist in
331 current climate models, we assessed the performances of CMIP5 models in
332 representing these feedbacks. As indicated by the traditional feedback calculations as
333 well as by a skill score designed for further evaluating the model simulations for the
334 entire tropical Pacific, the major characteristics of the feedbacks from cloud and water
335 vapor over the tropical Pacific regions in response to ENSO forcing are generally well
336 captured by CMIP5 models. The skill in simulating the SWCRF feedback is found to
337 be significantly higher in CMIP5 models than in CMIP3 models. The most impressive
338 improvement in this regard was noted in the NCAR Climate System Model (Gent et al.
339 2011). Nevertheless, a large spread among models and a notable difference between
340 the models and observations still exist in the historical runs of CMIP5 models. In
341 addition, most of the CMIP5 models do not have significant improvements in
342 simulating the water vapor feedback. More efforts need to be made to alleviate the
343 positive bias in the water vapor feedback that still remains prevalent in CMIP5

344 models.

345 Given the fact that the biases from many other feedbacks correlate with the bias
346 in the SWCRF feedback, 18 CMIP5 models were further classified as the HSM and
347 the LSM according to their scores in simulating the SWCRF feedback in the historical
348 runs. By analyzing the differences of climatological mean tropical Pacific SST
349 between the HSM and the LSM, we reveal that the LSM tend to have a more severe
350 excessive cold tongue than HSM. Further analysis demonstrates that the excessive
351 cold tongue generally leads to a westward shift and a significant weakening of the
352 maximum response to El Niño warming. Among the coupled models, there is a
353 correlation between the cold bias in the equatorial cold-tongue and the bias in the
354 SWCRF feedback.

355 There are some improvements in the AMIP runs in simulating the SWCRF
356 feedback from CMIP3 models to CMIP5 models, but they remain small relative to the
357 inter-model differences. Nonetheless, understanding the causes for these
358 improvements is important as we have noted that a small gain in the AMIP runs
359 generally result in a much reduced bias in the coupled runs. Factors that are
360 potentially responsible for the improvements in the AMIP runs involve the improved
361 convection parameterization, the cloud micro-physics schemes and the model
362 resolutions (Zhang et al. 1998; Lin 2007; Li and G. J. Zhang 2008; Guilyardi et al.
363 2009; Lloyd et al. 2009, 2011, 2012). For example, there are some improvements of
364 SWCRF feedback in the atmospheric component model of FGOALS-g2 (named

365 GAMIL2) compared to its previous version (named GAMIL1) (Li et al. 2012a). It has
366 been pointed out by Li et al. (2012b, submitted) that the biases of shortwave and
367 longwave cloud forcing distribution are reduced in GAMIL2 and the strength of
368 atmospheric response to El Niño warming (i.e., regression coefficient between
369 Southern Oscillation Index (SOI) and Niño3 SSTA) is much closer to the observations
370 than GAMIL1. The former arises from the updated cloud macro/micro-physics
371 schemes and their coordination with other schemes (e.g., the convection
372 parameterization). And the latter is related to the artificially amplified liquid water
373 path (LWP) in GAMIL1, which leads to an unrealistic response of the SWCRF and
374 other corresponding variables to ENSO cycle. To be sure, factors responsible for the
375 improvements in AMIP runs may vary with models or be model dependent.

376

377 **Acknowledgements.** We would like to thank two anonymous reviewers for their
378 constructive and insightful comments. The discussion with Dr. Li about the model
379 FGOLAS-g2 is gratefully acknowledged. This study is jointly supported by the
380 “Strategic Priority Research Program Climate Change: Carbon Budget and Relevant
381 Issues” of the Chinese Academy of Sciences (Grant No. XDA05110301), and the
382 National Key Program for Developing Basic Sciences Grant No. 2010CB950502, the
383 US National Science Foundation Climate Dynamics Program under AGS 0852329
384 and by grants from NOAA office of global programs (MAPP and ESS). We also

385 acknowledge the modeling groups, the PCMDI and the World Climate Research
386 Program's (WCRP's) Working Group on Coupled Modeling (WGCM) for their roles
387 in making available the WCRP CMIP5 multimodel dataset. Support of this dataset is
388 provided by the Office of Science, U.S. Department of Energy.

389 **References**

- 390 Bony, S., and J.-L. Dufresne, 2005: Marine boundary layer clouds at the heart of
391 tropical cloud feedback uncertainties in climate models. *Geophys. Res. Lett.*, **32**,
392 L20806, doi:10.10292005GL023851.
- 393 —, and Coauthors, 2006: How well do we understand and evaluate climate change
394 feedback processes? *J. Climate*, **19**, 3445–3482.
- 395 Cess, R. D., and Coauthors, 1990: Intercomparison and interpretation of climate
396 feedback processes in 19 atmospheric general circulation models. *J. Geophys.*
397 *Res.*, **95**, 16601–16615.
- 398 —, and Coauthors, 1996: Cloud feedback in atmospheric general circulation
399 models: An update. *J. Geophys. Res.*, **101**, 12791–12794.
- 400 Cronin, M. F., N. A. Bond, C. W. Fairall, and R. A. Weller, 2006: Surface cloud
401 forcing in the east Pacific stratus deck/cold tongue/ITCZ complex. *J.*
402 *Climate*, **19**, 392–409.
- 403 Davey, M. K., and Coauthors, 2002: STOIC: A study of coupled model climatology
404 and variability in tropical ocean regions. *Climate Dyn.*, **18**, 403–420.
- 405 Gent, P. R., and Coauthors, 2011: The Community Climate System Model version 4. *J.*
406 *Climate*, **24**, 4973–4991.
- 407 Graham, N. E., and T. P. Barnett, 1987: Sea surface temperature, surface wind
408 divergence, and convection over tropical oceans. *Science*, **238**, 657–659.

409 Guilyardi E, P. Braconnot, F.-F. Jin, S. T. Kim, M. Kolasinski, T. Li, and I. Musat,
410 2009: Atmosphere feedbacks during ENSO in a coupled GCM with a modified
411 atmospheric convection scheme. *J. Climate*, **22**, 5698–5718.

412 Harrison, E. F., P. Minnis, B. R. Barkstrom, V. Ramanathan, R. D. Cess, and G. G.
413 Gibson, 1990: Seasonal variation of cloud radiative forcing derived from the
414 Earth Radiation Budget Experiment. *J. Geophys. Res.*, **95**, 18 687–18 703.

415 Hartmann, D. L., and D. A. Short, 1980: On the use of Earth radiation budget statistics
416 for studies of clouds and climate, *J. Atmos. Sci.*, **37**, 1233-1250.

417 Houghton, J. T., Y. Ding, D. J. Griggs, M. Noguera, P. J. van der Linden, X. Dai, K.
418 Maskell, and C. A. Johnson, Eds., 2001: *Climate Change 2001: The Scientific*
419 *Basis*. Cambridge University Press, 881 pp.

420 Hirota, N., Y. N. Takayabu, M. Watanabe, and M. Kimoto, 2011: Precipitation
421 reproducibility over tropical oceans and its relationship to the double ITCZ
422 problem in CMIP3 and MIROC5 climate models. *J. Climate*, **24**, 4859–4873.

423 Huffman, G. J., R. F. Adler, D. T. Bolvin, and G. Gu, 2009: Improving the global
424 precipitation record: GPCP version 2.1. *Geophys. Res. Lett.*, **36**, L17808,
425 doi:10.1029/2009GL040000.

426 Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G.
427 L. Potter, 2002: NCEP–DOE AMIP-II Reanalysis (R-2). *Bull. Amer. Meteor. Soc.*,
428 **83**, 1631–1643.

429 Kiehl, J. T., and K. E. Trenberth, 1997: Earth’s annual global mean energy budget.

430 *Bull. Amer. Meteor. Soc.*, **78**, 197–208.

431 Latif , and Coauthors, 2001: ENSIP: The El Niño simulation inter-comparison project.
432 *Climate Dyn.*, **18**, 255–276.

