Enhanced future warming constrained by past trends in the equatorial Pacific sea surface temperature gradient

Masahiro Watanabe^{1*}, Jean-Louis Dufresne², Yu Kosaka³, Thorsten Mauritsen⁴ & Hiroaki Tatebe⁵

1: Atmosphere and Ocean Research Institute, University of Tokyo, Chiba, Japan

2: Laboratoire de Météorologie Dynamique, Institut Pierre et Simon Laplace, Paris, France

3: The Research Center for Advanced Science and Technology, the University of Tokyo, Tokyo, Japan

4: Department of Meteorology, Stockholm University, Stockholm, Sweden

5: Japan Agency for Marine-Earth Science and Technology, Kanagawa, Japan

* Corresponding author:

Masahiro Watanabe, Atmosphere and Ocean Research Institute, the University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8568, Japan.

Phone: 81-4-7136-4387, Fax: 81-4-7136-4375

E-mail: hiro@aori.u-tokyo.ac.jp

Nature Climate Change

Submitted on 5 December, 2019; Revised on X March, 2020

The zonal gradient in the equatorial Pacific sea surface temperature (SST), high in the west and low in the east, is known to be a pacemaker of global warming¹. The zonal SST gradient has strengthened since the mid-20th century², but the cause is controversial because it is not reproduced by a majority of Coupled Model Intercomparison Project Phase 5 (CMIP5) climate models^{3–6}, which instead suggest a weakening of the zonal SST gradient from the past to the future⁷. This discrepancy between observations and models has to be reconciled not only for attributing the past climate change but also for assessing Earth's climate sensitivity^{8,9}. Here we show using large ensemble (LE) simulations by four different climate models, in addition to CMIP5 models, that the intensified trend in the Pacific zonal SST gradient for 1951–2010 can be captured by some realisations, suggesting that it could arise from internal climate variability. Models and members that simulate the past strengthening of the SST gradient in CMIP5 and LE commonly exhibit reversed trends in future projections, when the rate of global-mean temperature rise is likely to amplify by 9–30% with the larger values occurring in low-emission scenarios.

Atmosphere-ocean state in the tropical Pacific is characterised by the zonal contrast between high SST in the warm pool and low SST in the cold tongue, tied with the Walker circulation accompanying surface easterlies. This mean state is maintained by the well-known Bjerknes feedback¹⁰, which acts against radiative heating that homogenises SST zonally. Specifically, equatorial easterlies pile warm water up to the west and cause upwelling to cool the eastern Pacific. The strengthened SST gradient, in turn, amplifies the easterlies by

enhancing atmospheric convection over the warm pool that excites the zonal gradient in sea level pressure. The same feedback explains how El Niño–Southern Oscillation (ENSO) can be excited at interannual time scales^{11,12}. Because of the ENSO variability, detecting longer term changes in the zonal SST gradient in 20th century observations remains a challenge.

In future projections by CMIP5 Representative Concentration Pathways (RCP) simulations¹³, SST in the equatorial central–eastern Pacific warms more than the western Pacific with a certain degree of uncertainty^{7,14,15}, and the change is sometimes referred to as the ‘El Niño-like’ pattern. There are several hypotheses that explain the long-term SST pattern change; an ocean dynamical thermostat which prevents excessive warming in the warm pool¹⁶ and surface wind-evaporation-SST feedback suppressing the western basin warming¹⁷. The weakening of the zonal SST gradient in future projections is coupled with a slowdown of the Walker circulation, which has been supported from both a theory¹⁸ and climate model simulations^{19,20}.

The CMIP5 historical simulations similarly show the weakening of the SST gradient during the 20th century³. However, it is not supported by observations, which commonly show that the SST gradient has strengthened since the middle of the 20th century² (Supplementary Fig. 1), consistent with a strengthening of the Walker circulation as suggested from atmospheric data²¹. There exist observational uncertainty²² and a dependence of the trend magnitude on chosen period⁶, but the discrepancy between observations and climate models is obvious and critical not only for assessing the reliability of future climate changes in models but also for

estimating the Earth's climate sensitivity (ECS) from historical records. Recent studies show that cloud and lapse rate feedbacks to doubling of CO₂ concentration depend on spatial inhomogeneity in sea surface warming^{8,9,23}, called the 'pattern effect'²⁴. Specifically, if the SST gradient in the tropical Pacific is strengthened, these feedbacks tend to be weak or change the sign from positive to negative, leading to a low estimate of ECS, and vice versa. Thus, reconciling past zonal SST gradient change may aid in narrowing the range of ECS estimates from historical observations and CMIP5 models.