433 Li, G. and G. J. Zhang, 2008: Understanding biases in shortwave cloud radiative
434 forcing in the National Center for Atmospheric Research Community
435 Atmosphere Model (CAM3) during El Niño. *J. Geophys. Res.*, **113**, D02103,
436 doi:10.1029/2007JD08963.

437 Li, G. and G. J. Zhang, 2008: Understanding biases in shortwave cloud radiative
438 forcing in the National Center for Atmospheric Research Community
439 Atmosphere Model (CAM3) during El Niño. *J. Geophys. Res.*, **113**, D02103,
440 doi:10.1029/2007JD08963.

441 Li Lijuan, and Coauthors, 2012a: The Flexible Global Ocean-Atmosphere-Land
442 System Model: Grid-point Version 2: FGOASLS-g2. *Adv. Atmos. Sci.*, in press.

443 —, and Coauthors, 2012b: Development and Evaluation of grid-point Atmospheric
444 Model of IAP LASG version 2 (GAMIL2). *Submitted to Adv. Atmos. Sci.*

445 Lin, J-L., 2007: The double-ITCZ problem in IPCC AR4 coupled GCMs:
446 Ocean–atmosphere feedback analysis. *J. Climate*, **20**, 4497–4525.

447 Lloyd, J., E. Guilyardi, H. Weller, and J. Slingo, 2009: The role of atmosphere
448 feedbacks during ENSO in the CMIP3 models. *Atmos. Sci. Lett.*, **10**, 170–176.

449 —, — and —, 2011: The role of atmosphere feedbacks during ENSO in the
450 CMIP3 models. Part II: Using AMIP runs to understand the heat flux feedback

451 mechanisms. *Climate Dyn.*, **37**, 1271–1292.

452 ———, ——— and ———, 2012: The Role of Atmosphere Feedbacks during ENSO in
453 the CMIP3 Models. Part III: The Shortwave Flux Feedback. *J. Climate*, **25**,
454 4275–4293.

455 Manabe, S., and R. T. Wetherald, 1967: Thermal equilibrium of the atmosphere with a
456 given distribution of relative humidity. *J. Atmos. Sci.*, **24**, 241–259.

457 Mechoso, C. R., and Coauthors, 1995: The seasonal cycle over the tropical Pacific in
458 coupled ocean–atmosphere general circulation models. *Mon. Wea. Rev.*, **123**,
459 2825–2838.

460 Randall, D. A., and Coauthors, 2007: Climate models and their evaluation. *Climate*
461 *Change 2007: The Physical Science Basis*. S. Solomon et al., Eds., Cambridge
462 University Press, 589–662.

463 Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P.
464 Rowell, E. C. Kent, and A. Kaplan, 2003: Global analyses of sea surface
465 temperature, sea ice, and night marine air temperature since the late nineteenth
466 century. *J. Geophys. Res.*, **108**, 4407, doi:10.1029/2002JD002670.

467 Schneider, Edwin K., Ben P. Kirtman, Richard S. Lindzen, 1999: Tropospheric Water
468 Vapor and Climate Sensitivity. *J. Atmos. Sci.*, **56**, 1649–1658.

469 Shine, K. P., and A. Sinha, 1991: Sensitivity of the Earth’s climate to height dependent
470 changes in the water vapor mixing ratio. *Nature*, **354**, 382–384.

471 Soden, B. J., 1997: Variations in the tropical greenhouse effect during El Niño. *J.*
472 *Climate*, **10**, 1050–1055.

473 ———, and I. Held, 2006: An assessment of climate feedbacks in coupled
474 ocean–atmosphere models. *J. Climate*, **19**, 3354–3360.

475 ———, I. M. Held, R. Colman, K. M. Shell, J. T. Kiehl, and C. A. Shields, 2008:
476 Quantifying climate feedbacks using radiative kernels. *J. Climate*, **21**,
477 3504–3520.

478 Stephens, G. L., 2005: Cloud feedbacks in the climate system: A critical review. *J.*
479 *Climate*, **18**, 237–273.

480 Sun, D.-Z., and I. M. Held, 1996: A comparison of modeled and observed
481 relationships between interannual variations of water vapor and temperature. *J.*
482 *Climate*, **9**, 665–675.

483 ———, J. Fasullo, T. Zhang, and A. Roubicek, 2003: On the radiative and dynamical
484 feedbacks over the equatorial Pacific cold tongue. *J. Climate*, **16**, 2425–2432.

485 ———, and Coauthors, 2006: Radiative and dynamical feedbacks over the equatorial
486 cold tongue: Results from nine atmospheric GCMs. *J. Climate*, **19**, 4059–4074.

487 ———, Y. Yu, and T. Zhang, 2009: Tropical water vapor and cloud feedbacks in
488 climate models: A further assessment using coupled simulations. *J. Climate*, **22**,
489 1287–1304.

490 Taylor, K. E., 2001: Summarizing multiple aspects of model performance in a single
491 diagram. *J. Geophys. Res.*, **106** (D7), 183–7192.

492 Taylor, K.E., R.J. Stouffer, G.A. Meehl, 2012: An Overview of CMIP5 and the
493 experiment design. *Bull. Amer. Meteor. Soc.*, **93**, 485-498.

494 Xie, Shang-Ping, Clara Deser, Gabriel A. Vecchi, Jian Ma, Haiyan Teng, Andrew T.
495 Wittenberg, 2010: Global Warming Pattern Formation: Sea Surface Temperature
496 and Rainfall*. *J. Climate*, **23**, 966–986.

497 Zhang, G. J., J. T. Kiehl, and P. J. Rasch, 1998: Response of climate simulation to a
498 new convective parameterization in the National Center for Atmospheric
499 Research Community Climate Model (CCM3). *J. Climate*, **11**, 2097-2115.

500 Zhang, T., and D. -Z. Sun, 2006: Response of water vapor and clouds to El Niño
501 warning in three National Center for Atmospheric Research atmospheric models.
502 *J. Geophys. Res.*, **111**, D17103, doi:10.1029/2005JD006700.

503 ——— and ———, 2008: What Causes the Excessive Response of the Clear-Sky
504 Greenhouse Effect to El Niño Warming in the NCAR Community Atmosphere
505 Models? *J. Geophys. Res.*, **113**, D02108, doi:10.1029/2007JD009247.

506 Zhang, Y., W. B. Rossow, A. A. Lacis, V. Oinas, and M. I. Mishchenko, 2004:
507 Calculation of radiative fluxes from the surface to top of atmosphere based on
508 ISCCP and other global data sets: Refinements of the radiative transfer model
509 and the input data. *J. Geophys. Res.*, **109**, D19105, doi:10.1029/2003JD004457.

510 Zhu, P., J. J. Hack, J. T. Kiehl, and C. S. Bretherton, 2007: Climate sensitivity of
511 tropical and subtropical marine low cloud amount to ENSO and global warming
512 due to doubled CO₂. *J. Geophys. Res.*, **112**, D17108,

514 **Table Captions**

515 **Table 1** List of CMIP5 models analyzed in this study.

516 **Table 2** Tropical cloud and water vapor feedbacks from observations and coupled

517 models: the water vapor feedback ($\frac{\partial(G_a)}{\partial T}$), the LWCRF feedback ($\frac{\partial(C_l)}{\partial T}$), the

518 SWCRF feedback ($\frac{\partial(C_s)}{\partial T}$), the feedback from the atmospheric transport ($\frac{\partial(D_a)}{\partial T}$), the

519 net atmospheric feedback ($\frac{\partial(F_a)}{\partial T} = \frac{\partial(G_a)}{\partial T} + \frac{\partial(C_l)}{\partial T} + \frac{\partial(C_s)}{\partial T} + \frac{\partial(D_a)}{\partial T}$), and the

520 feedback from the net surface heat flux into the ocean ($\frac{\partial(F_s)}{\partial T}$). The values of these

521 feedbacks are obtained through a linear regression using the interannual variations of

522 SSTA and the corresponding fluxes over the equatorial Pacific (5°S-5°N,

523 150°E-110°W). The results in the parentheses are obtained from the corresponding

524 AMIP runs. The values in the last two rows are ensemble model mean for two groups

525 of models—High Score Model (HSM) and Low Score Model (LSM). The definition

526 of HSM and LSM are provided in Section 4. Boldface indicates the results from the

527 models that are in the HSM. For convenient comparison, some results from CMIP3

528 models (derived from Table 2 in Sun2009) are also shown but with italicization.

529

530 **Figure Captions**

531 **Fig. 1** Response of the shortwave cloud radiative forcing (SWCRF) to El Niño
532 warming ($\frac{\partial(C_s)}{\partial T}$) (shading; units: $W m^{-2} K^{-1}$) for (a) the observations, (b)-(s) the
533 AMIP runs of CMIP5 models, (t) the ensemble model mean of the HSM and (u) the
534 ensemble model mean of the LSM. The labels at the top right corner are the models'
535 scores for their simulation of the SWCRF feedback.

536 **Fig. 2** Same as Fig. 1, but for the corresponding historical runs.

537 **Fig. 3** Scatter diagrams (a)-(c) showing the relationship in the historical runs between
538 the biases of: the SWCRF feedback and (a) the precipitation feedback, (b) the
539 LWCRF feedback, and (c) the feedback from ω at 500 hPa in historical runs. (d) The
540 relationship between the SWCRF feedback in AMIP runs and the SWCRF feedback in
541 historical runs. The dashed line indicates the score value (0.5) used for sorting the
542 models into two groups (ie, the HSM and the LSM). The red (blue) ones indicate the
543 HSM (LSM).

544 **FIG.4.** Diagram displaying the relationship between the skill scores of the simulated
545 SWCRF feedbacks and the standard deviation of Niño3 SSTA in the coupled runs.
546 The radial distance from the origin is the ratio of the simulated Niño3 SSTA standard
547 deviation against the observed. The red open circles indicate the results in the HSM
548 and the blue open squares indicate the results in the LSM. The skill scores of the
549 SWCRF feedback in the historical runs are given by the azimuthal position of the
550 open circles. The red (blue) ones indicate the HSM (LSM).

551 **Fig. 5** Tropical mean SST (shading; units: °C) for (a) the observation, (b) the ensemble
552 mean of the HSM, and (c) the ensemble mean of the LSM. (d) The differences in
553 ensemble mean SST between the two groups (LSM—HSM), and the biases of (e) the
554 HSM and (f) the LSM. All the model results are derived from the historical runs of
555 CMIP5 models.

556 **FIG.6** Scatter diagram showing the relationship between the skill score of the
557 SWCRF feedback and the cold-tongue biases as measured by SST difference from the
558 observations over the equatorial Pacific (5°S-5°N, 150°E-110°W) in the historical
559 runs (units: °C). The red (blue) ones indicate the HSM (LSM).

560 **FIG.7.** The net atmospheric feedback ($\frac{\partial(F_a)}{\partial T}$) (shading; units: W m⁻² K⁻¹), which are
561 derived from (a) OAFlux and ISCCP, (b) the HSM historical runs, (c) the LSM
562 historical runs.

563 **FIG.8.** Scatter diagram showing the values of skill scores of various feedbacks for the
564 AMIP and historical runs of the HSM and the LSM. The Ga, Cl, Ga+Cl, Cs and Fs on
565 the X axis indicate the water vapor feedback ($\frac{\partial(G_a)}{\partial T}$), the LWCRF feedback ($\frac{\partial(C_l)}{\partial T}$),
566 the feedback of the total greenhouse effect of water vapor and clouds ($\frac{\partial(G_a+C_l)}{\partial T}$), the
567 SWCRF feedback ($\frac{\partial(C_s)}{\partial T}$) and the feedback from the net surface heat flux into the
568 ocean ($\frac{\partial(F_s)}{\partial T}$), respectively. Note that coupling generally lowers the score of AMIP
569 runs, and the LSM is more affected than the HSM by the coupling in their simulation
570 of the feedbacks.

571 **Table 1** List of CMIP5 models analyzed in this study.

Model	Institution
ACCESS1-0	CSIRO (Commonwealth Scientific and Industrial Research Organisation, Australia), and BOM (Bureau of Meteorology, Australia)
BCC-CSM1-1	Beijing Climate Center(BCC),China Meteorological Administration, China
BNU-ESM	GCESS, BNU, Beijing, China
CanESM2	CCCma (Canadian Centre for Climate Modeling and Analysis, Victoria, BC, Canada)
CCSM4	NCAR (National Center for Atmospheric Research) Boulder, CO, USA
CNRM-CM5	CNRM (Centre National de Recherches Meteorologiques, Meteo-France, Toulouse, France) and CERFACS (Centre Europeen de Recherches et de Formation Avancee en Calcul Scientifique, Toulouse, France)
CSIRO-Mk3-6-0	Australian Commonwealth Scientific and Industrial Research Organization (CSIRO) Marine and Atmospheric Research (Melbourne, Australia) in collaboration with the Queensland Climate Change Centre of Excellence (QCCCE) (Brisbane,

	Australia)
FGOALS-g2	IAP (Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China) and THU (Tsinghua University)
FGOALS-s2	IAP (Institute of Atmospheric Physics), CAS (Chinese Academy of Sciences), Beijing, China
GISS-E2-R	NASA/GISS (Goddard Institute for Space Studies) New York, NY
HadGEM2-ES	Met Office Hadley Centre, Fitzroy Road, Exeter, Devon, EX1 3PB, UK
INMCM4	INM (Institute for Numerical Mathematics, Moscow, Russia)
IPSL-CM5A-LR	IPSL (Institut Pierre Simon Laplace, Paris, France)
MIROC5	AORI (Atmosphere and Ocean Research Institute, The University of Tokyo, Chiba, Japan), NIES (National Institute for Environmental Studies, Ibaraki, Japan), JAMSTEC (Japan Agency for Marine-Earth Science and Technology, Kanagawa, Japan)
MPI-ESM-LR	Max Planck Institute for Meteorology
MPI-ESM-MR	Max Planck Institute for Meteorology
MRI-CGCM3	MRI (Meteorological Research Institute, Tsukuba, Japan)

572

573 **Table 2** Tropical cloud and water vapor feedbacks from observations and coupled models: the water vapor feedback ($\frac{\partial(G_a)}{\partial T}$), the LWCRF feedback
 574 ($\frac{\partial(C_l)}{\partial T}$), the SWCRF feedback ($\frac{\partial(C_s)}{\partial T}$), the feedback from the atmospheric transport ($\frac{\partial(D_a)}{\partial T}$), the net atmospheric feedback
 575 ($\frac{\partial(F_a)}{\partial T} = \frac{\partial(G_a)}{\partial T} + \frac{\partial(C_l)}{\partial T} + \frac{\partial(C_s)}{\partial T} + \frac{\partial(D_a)}{\partial T}$), and the feedback from the net surface heat flux into the ocean ($\frac{\partial(F_s)}{\partial T}$). The values of these feedbacks are
 576 obtained through a linear regression using the interannual variations of SSTA and the corresponding fluxes over the equatorial Pacific (5°S-5°N,
 577 150°E-110°W). The results in the parentheses are obtained from the corresponding AMIP runs. The values in the last two rows are ensemble model mean
 578 for two groups of models—High Score Model (HSM) and Low Score Model (LSM). The definition of the HSM and the LSM are provided in Section 4.
 579 Boldface indicates the results from the models that are in the HSM. For convenient comparison, some results from CMIP3 models (derived from Table 2 in
 580 Sun2009) are also shown but with italicization.