The observed strengthening of the Pacific zonal SST gradient can be due to either an internal variability of the climate system or a forced response to greenhouse gases (GHGs) and other agents such as aerosols and solar activity. On one hand, individual model simulations do not necessarily reproduce observations if internal variability is the main driver of the observed trend, which should instead be covered by the whole ensemble. On the other hand, the model ensemble mean will capture the observed trend if it is attributable to the external radiative forcing given a possibility that cooling by the mean upwelling in the cold tongue retards the transient warming in this region. Both these cases however may not apply when models have biases due to systematic errors. A similar discussion has continued for understanding the cause of the global warming 'hiatus'²⁵ that lasted about 15 years since around 2000. The differences lie in the length of the linear trend (60 years for the present analysis), and the increasing influence of radiative forcing is expected at this time scale.

Analyses of the CMIP5 historical simulations show that the observed strengthening of the zonal SST gradient is outside of the ensemble of simulated trends, suggesting a primary role of external forcing whose impact is not adequately calculated in climate models^{3,4}. A recent study, based on simulations using a simplified atmosphere-ocean model with observed mean states, concluded that the GHG-forced response in CMIP5 models would be consistent with the observations if the bias in mean SST, which is too cold in the eastern Pacific, did not exist⁵. However, this conclusion raises another inconsistency with our current understanding that the SST gradient (and the Walker circulation) would be weakened in the long-term response to global warming¹⁴. Furthermore, it is plausible that the ensemble of opportunity by CMIP5 simulations insufficiently samples the spread due to internal variability. Indeed, increased number of realisations, using either long pre-industrial control runs or a large set of historical simulations by a single climate model, made it possible to include ensemble members that show both weakening and strengthening of the Walker circulation, the latter being similar to the observed feature^{6,26}.

To examine whether the observed changes in the equatorial Pacific zonal SST gradient, hereafter denoted as $\Delta\text{SST}_{\text{eq}}$ (east minus west, see Methods), reflect internal variability, we analysed a combined set of large ensembles (LEs) from four different climate models having 220 members in total, in addition to CMIP5 historical runs, and compared them with six SST data sets (Methods). Each LE consists of historical runs made by perturbing the initial condition and thus the ensemble deviations solely represent internal variability. We demonstrate that the

large set of simulations covers well the observed trend of $\Delta\text{SST}_{\text{eq}}$ for 1951–2010 although the ensemble shows a bias towards weakening $\Delta\text{SST}_{\text{eq}}$; the result indicates that the observed strengthening of the SST gradient can be caused by internal variability at the interdecadal time scale. Furthermore, the SST gradient will turn to a weakening by the end of this century in all RCP scenarios, during which a rate of global-mean surface air temperature (GSAT) increase is temporarily amplified.

In the four equatorial ocean strips (Indian Ocean, western and eastern Pacific, and Atlantic), observed SST has increased by about 0.6 K over 1951–2010 (Supplementary Fig. 2). It is well reproduced by the CMIP5 ensemble-mean anomalies except for the eastern Pacific where models overestimate the observed warming: 1.1 versus 0.78 K per century. While the differences in observed and simulated SST trends are apparently small, the trend in $\Delta\text{SST}_{\text{eq}}$ is quite distinct; all observational data sets show the negative trend but only a few models do the same (Fig. 1a and Supplementary Fig. 3). The observed $\Delta\text{SST}_{\text{eq}}$ trend is robust when shifting the 60-year period from 1951–2010 to 1960–2019 (Supplementary Fig. 4), but the magnitude is systematically different between interpolated analyses (HadISST, COBE-SST2, ERSSTv5, Kaplan) and un-interpolated data sets (HadSST3 and ICOADS). The ensemble of six data sets leads to the value of -0.33 ± 0.16 K per century (the range representing the standard deviation), which may overestimate the true trend given potential errors in the SST analysis²².

While CMIP5 models show a large inter-model spread, the average of the observed $\Delta\text{SST}_{\text{eq}}$ trends is outside the 5–95% range of the CMIP5-based trends (bars in Fig. 1b). This is

apparently consistent with the argument that models fail to represent the externally forced signal, which may explain the observed ΔSST_{eq} trend^{4,5}. However, probability distributions of ΔSST_{eq} trends from 220-member LEs, calculated for individual models, capture the observations (Fig. 1b). The ensemble-mean trends are either positive or negative among the four LEs, indicating a model uncertainty in the externally forced response. Whereas the agreement with the IPSL-CM6A-LR and MIROC6 ensembles are only marginally consistent with observations, CESM1 and MPI-ESM1.1 show 10 and 40% of the ensemble members exceeding the observed ΔSST_{eq} trend. Thus, the CMIP5 ensemble clearly underestimates the uncertainty range arising from natural internal variability, and LEs suggest that the observed negative ΔSST_{eq} trend can be caused by internal variability at the interdecadal time scale. The features found in Fig. 1 change little when interannual ENSO signals were removed before calculating the linear trends (Methods and [Supplementary Fig. 5](#)).

As Bjerknes feedback operates at any time scale, the change in ocean subsurface temperature should be different between models that reproduce or fail the observed ΔSST_{eq} trend. When we select six CMIP5 models that show the SST gradient strengthening (called S-models) and also weakening (W-models), we see a contrast in their means and spreads: -0.24 ± 0.1 and 0.39 ± 0.1 K per century (red and blue bars in Fig. 1a). Linear trends in the equatorial ocean temperature for 1951–2010 from observations, CMIP5 ensemble mean, and averages of S- and W-models, reveal an intriguing difference in addition to an overall similarity (Fig. 2).

The most prominent feature of the observed ocean temperature trend is the contrast between warming near the surface and cooling near the thermocline (Fig. 2a). The subsurface cooling is indicative of a rising of the thermocline from the dateline to the eastern Pacific. This causes the upwelling to cool the surface layer in the eastern basin, consistent with the negative ΔSST_{eq} trend. The CMIP5 ensemble-mean trend qualitatively simulates this subsurface cooling, which is regarded as a forced oceanic response (Fig. 2b). The shoaling trend of the equatorial thermocline has been reported in the literature⁷ and is accompanied by a deepening of the thermocline in subtropical oceans as measured by the trend in a depth of 20 °C isotherm in both observations and the CMIP5 ensemble mean (Supplementary Fig. 6).

Compared to the vertical structure of the observed temperature trend, the surface layer is thicker besides subsurface cooling weaker and shifted to a deeper layer in the CMIP5 ensemble mean. These models suggest that the upwelling does not effectively cool the eastern Pacific despite the rising of the thermocline. This is more clearly identified in W-models where subsurface cooling occurs only weakly in the western Pacific, in contrast to a large subsurface cooling in the eastern Pacific seen in S-models (Fig. 2c, d). The difference between S- and W-models may partly be attributed to a different structure of the externally forced response, but a quite similar contrast in the magnitude and location of subsurface cooling has been identified by categorising CESM1 LE into members that show a strengthening and a weakening of the Walker circulation²⁵. This suggests that the difference between S- and W-models is primarily due to the effect of internal variability.

A dominant low-frequency mode of internal variability in the Pacific is the Interdecadal Pacific Oscillation (IPO), featuring a similar spatial pattern to ENSO but modulating the eastern Pacific SST and hence $\Delta\text{SST}_{\text{eq}}$ at interdecadal time scales²⁷ (Supplementary Figs. 7-9). The phase of the IPO varies irregularly²⁸ and has changed more than once during 1951–2010 in both observations and each member of the historical ensemble. However, S-models tend to show a positive phase in the early period and a negative phase in the late period, and vice versa for W-models. The difference in the IPO trend is subtle, but it is clearer in the MPI LE. While irregularity of the IPO results in uncertainty to a certain degree, its turnabout is expected to occur in coming decades. The IPO phase reversal, in addition to the forced ‘El Niño-like’ pattern in future projections, is likely to amplify the positive $\Delta\text{SST}_{\text{eq}}$ trend for 2001–2050 in S-models but suppress it in W-models (Supplementary Fig. 10).

Indeed, the linear $\Delta\text{SST}_{\text{eq}}$ trends calculated using 50-year moving segments change the sign from negative to positive by the mid-21st century in S-models while they are positive throughout the period in W-models; their difference is maximum for 1991–2040 (Fig. 3a). Because the IPO predictability is less than a decade²⁹, this result does not indicate that the future phase of the IPO can be predicted by its past state but rather suggests that the recovery from negative or positive phase during the historical period influences the future multi-decadal trend in $\Delta\text{SST}_{\text{eq}}$. A very similar result was obtained from a single LE using MPI-ESM1.1, supporting the premise that near-term change in $\Delta\text{SST}_{\text{eq}}$ is modulated by internal variability (Fig. 3b). In both CMIP5 and the MPI LE, the $\Delta\text{SST}_{\text{eq}}$ trends are not well separated between

S- and W-groups in the late 21st century. This indicates that the amplification of weakening the SST gradient is a temporary phenomenon.

Previous debates on the global warming hiatus have made it clear that the equatorial Pacific SST pattern, as measured by $\Delta\text{SST}_{\text{eq}}$, is an important driver of GSAT to slow down or accelerate the radiatively forced response^{1,21,25,30}. Therefore, the evolution of multi-decadal trends in GSAT is expected to reveal a discernible difference between S- and W-models (Fig. 4). Consistent with the reversal of the $\Delta\text{SST}_{\text{eq}}$ trend in the future, a positive GSAT trend in S-models is larger than in W-models in all the RCP simulations, although the timing of this amplified global warming is dependent on the scenarios. In a stabilisation scenario of RCP2.6, the rates of warming are clearly different between S- and W-models (larger in the former than the latter) by about 2050, but the difference eventually diminishes thereafter (Fig. 4a). With a strong GHG forcing in RCP8.5, the difference in GSAT trend between S- and W-models appears throughout the century (Fig. 4c). Yet, the implications of attributing historical changes in zonal SST gradient is evident; the rate of global warming will be temporarily amplified if climate models reproduce the observed strengthening of the SST gradient for 1951–2010. S-models show the greatest amplification in the 50-year trend by 29.6% in RCP2.6 for 1981–2030 while the least by 8.9% in RCP8.5 for 2051–2100. This dependence on scenario can be interpreted in terms of the signal-to-noise ratio; the influence of internal variability (noise) on future warming rates is larger if the GHG-induced warming (signal) is smaller.