581

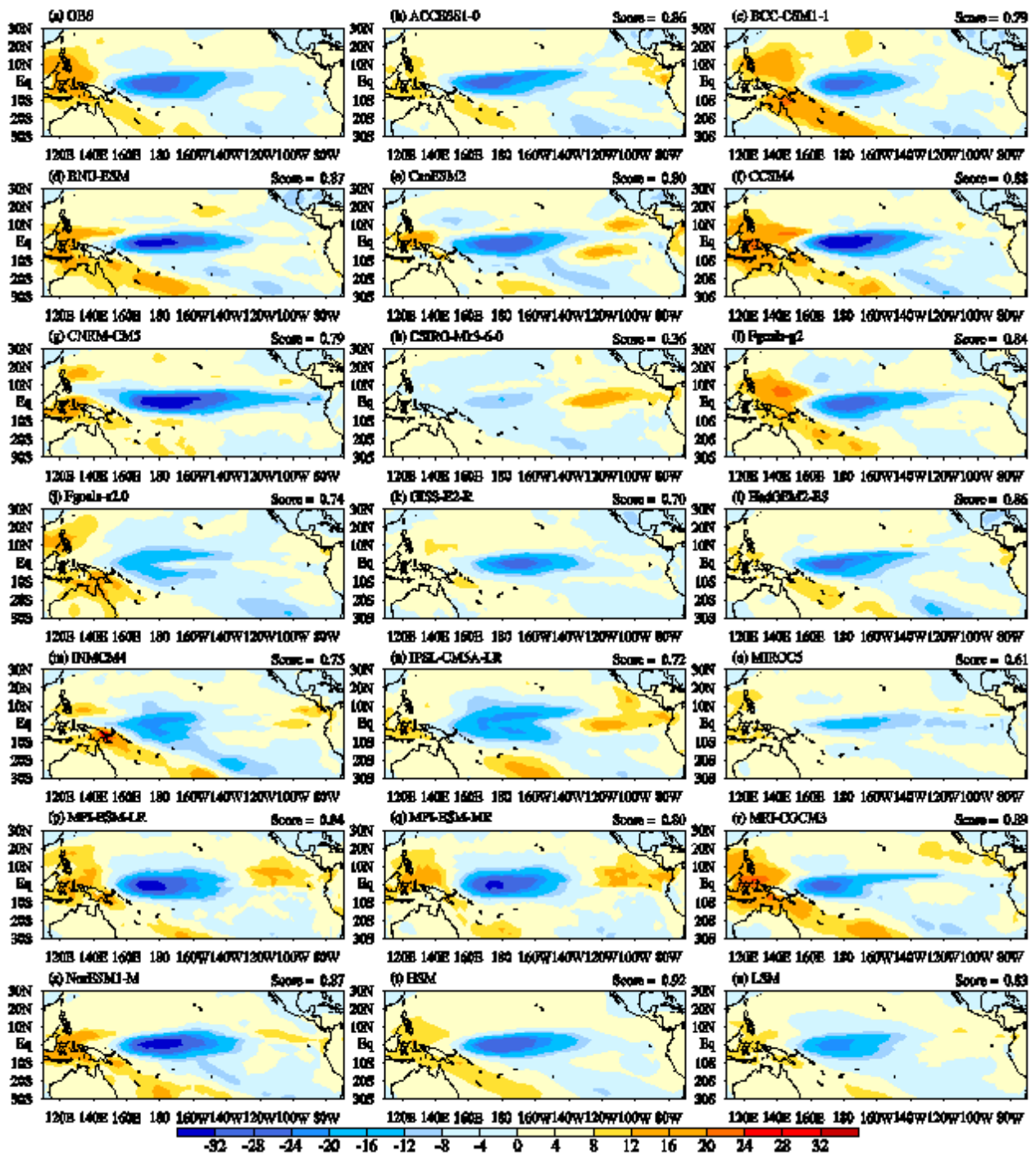
Feedback(W m⁻² K⁻¹)

Name of process	$\frac{\partial(G_a)}{\partial T}$	$\frac{\partial(C_1)}{\partial T}$	$\frac{\partial(G_a+C_1)}{\partial T}$	$\frac{\partial(C_s)}{\partial T}$	$\frac{\partial(C_1+C_s)}{\partial T}$	$\frac{\partial(D_a)}{\partial T}$	$\frac{\partial(F_a)}{\partial T}$	$\frac{\partial(F_s)}{\partial T}$
Observation	6.61±0.29	12.64±0.48	19.25±0.61	-13.33±0.68	-0.68±0.36	-20.28±1.01	-14.36±1.20	-21.43±1.04
ACCESS1-0	9.18±0.29 (9.96±0.26)	10.53±0.48 (12.33±0.51)	19.71±0.73 (22.29±0.73)	-9.16±0.56 (-11.73±0.60)	1.38±0.25 (0.60±0.26)	-17.57±0.91 (-23.05±0.93)	-7.02±0.89 (-12.48±0.95)	-12.21±0.93 (-19.00±0.98)
BCC-CSM1-1	7.74±0.15 (5.99±0.14)	8.37±0.36 (7.76±0.39)	16.11±0.46 (13.74±0.49)	-5.48±0.42 (-8.67±0.54)	2.89±0.27 (-0.91±0.35)	-16.36±0.81 (-20.29±0.78)	-5.73±0.83 (-15.22±0.88)	-12.74±0.90 (-22.16±0.93)
BNU-ESM	8.95±0.11 (9.11±0.18)	9.54±0.28 (12.53±0.44)	18.49±0.38 (21.65±0.59)	-7.01±0.48 (-12.74±0.69)	2.53±0.26 (-0.21±0.37)	-15.64±0.49 (-22.91±1.03)	-4.16±0.62 (-14.01±1.23)	-10.45±0.64 (-20.13±1.23)
CanESM2	8.48±0.14 (10.54±0.22)	9.35±0.38 (14.97±0.56)	17.83±0.50 (25.50±0.73)	-1.47±0.41 (-10.48±0.75)	7.88±0.25 (4.48±0.38)	-17.60±0.71 (-24.47±0.93)	-1.23±0.63 (-9.45±0.98)	-7.09±0.64 (-16.11±1.00)
CCSM4	9.16±0.13 (9.20±0.20)	11.86±0.35 (13.47±0.53)	21.01±0.44 (22.67±0.67)	-12.73±0.43 (-14.15±0.72)	-0.88±0.23 (-0.68±0.45)	-15.25±0.58 (-18.12±0.87)	-6.97±0.61 (-9.60±0.97)	-13.33±0.64 (-16.45±1.01)
<i>CCSM3</i>	8.95±0.15 (10.18±0.19)	10.26±0.47 (12.93±0.51)	19.21±0.57 (21.11±0.63)	-3.69±0.58 (-10.31±0.63)	6.57±0.38 (-2.62±0.42)	-8.50±0.54 (-5.39±0.80)	7.02±0.49 (7.42±0.83)	-0.32±0.51 (0.42±0.83)
CNRM-CM5	9.44±0.17	9.70±0.30	19.14±0.44	-17.95±0.57	-8.25±0.33	-13.55±0.73	-12.36±0.84	-18.40±0.86