The small inter-model spread of the $\Delta\text{SST}_{\text{eq}}$ trend in CMIP5 may be partly due to an underestimation of the SST variance associated with the IPO^{28,31} (Supplementary Figs. 8 and 11). The four LEs except IPSL-CM6A-LR simulate the IPO with a larger amplitude than many CMIP5 models, and it might be a reason why LEs successfully captured the observed $\Delta\text{SST}_{\text{eq}}$ trend.

Our results do not exclude the possibility that climate models systematically fail to simulate a forced component in the $\Delta\text{SST}_{\text{eq}}$ trend⁵ given that the observed large trend lies outside the CMIP5 range (Fig. 1). This may not happen due solely to an extreme internal fluctuation at interdecadal time scales⁶. However, additional analyses show that the failure to reproduce the observed trend only spans over the recent two decades, but not the entire 60-year period (Supplementary Figs. 12 and 13). For 1991–2010, the observed $\Delta\text{SST}_{\text{eq}}$ trend is far stronger than the simulated trends in CMIP5 and even any of LEs. This period corresponds to the hiatus when the Pacific trade winds were stronger than ever since the early 20th century²¹. The intensified $\Delta\text{SST}_{\text{eq}}$ trend for 1991–2010, 15 times as large as the trend for 1951–2010, has been suggested to be forced partially by sulphate aerosols³² or the Atlantic warming³³. Although such a process might have not been represented by climate models well, our results indicate that considering internal climate variability is critical for understanding long-term changes in the Pacific zonal SST gradient and global-mean temperature not only for the past decades but also for the future projection.

References

1. Kosaka, Y. & Xie S.-P. The tropical Pacific as a key pacemaker of the variable rates of global warming. *Nature Geosci.* **9**, 669–673 (2016).
2. Solomon, A. & Newman, M. Reconciling disparate twentieth-century Indo-Pacific Ocean temperature trends in the instrumental record. *Nat. Clim. Change* **2**, 691–699 (2012).
3. Coats, S. & Karnauskas, K. B. Are simulated and observed twentieth century tropical Pacific sea surface temperature trends significant relative to internal variability? *Geophys. Res. Lett.* **44**, 9928–9937 (2017).
4. Luo, J.-J., Wang, G. & Dommenges, D. May common model biases reduce CMIP5’s ability to simulate the recent Pacific La Niña-like cooling? *Clim. Dynam.* **50**, 1335–1351 (2018).
5. Seager, R., Cane, M., Henderson, N. et al. Strengthening tropical Pacific zonal sea surface temperature gradient consistent with rising greenhouse gases. *Nat. Clim. Change* **9**, 517–522 (2019).
6. Bordbar, M. H., Martin, T., Latif, M. & Park, W. Role of internal variability in recent decadal to multidecadal tropical Pacific climate changes. *Geophys. Res. Lett.* **44**, 4246–4255 (2017).
7. Collins, M., An, S., Cai, W. et al. The impact of global warming on the tropical Pacific Ocean and El Niño. *Nature Geosci.* **3**, 391–397 (2010).

- 217 8. Andrews, T., Gregory, J. M., Paynter, D. et al. Accounting for changing temperature
218 patterns increases historical estimates of climate sensitivity. *Geophys. Res. Lett.* **45**, 8490–
219 8499 (2018).
- 220 9. Ceppi, P. & Gregory, J. M. Relationship of tropospheric stability to climate sensitivity and
221 Earth’s observed radiation budget. *Proc. Nat. Acad. Sci.* **114**, 13126–13131 (2017).
- 222 10. Bjerknes, J. Atmospheric teleconnections from the equatorial Pacific. *Mon. Weather Rev.*
223 **97**, 162–172 (1969).
- 224 11. Jin, F.-F. Tropical ocean-atmosphere interaction, the Pacific cold tongue, and the El Niño
225 Southern Oscillation. *Science* **274**, 76–78 (1996).
- 226 12. Timmermann, A., An, S.-I., Kug, J.-S. et al. El Niño–Southern Oscillation complexity.
227 *Nature* **559**, 535–545 (2018).
- 228 13. Taylor, K. E., Stouffer, R. J. & Meehl, G. A. An overview of CMIP5 and the experiment
229 design. *Bull. Amer. Meteor. Soc.* **93**, 485–498 (2012).
- 230 14. Christensen, J. H. et al. Climate phenomena and their relevance for future regional Climate
231 change. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working*
232 *Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.*
233 Stocker, T. F. et al. (eds.), Cambridge University Press (2013).
- 234 15. Kim, S. T., Cai, W., Jin, F.-F. et al. Response of El Niño sea surface temperature variability
235 to greenhouse warming. *Nat. Clim. Change* **4**, 786–790 (2014).

- 236 16. Clement, A., Seager, R., Cane, M. A. & Zebiak, S. E. An ocean dynamical thermostat. *J.*
237 *Climate* **9**, 2190–2196 (2019).
- 238 17. Xie, S.-P., Deser, C. Vecchi, G. A. et al. Global warming pattern formation: Sea surface
239 temperature and rainfall. *J. Climate* **23**, 966–986 (2010).
- 240 18. Held, I. M. & Soden, B. J. Robust responses of the hydrological cycle to global warming.
241 *J. Climate* **19**, 5686–5699 (2006).
- 242 19. Vecchi, G. A., Soden, B. J., Wittenberg, A. T. et al. Weakening of tropical Pacific
243 atmospheric circulation due to anthropogenic forcing. *Nature* **441**, 73–76 (2006).
- 244 20. Chadwick, R., Boutle, I. & Martin, G. Spatial patterns of precipitation change in CMIP5:
245 Why the rich do not get richer in the tropics. *J. Climate* **26**, 3803–3822 (2013).
- 246 21. England, M. H., McGregor, S., Spence, P. et al. Recent intensification of wind-driven
247 circulation in the Pacific and the ongoing warming hiatus. *Nat. Clim. Change* **4**, 222–227
248 (2014).
- 249 22. Deser, C., Phillips, A. S. & Alexander, M. A. Twentieth century tropical sea surface
250 temperature trends revisited. *Geophys. Res. Lett.* **37**, L10701 (2010).
- 251 23. Zhou, C., Zelinka, M. D. & Klein, S. A. Impact of decadal cloud variations on the Earth’s
252 energy budget. *Nat. Geosci.* **9**, 871–874 (2016).
- 253 24. Dong, Y., Proistosescu, C., Armour, K. C & Battisti, D. S. Attributing historical and future
254 evolution of radiative feedbacks to regional warming patterns using a Green’s function
255 approach: The preeminence of the western Pacific. *J. Climate* **32**, 5471–5491 (2019).

- 256 25. Fyfe, J. C., Meehl, G. A., England, M. H. et al. Making sense of the early-2000s warming
257 slowdown. *Nat. Clim. Change* **6**, 224–228 (2016).
- 258 26. Chung, E.-S., Timmermann, A., Soden, B. J. et al. Reconciling opposing Walker
259 circulation trends in observations and model projections. *Nat. Clim. Change* **9**, 405–412
260 (2019).
- 261 27. Power, S., Casey, T., Folland, C. et al. Inter-decadal modulation of the impact of ENSO on
262 Australia. *Clim. Dynam.* **15**, 319–324 (1999).
- 263 28. Henley, B. J., Meehl, G. A., Power, S. et al. Spatial and temporal agreement in climate
264 model simulations of the Interdecadal Pacific Oscillation. *Env. Res. Lett.* **12**, 044011
265 (2017).
- 266 29. Meehl, G., Hu, A. & Teng, H. Initialized decadal prediction for transition to positive phase
267 of the Interdecadal Pacific Oscillation. *Nat. Comm.* **7**, 11718 (2016).
- 268 30. Bordbar, M. H., England, M. H., Gupta, A. S. et al. Uncertainty in near-term global surface
269 warming linked to tropical Pacific climate variability. *Nat. Comm.* **10**, 1990 (2019).
- 270 31. Kociuba, G. & Power, S. B. Inability of CMIP5 models to simulate recent strengthening
271 of the Walker circulation: Implications for projections. *J. Climate* **28**, 20–35 (2015).
- 272 32. Takahashi, C. & Watanabe, M. Pacific trade winds accelerated by aerosol forcing over the
273 past two decades. *Nat. Clim. Change* **6**, 768–772 (2016).
- 274 33. McGregor, S. et al. Recent Walker circulation strengthening and Pacific cooling amplified
275 by Atlantic warming. *Nat. Clim. Change* **4**, 888–892 (2014).

276

277 **Acknowledgements**

278 We acknowledge the modelling groups, the PCMDI, and the WCRP's WGCM for their efforts
279 in making the CMIP5 multi-model data set available. M.W., Y.K. and H.T. were supported by
280 the Grant-in-Aid 26247079 and the Integrated Research Program for Advancing Climate
281 Models from the Ministry of Education, Culture, Sports, Science and Technology (MEXT),
282 Japan. T.M. acknowledges funding from the European Research Council grant #770765 and
283 European Union Horizon 2020 project #820829.

284

285 **Author Contributions**

286 M.W. designed the research and wrote the paper. H.T. conducted the MIROC large ensemble
287 experiments. J.-L.D., Y.K. and T.M. helped analysing the large ensemble simulations. All
288 authors discussed the results and commented on the manuscript.