	(8.24±0.19)	(9.49±0.35)	(17.73±0.50)	(-17.93±0.62)	(-8.44±0.38)	(-19.67±1.15)	(-19.87±1.30)	(-26.31±1.30)
<i>CNRM-CM3</i>	7.97±0.10	8.21±0.30	16.18±0.40	-9.53±0.48	-1.32±0.27	-7.67±0.48	-1.03±0.56	-4.71±0.49
	(10.30±0.25)	(13.74±0.67)	(24.04±0.89)	(-20.21±0.94)	(-6.47±0.54)	(-10.82±1.22)	(-6.99±1.25)	(-8.82±1.11)
<i>CSIRO-Mk3-6-0</i>	7.30±0.15	7.35±0.32	14.66±0.41	2.92±0.34	10.28±0.28	-13.45±0.67	4.14±0.67	-2.06±0.70
	(9.84±0.23)	(11.10±0.50)	(20.93±0.69)	(-0.60±0.36)	(10.50±0.37)	(-6.33±0.69)	(14.01±0.59)	(7.62±0.61)
FGOALS-g2	6.72±0.09	8.52±0.24	15.24±0.70	-14.46±0.40	-5.95±0.33	-16.65±0.47	-15.88±0.75	-20.96±0.88
	(6.26±0.16)	(6.51±0.26)	(12.77±0.38)	(-10.43±0.49)	(-3.91±0.38)	(-17.13±0.74)	(-14.79±0.97)	(-20.90±0.97)
<i>FGLALS-g1</i>	7.84±0.04	7.22±0.13	15.16±0.16	-2.25±0.12	-5.08±0.08	-12.55±0.23	-0.36±0.20	-6.11±0.20
	(N/A)	(N/A)	(N/A)	(N/A)	(N/A)	(N/A)	(N/A)	(N/A)
<i>FGOALS-s2</i>	9.50±0.18	5.59±0.24	15.09±0.39	-1.49±0.59	4.10±0.45	-18.66±0.64	-5.06±0.80	-9.73±0.84
	(9.85±0.26)	(6.19±0.35)	(16.04±0.55)	(-7.41±0.86)	(-1.23±0.61)	(-22.16±0.92)	(-13.53±1.06)	(-20.28±1.12)
GISS-E2-R	9.86±0.27	10.76±0.48	20.61±0.71	-10.56±0.42	0.19±0.34	-15.86±0.90	-5.81±0.79	-10.13±0.84
	(9.46±0.22)	(9.19±0.29)	(18.65±0.47)	(-9.91±0.33)	(-0.72±0.26)	(-15.70±0.80)	(-6.96±0.76)	(-14.03±0.82)
<i>GISS-EH</i>	7.07±0.24	0.54±0.36	7.61±0.47	1.70±0.58	2.25±0.50	N/A	N/A	N/A
	(N/A)	(N/A)	(N/A)	(N/A)	(N/A)	(N/A)	(N/A)	(N/A)
HadGEM2-ES	8.77±0.24	8.44±0.41	17.21±0.60	-2.40±0.57	6.04±0.31	-15.72±0.79	-0.91±0.81	-6.44±0.83

	(9.43±0.25)	(11.44±0.47)	(20.87±0.68)	(-11.06±0.54)	(0.39±0.24)	(-22.33±0.91)	(-12.51±0.91)	(-18.99±0.96)
<i>HadGEM1</i>	5.95±0.21	1.95±0.19	7.90±0.37	3.80±0.44	5.75±0.36	-5.70±0.52	6.00±0.69	-0.39±0.67
	(11.06±0.31)	(14.14±0.57)	(25.21±0.85)	(-14.02±0.68)	(-0.13±0.24)	(N/A)	(N/A)	(N/A)
INMCM4	8.96±0.19	3.51±0.26	12.47±0.40	0.60±0.55	4.12±0.37	-11.07±0.53	2.00±0.67	-4.15±0.66
	(8.34±0.16)	(6.08±0.29)	(14.42±0.41)	(-7.03±0.52)	(-0.95±0.31)	(-17.67±0.68)	(-10.29±0.79)	(-16.62±0.79)
IPSL-CM5A-LR	7.52±0.11	6.77±0.46	14.29±0.53	0.26±0.35	7.03±0.24	-11.32±0.60	3.23±0.46	-2.50±0.47
	(9.68±0.17)	(13.44±0.69)	(23.13±0.78)	(-8.17±0.69)	(5.28±0.40)	(-17.89±0.96)	(-2.93±0.85)	(-9.29±0.87)
<i>IPSL-CM4</i>	9.40±0.11	13.07±0.46	22.47±0.51	-9.56±0.54	3.51±0.33	-18.09±0.60	-5.18±0.61	-11.66±0.59
	(9.66±0.15)	(16.26±0.50)	(25.92±0.59)	(-13.94±0.62)	(2.32±0.33)	(-19.52±0.65)	(-7.53±0.65)	(-13.63±0.64)
MIROC5	8.65±0.12	9.57±0.26	18.22±0.37	-6.75±0.23	2.82±0.15	-15.76±0.50	-4.30±0.46	-10.53±0.47
	(9.16±0.19)	(9.69±0.45)	(18.85±0.59)	(-6.16±0.42)	(3.53±0.27)	(-15.50±1.19)	(-2.81±1.13)	(-9.33±1.16)
<i>MIROC3.2(hires)</i>	7.47±0.20	8.76±0.52	16.23±0.68	-1.32±0.56	7.45±0.30	-18.20±0.75	-3.28±0.69	-8.78±0.69
	(7.99±0.13)	(10.58±0.36)	(18.58±0.46)	(-4.21±0.46)	(6.38±0.24)	(-19.93±0.69)	(-5.57±0.74)	(-10.56±0.68)
MPI-ESM-LR	6.63±0.15	4.91±0.25	11.54±0.37	-3.41±0.31	1.49±0.19	-10.59±0.47	-2.47±0.43	-7.40±0.44
	(9.61±0.25)	(13.22±0.57)	(22.84±0.77)	(-12.25±0.64)	(0.97±0.29)	(-22.67±1.18)	(-12.08±1.19)	(-18.70±1.23)
MPI-ESM-MR	7.28±0.24	4.37±0.37	11.64±0.57	-3.54±0.45	0.82±0.26	-9.08±0.75	-0.98±0.64	-6.26±0.70

<i>ECHAM5/MPI</i>	(9.62±0.24)	(12.27±0.50)	(21.89±0.68)	(-11.82±0.61)	(0.45±0.30)	(-21.32±1.08)	(-11.24±1.10)	(-17.86±1.11)
	<i>7.90±0.13</i>	<i>6.77±0.28</i>	<i>14.67±0.39</i>	<i>-4.37±0.44</i>	<i>2.40±0.23</i>	<i>-14.62±0.49</i>	<i>-4.32±0.54</i>	<i>-10.33±0.52</i>
	(10.56±0.25)	(17.15±0.65)	(27.71±0.87)	(-20.31±1.01)	(-3.16±0.52)	(-27.20±1.08)	(-19.80±1.30)	(-24.81±1.25)
MRI-CGCM3	8.57±0.24	7.10±0.43	15.68±0.62	1.51±0.68	8.61±0.46	-15.30±0.91	1.89±1.05	-4.23±1.07
	(8.94±0.22)	(10.33±0.44)	(19.27±0.61)	(-9.15±0.63)	(1.18±0.36)	(-20.89±0.94)	(-10.77±1.06)	(-17.08±1.07)
NorESM1-M	9.98±0.19	13.51±0.58	23.49±0.70	-14.27±0.68	-7.60±0.41	-15.41±0.78	-6.19±0.84	-11.44±0.87
	(10.20±0.22)	(13.98±0.67)	(24.19±0.82)	(-13.80±0.84)	(0.19±0.49)	(-13.96±0.96)	(-3.57±1.08)	(-10.12±1.09)
HSM	8.87±0.18	10.14±0.39	19.00±0.57	-10.45±0.50	-1.07±0.30	-15.78±0.72	-7.23±0.78	-12.90±0.82
	(8.65±0.20)	(10.74±0.43)	(19.40±0.59)	(-12.27±0.60)	(-1.52±0.35)	(-19.24±0.91)	(-12.11±1.01)	(-18.68±1.03)
LSM	8.10±0.17	6.50±0.33	14.60±0.46	-1.26±0.43	5.24±0.29	-13.65±0.64	-0.31±0.65	-6.00±0.67
	(9.51±0.22)	(10.81±0.48)	(20.32±0.65)	(-8.12±0.61)	(2.69±0.37)	(-18.77±0.95)	(-6.57±0.97)	(-13.07±1.00)



582

583

584

585

FIG.1. Response of the shortwave cloud radiative forcing (SWCRF) to El Niño warming ($\frac{\partial(C_s)}{\partial T}$) (shading; units: $W m^{-2} K^{-1}$) for (a) the observations, (b)-(s) the AMIP runs of CMIP5 models, (t) the ensemble model mean of the HSM and (u) the ensemble model mean of the LSM. The labels at the top

586 right corner are the models' scores for their simulation of the SWCRF feedback.