289

290 **Additional information**

291 Supplementary information is available in the online version of the paper. Reprints and
292 permissions information is available online at www.nature.com/reprints. Correspondence and
293 requests for materials should be addressed to M.W. (hiro@aori.u-tokyo.ac.jp)

294

295 **Competing financial interests**

The authors declare that they have no competing financial interests.

Methods

Observed SST and ocean temperature data sets. To evaluate the observed SST trends, we analysed six different SST data sets from COBE-SST2³⁴, the Hadley Centre Sea Ice and SST version 1.1 (HadISST1.1)³⁵, the Hadley Centre SST version 3.1 (HadSST3.1)³⁶, the National Oceanic and Atmospheric Administration Extended Reconstructed SST version 5 (ERSSTv5)³⁷, the International Comprehensive Ocean-Atmosphere Data Set (ICOADS)³⁸, and the Kaplan Extended SST version 2³⁹. Two out of six data sets (HadSST3 and ICOADS) are uninterpolated analyses, i.e., without any spatial/temporal smoothing or interpolation, and may be more reliable than the interpolated analyses²². The data are not available after 2014 in COBE-SST2 and after 2016 in HadISST, but other four archives provide data available until 2019. Annual mean data for 1951–2010 were analysed for all the data sets, and the linear trend was calculated using annual anomalies relative to the 1971–2000 mean. We also used the gridded ocean temperature data compiled with COBE-SST2, having 14 vertical levels above 500 m⁴⁰.

Large ensemble historical simulations. For comparing the SST trend with observations and CMIP5 simulations, we analysed large ensemble (LE) climate simulations from Community Earth System Model Version 1 (CESM1)⁴¹, Max Planck Institute for Meteorology Earth System Model version 1.1 (MPI-ESM1.1)⁴², Model for Interdisciplinary Research on Climate version

6 (MRIOC6)⁴³, and Institut Pierre et Simon Laplace Earth System Model version 6A (IPSL-
CM6A-LR)⁴⁴. Each LE was generated by repeating climate simulations with the identical
external forcing but perturbing their initial conditions. The CESM1 LE has 40 members driven
by the CMIP5 historical and RCP8.5 forcing, and provide data publicly via the NCAR Large
Ensemble project. The MPI-ESM1.1 has 100 members driven by the CMIP5 forcing as well
and the data are open to the research community. We used the surface skin temperature data,
which is nearly the same as SST, from historical and RCP8.5 runs. LEs from other two,
MIROC6 including 50 members and IPSL-CM6-LR including 30 members, have recently been
produced using the CMIP6 version models and external forcing for the CMIP6 historical
simulation for 1850–2015⁴⁵. Some of the radiative forcing agents have been updated in CMIP6,
but the updates to models presumably have a larger impact on the simulations. Given that the
CMIP6-based LEs are not systematically different from the CMIP5-based LEs (Fig. 1b), we
treated the four LEs equally.

CMIP5 Models. We employed 27 climate model simulations from the CMIP5 archive¹³. We
used one member of the historical and three RCP experiments (RCP2.6, 4.5, and 8.5) for each
model. Annual mean anomalies, relative to the 1971–2000 mean, were first compiled on a
regular 2.5°×2.5° grid. The linear trends for 1951–2010 were then calculated using the annual
mean anomalies from historical runs for 1951–2005 combined with RCP4.5 simulations for
2006–2010. The results have been confirmed to change little when using other RCP simulations.
The models used in the present study are listed in Supplementary Table 1, wherein the ΔSST_{eq}

trend for 1951–2010 is also presented. Although historical and RCP4.5 simulations are available for more than 40 models^{3,5}, in this study we used only those models that provided subsurface ocean temperature data. We did not use GSAT for the RCP2.6 simulations by CMCC-CM and CESM1-BGC because the data were not available.

Metric for the SST pattern change in the tropical Pacific. The zonal SST gradient in the equatorial Pacific, $\Delta\text{SST}_{\text{eq}}$, is defined as the difference between the western Pacific (110°E–180°, 5°S–5°N) and the eastern Pacific (180°–80°W, 5°S–5°N), so that the positive change in $\Delta\text{SST}_{\text{eq}}$ represents a weakening of the zonal gradient. The eastern box is very similar to the Niño 3 region (150°W–90°W, 5°S–5°N) but only slightly wider. While anomaly in the western Pacific is often ignored in the ENSO theory assuming a radiative-convective equilibrium over the warm pool¹¹, the $\Delta\text{SST}_{\text{eq}}$ trend for 1951–2010 depends on SST changes in both regions (Supplementary Figs. 2 and 7).

Separation between interannual and decadal/interdecadal anomalies. Most of analyses are based on annual mean anomalies, but some of analyses were conducted after the anomalies were separated into high- and low-frequency components using a 10-year lowpass filter (Supplementary Figs. 5, 7-9, 11). Namely, filtered anomalies are regarded as low-frequency components that contain decadal/interdecadal variability and linear trends, whereas the residuals are referred to as high-frequency components representing the interannual variability.

Statistical significance and uncertainty estimates. We evaluated if the linear trend is statistically significant against fluctuations around the trend, by using the Student's t-test. The

result is shown where the pattern of linear trends is presented (Fig. 2, Supplementary Figs. 1, 6, and 10). For model ensembles, we also calculate the level of agreement across models, i.e., number of models that reveal the same sign of the trend. This is another measure of confidence, and more than 80% of models agree with the sign where the linear trend is statistically significant at the 95% level (Supplementary Fig. 6). For estimating uncertainty of the linear trend, we plotted either the 66% or 90% range for samples assuming a Gaussian distribution. We compared probability density functions of linear trends in each LE (Fig. 1b) with those estimated using a non-parametric least-squares cross validation technique⁴⁶, and confirmed that the Gaussian distribution is a good approximation (Supplementary Fig. 14). Small sample sizes for observations leads to a potential error in the Gaussian fitting, but we applied the same method for comparing with uncertainty ranges in GCMs.

Data and code availability

COBE-SST2, ERSSTv5, ICOADS, and Kaplan SST data sets are all available from the NOAA/OAR/ESRL PSD website (<https://www.esrl.noaa.gov/psd/data/gridded/>). Two other SST data sets compiled at the Met Office Hadley Centre are all available from <https://www.metoffice.gov.uk/hadobs/>. The CMIP5 model output analysed in this study is available from the Earth System Grid Federation server (<https://esgf-node.llnl.gov/search/cmip5/>). The CESM Large Ensemble project simulation output can be obtained from <http://www.cesm.ucar.edu/projects/community-projects/LENS/data-sets.html>. The MPI-ESM1.1 large ensemble data are available at the MPI Grand Ensemble project

376 website (<https://www.mpimet.mpg.de/en/grand-ensemble/>). The IPSL-CM6-LR and MIROC6
377 large ensemble simulation data are not accessible via website, but are available upon request.
378 The Fortran codes used for creating plots in this study can also be made available on request.
379

380 34. Hirahara, S., Ishii, M. & Fukuda, Y. Centennial-scale sea surface temperature analysis and
381 its uncertainty. *J. Climate* **27**, 57–75 (2014).

382 35. Rayner, N. A. et al. Global analyses of sea surface temperature, sea ice, and night marine
383 air temperature since the late nineteenth century. *J. Geophys. Res.* **108**, 4407 (2003).

384 36. Kennedy J. J., Rayner, N. A., Smith, R. O., Saunby, M. & Parker, D.E. Reassessing biases
385 and other uncertainties in sea-surface temperature observations since 1850 Part 1:
386 Measurement and sampling errors. *J. Geophys. Res.* **116**, D14103 (2011).

387 37. Huang, B. et al. Extended reconstructed sea surface temperature, version 5 (ERSSTv5):
388 Upgrades, validations, and intercomparisons. *J. Climate* **30**, 8179–8205 (2017).

389 38. Woodruff, S. D., Worley, S. J., Lubker, S. J. et al. ICOADS release 2.5: Extensions and
390 enhancements to the surface marine meteorological archive. *Int. J. Climatol.* **31**, 951–967
391 (2011).

392 39. Kaplan, A., Cane, M. A., Kushnir, Y. et al. Analyses of global sea surface
393 temperature 1856–1991. *J. Geophys. Res.* **103**, 18567–18589 (1998).

394 40. Ishii, M. & Kimoto, M. Reevaluation of historical ocean heat content variations with time-
395 varying XBT and MBT depth bias corrections. *J. Oceanogr.* **65**, 287–299 (2009).

41. Kay, J. E. et al. The Community Earth System Model (CESM) large ensemble project: A community resource for studying climate change in the presence of internal climate variability. *Bull. Amer. Meteor. Soc.* **96**, 1333–1349 (2015).
42. Maher, N. et al. The Max Planck Institute grand ensemble: Enabling the exploration of climate system variability. *J. Adv. Model. Earth Sys.* **11**, 1–21 (2019).
43. Tatebe, H. et al. Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity in MIROC6. *Geo. Model Dev.* **12**, 2727–2765 (2019).
44. Boucher, O. et al. Presentation and evaluation of the IPSL-CM6A-LR climate model. *J. Adv. Model. Earth Sys.* submitted (2019).
45. Eyring, V. et al. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.*, **9**, 1937–1958 (2016).
46. Kimoto, M. & Ghil, M. Multiple flow regimes in the Northern Hemisphere winter. Part I: Methodology and hemispheric regimes. *J. Atmos. Sci.* **50**, 2625–2643 (1993).