587

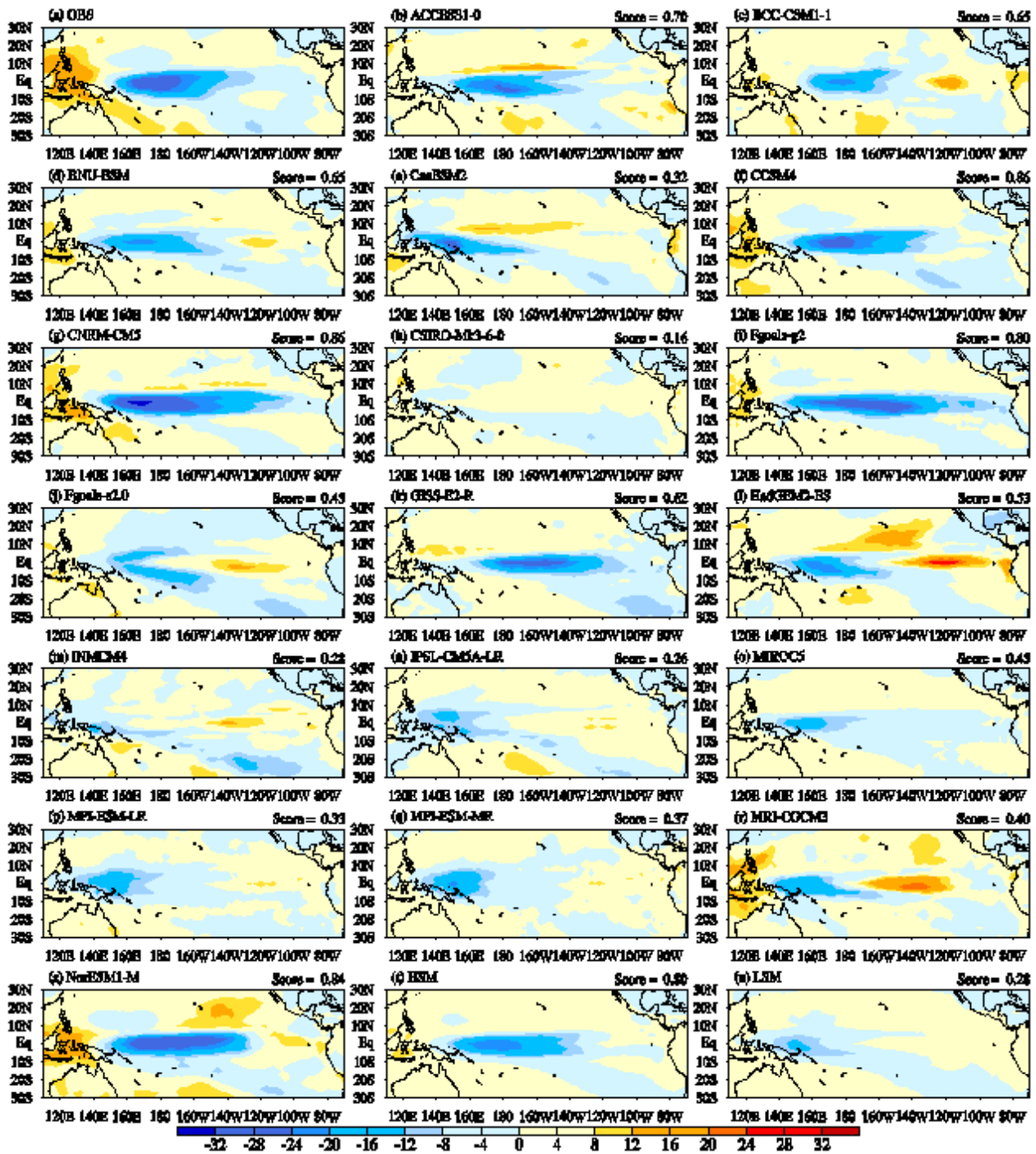


FIG.2. Same as Fig. 1, but for the corresponding historical runs.

591

Figure 3(a) and (b) are scatter plots showing the relationship between SWCRF feedback in historical runs (y-axis, 0 to 1.0) and other feedback metrics (x-axis, 0.0 to 1.0). Plot (a) compares SWCRF feedback with precipitation feedback, while plot (b) compares it with LWCRF feedback. Both plots feature 18 numbered data points representing different climate models. The legend for both plots is as follows:

- 1 ACCESS1-0
- 2 BCC-CBM1-1
- 3 BNU-BSM
- 4 CanESM2
- 6 CCSM4
- 6 CNRM-CM5
- 7 CSIRO-Mk3.6-0
- 8 Fgoals-g2
- 9 Fgoals-g2
- 10 GISS-ER-R
- 11 HadCM3-ES
- 12 INMCM4
- 13 IPSL-CM5A-LR
- 14 MIROC3
- 16 MPI-ESM-LR
- 16 MPI-ESM-MR
- 17 MRI-CGCM3
- 18 NorESM1-M

592

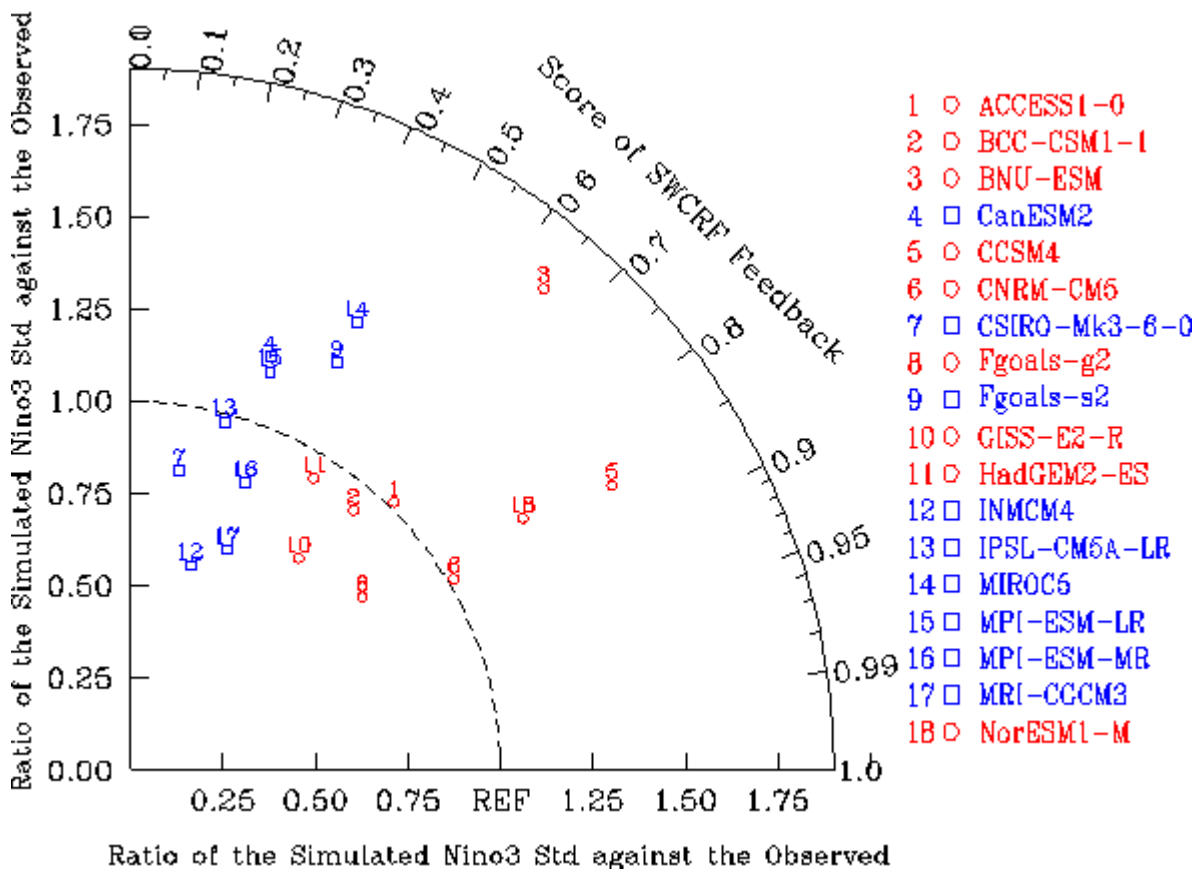
Figure 3(c) and (d) are scatter plots showing the relationship between SWCRF feedback in historical runs (y-axis, 0 to 1.0) and other feedback metrics (x-axis, 0.0 to 1.0). Plot (c) compares SWCRF feedback with Omega(500hPa) feedback, while plot (d) compares it with SWCRF feedback in AMIP runs. Both plots feature 18 numbered data points representing different climate models. The legend for both plots is as follows:

- 1 ACCESS1-0
- 2 BCC-CBM1-1
- 3 BNU-BSM
- 4 CanESM2
- 6 CCSM4
- 6 CNRM-CM5
- 7 CSIRO-Mk3.6-0
- 8 Fgoals-g2
- 9 Fgoals-g2
- 10 GISS-ER-R
- 11 HadCM3-ES
- 12 INMCM4
- 13 IPSL-CM5A-LR
- 14 MIROC3
- 16 MPI-ESM-LR
- 16 MPI-ESM-MR
- 17 MRI-CGCM3
- 18 NorESM1-M

In plot (d), a horizontal dashed line is drawn at a y-axis value of 0.5, indicating the score used for sorting models into HSM and LSM groups.

593 FIG.3. Scatter diagrams (a)-(c) showing the relationship in the historical runs between the biases of: the
 594 SWCRF feedback and (a) the precipitation feedback, (b) the LWCRF feedback, and (c) the feedback
 595 from ω at 500 hPa in historical runs. (d) The relationship between the SWCRF feedback in AMIP runs
 596 and that in historical runs. The dashed line indicates the score value (0.5) used for sorting the models
 597 into two groups (ie, the HSM and the LSM). The red (blue) ones indicate the HSM (LSM).
 598

41



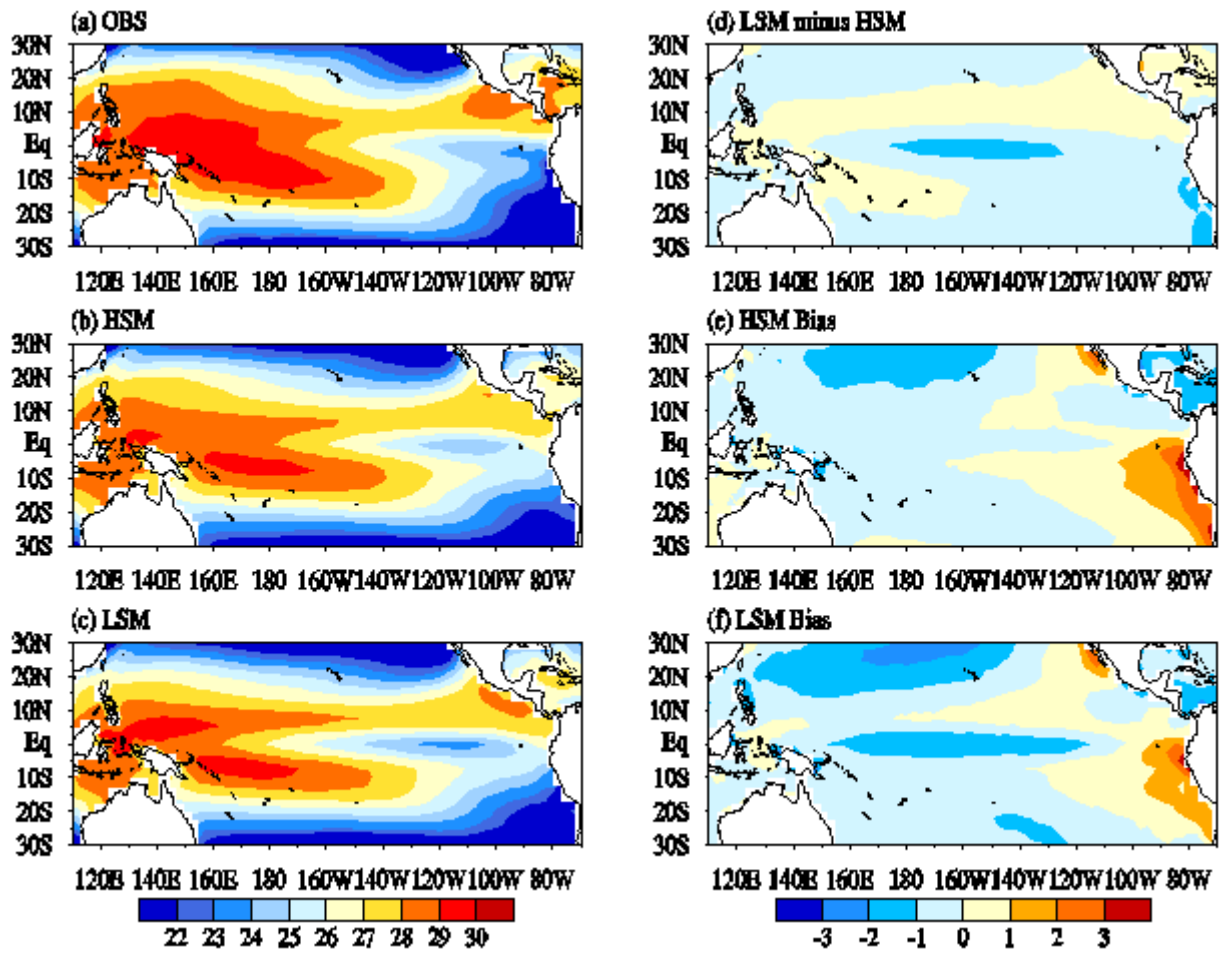
599

600

601 FIG.4. Diagram displaying the relationship between the skill scores of the simulated SWCRF feedbacks
 602 and the standard deviation of Niño3 SSTA in the coupled runs. The radial distance from the origin is the
 603 ratio of the simulated Niño3 SSTA standard deviation against the observed. The red open circles indicate
 604 the results in the HSM and the blue open squares indicate the results in the LSM. The skill scores of the
 605 SWCRF feedback in the historical runs are given by the azimuthal position of the open circles. The red
 606 (blue) ones indicate the HSM (LSM).