Figure captions

Figure 1 | Linear trends in the equatorial Pacific zonal SST gradient ($\Delta\text{SST}_{\text{eq}}$) during 1951–2010.

a, Trends in six observational data sets (black bars) and 27 CMIP5 models, arranged from strengthening (negative) to weakening (positive) $\Delta\text{SST}_{\text{eq}}$. The unit is K per century. Red bars indicate models showing the strengthening of $\Delta\text{SST}_{\text{eq}}$ consistent with the observations (S-models), whereas blue bars indicate models simulating large weakening trend opposite to the observations (W-models). **b**, Probability density function of $\Delta\text{SST}_{\text{eq}}$ trends in four large ensemble simulations: CESM1 (blue), IPSL-CM6-LR (green), MPI-ESM1.1 (pink), and MIROC6 (cyan). Numbers in parentheses indicate their ensemble size. Means and the 5–95% ranges for observations, CMIP5 models, and combined LEs are represented, respectively, by black, grey, and orange dots with bars at the left.

Figure 2 | Linear trends in the equatorial ocean temperature during 1951–2010.

The trend is calculated for the 5°S–5°N strip for **a**, observations, **b**, CMIP5 multi-model mean, **c**, CMIP5 S-model mean, and **d**, CMIP5 W-model mean. The unit is K per century. The grey shading represents the trend statistically significant at the 95% level. White contours indicate 19, 20, and 21 °C isotherms in the climatological states, measuring the depth of mean thermocline.

Figure 3 | Linear trends in $\Delta\text{SST}_{\text{eq}}$ for 50 years with a sliding window between 1951 and 2100.

a, Ensemble-mean values of the trend from the CMIP5 historical and RCP4.5 simulations by S-models (red), W-models (blue), and all models (grey) represented by squares and dashed curves, with error bars indicating the 66% range. The unit is K per century. The abscissa shows the end year for the 50-year segment. **b**, Same as **a**, but for the MPI-ESM1.1 100-member LE, in which ten S- and W-members are selected following the values of ΔSST_{eq} trend for 1951–2010.

Figure 4 | Linear trends in GSAT for 50 years with a sliding window between 1951 and 2100.

Ensemble-mean values of the trend from the CMIP5 S-models (red curve) and W-models (blue curve) and the 66% range (shading) for **a**, RCP2.6, **b**, RCP4.5, and **c**, RCP8.5 simulations. The abscissa shows the end year for the 50-year segment. The unit is K per century. The linear trends for 1951–2000 are calculated using the historical runs and are highlighted by squares.

Figure 1 | Linear trends in the equatorial Pacific zonal SST gradient (ΔSST_{eq}) during 1951–2010.

a, Trends in six observational data sets (black bars) and 27 CMIP5 models, arranged from strengthening (negative) to weakening (positive) ΔSST_{eq} . The unit is K per century. Red bars indicate models showing the strengthening of ΔSST_{eq} consistent with the observations (S-models), whereas blue bars indicate models simulating large weakening trend opposite to the observations (W-models). **b**, Probability density function of ΔSST_{eq} trends in four large ensemble simulations: CESM1 (blue), IPSL-CM6-LR (green), MPI-ESM1.1 (pink), and MIROC6 (cyan). Numbers in parentheses indicate their ensemble size. Means and the 5–95% ranges for observations, CMIP5 models, and combined LEs are represented, respectively, by black, grey, and orange dots with bars at the left.

Figure 2 | Linear trends in the equatorial ocean temperature during 1951–2010.

The trend is calculated for the 5°S–5°N strip for **a**, observations, **b**, CMIP5 multi-model mean, **c**, CMIP5 S-model mean, and **d**, CMIP5 W-model mean. The unit is K per century. The grey shading represents the trend statistically significant at the 95% level. White contours indicate 19, 20, and 21 °C isotherms in the climatological states, measuring the depth of mean thermocline.

Figure 3 | Linear trends in $\Delta S S T_{eq}$ for 50 years with a sliding window between 1951 and 2100.

a, Ensemble-mean values of the trend from the CMIP5 historical and RCP4.5 simulations by S-models (red), W-models (blue), and all models (grey) represented by squares and dashed curves, with error bars indicating the 66% range. The unit is K per century. The abscissa shows the end year for the 50-year segment. **b**, Same as **a**, but for the MPI-ESM1.1 100-member LE, in which ten S- and W-members are selected following the values of $\Delta S S T_{eq}$ trend for 1951–2010.

Figure 4 | Linear trends in GSAT for 50 years with a sliding window between 1951 and 2100.

Ensemble-mean values of the trend from the CMIP5 S-models (red curve) and W-models (blue curve) and the 66% range (shading) for **a**, RCP2.6, **b**, RCP4.5, and **c**, RCP8.5 simulations. The abscissa shows the end year for the 50-year segment. The unit is K per century. The linear trends for 1951–2000 are calculated using the historical runs and are highlighted by squares.