607

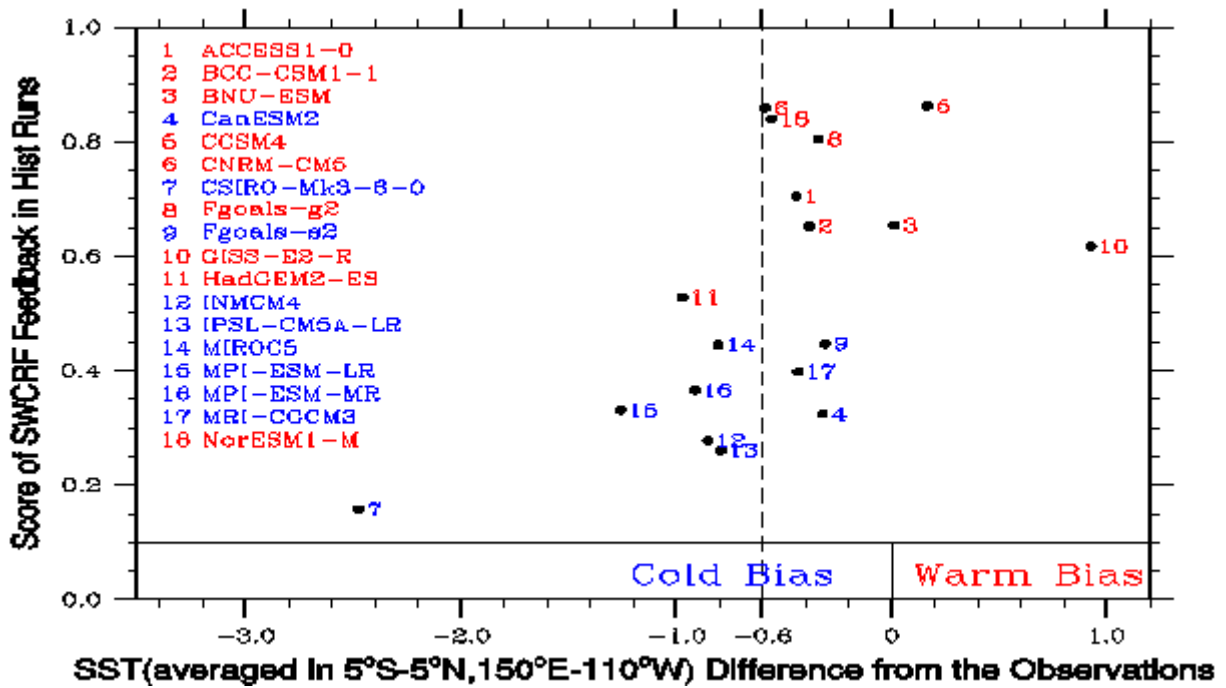
608



609 FIG.5. Tropical mean SST (shading; units: °C) for (a) the observation, (b) the ensemble mean of the
610 HSM, and (c) the ensemble mean of the LSM. (d) The differences in ensemble mean SST between the
611 two groups (LSM - HSM), and the biases of (e) the HSM and (f) the LSM. All the model results are
612 derived from the historical runs of CMIP5 models.

613

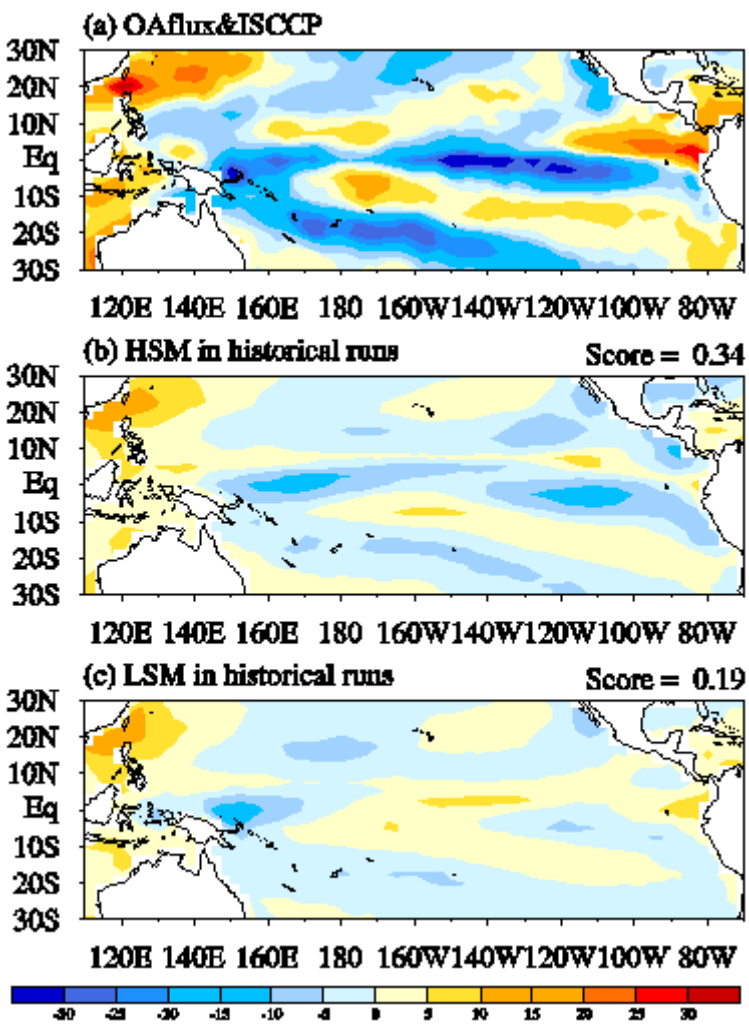
614



615

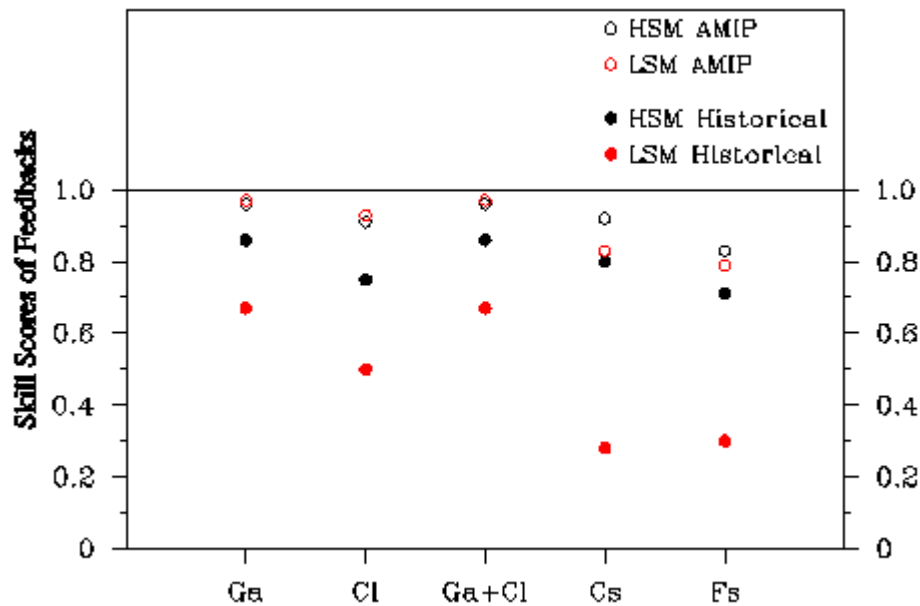
616 FIG.6. Scatter diagram showing the relationship between the skill score of the SWCRF feedback and the
617 cold-tongue biases as measured by SST difference from the observations over the equatorial Pacific
618 (5°S-5°N, 150°E-110°W) in the historical runs (units: °C). The red (blue) ones indicate the HSM (LSM).

619



620

621 FIG.7. The net atmospheric feedback ($\frac{\partial(F_a)}{\partial T}$) (shading; units: $W m^{-2} K^{-1}$), which are derived from (a)
622 OAFlux and ISCCP, (b) the HSM historical runs, (c) the LSM historical runs.



623

624 FIG.8. Scatter diagram showing the values of skill scores of various feedbacks for the AMIP and
 625 historical runs of the HSM and the LSM. The Ga, Cl, Ga+Cl, Cs and Fs on the X axis indicate the water
 626 vapor feedback ($\frac{\partial(G_a)}{\partial T}$), the LWCRF feedback ($\frac{\partial(C_l)}{\partial T}$), the feedback of the total greenhouse effect of
 627 water vapor and clouds ($\frac{\partial(G_a+C_l)}{\partial T}$), the SWCRF feedback ($\frac{\partial(C_s)}{\partial T}$) and the feedback from the net
 628 surface heat flux into the ocean ($\frac{\partial(F_s)}{\partial T}$), respectively. Note that coupling generally lowers the score of
 629 AMIP runs, and the LSM is more affected than the HSM by the coupling in their simulation of the
 630 feedbacks.