nature climate change

ARTICLES

https://doi.org/10.1038/s41558-021-01099-2

Robust decrease in El Niño/Southern Oscillation amplitude under long-term warming

Christopher W. Callahan[1,2], Chen Chen[3,4]✉, Maria Rugenstein[5,6], Jonah Bloch-Johnson[4,7], Shuting Yang[8] and Elisabeth J. Moyer[4]✉

El Niño/Southern Oscillation (ENSO) is the primary mode of interannual climate variability, and understanding its response to climate change is critical, but research remains divided on the direction and magnitude of that response. Some twenty-first-century simulations suggest that increased CO_2 strengthens ENSO, but studies suggest that on palaeoclimate timescales higher temperatures are associated with a reduced ENSO amplitude and a weaker Pacific zonal temperature gradient, sometimes termed a 'permanent El Niño'. Internal variability complicates this debate by masking the response of ENSO to forcing in centennial-length projections. Here we exploit millennial-length climate model simulations to disentangle forced changes to ENSO under transient and equilibrated conditions. On transient timescales, models show a wide spread in ENSO responses but, on millennial timescales, nearly all of them show decreased ENSO amplitude and a weakened Pacific zonal temperature gradient. Our results reconcile differences among twenty-first-century simulations and suggest that CO_2 forcing dampens ENSO over the long term.

El Niño/Southern Oscillation (ENSO) is the leading mode of interannual climate variability and has substantial socio-economic impacts, causing precipitation extremes in the Americas[1], changes in Pacific fish stocks[2] and droughts in Southeast Asia[3]. Understanding the response of ENSO to climate change is therefore of paramount importance for understanding climate impacts and adaptation options, but there is little consensus on the direction and magnitude of that response[4,5]. This forced response is difficult to predict because the feedbacks that drive ENSO are complex and multidirectional, balanced between positive dynamical and negative thermodynamical feedbacks, with substantial cloud-driven components[6,7]. Several studies using Coupled Model Intercomparison Project Phase 5 (CMIP5) models have suggested that ENSO strengthens in future high-CO_2 climates[8–10], but models can disagree[4–6,11] and some single-model studies suggest that ENSO weakens instead[12]. The presence of substantial unforced variability in ENSO means that resolution of this debate poses a challenge for existing modelling resources[13–17].

Despite disagreements over ENSO, climate models largely agree that under increased CO_2 the Eastern Pacific warms more than the Western Pacific, probably in tandem with a weakened Walker circulation, decreasing the zonal temperature gradient and making the equatorial Pacific more 'El Niño-like'[7,18–20]. Palaeoclimate studies also largely—though not universally—find that the warm, high-CO_2 Pliocene epoch is associated with a decreased Pacific zonal temperature gradient[21–23]. However, the relationship between mean Pacific sea surface temperatures (SSTs) and ENSO amplitude remains unclear[5,11]. Some studies hypothesize that a weakened zonal temperature gradient should dampen ENSO[24], but others hypothesize that it should strengthen it by reducing the barrier to convection in the Eastern Pacific[6,20], and some argue that El Niño-like variability may have been present in the Pliocene despite the reduced temperature gradient[25]. Moreover, the observed Pacific zonal temperature gradient has been strengthening in recent decades, rather than weakening, raising questions about whether models accurately simulate the tropical Pacific[26] or whether internal variability dominates the relatively short observational record[27].

In this study we exploit the Long Run Model Intercomparison Project (LongRunMIP), a new archive of 1,000-year or longer fully coupled climate model simulations[28], to distinguish the forced response of ENSO from internal variability and clarify the relationship between ENSO and the Pacific mean climate. We measure ENSO using quadratically detrended[9], area-averaged surface air temperature anomalies in the Niño3.4 region (Fig. 1a,b), and we define ENSO amplitude using the standard deviation of these anomalies (Fig. 1c)[16]. While some studies have suggested that different 'flavours' of ENSO may require different indices, or that its primary location may shift in future climates[10,29], our results are similar whether we use the Niño3 (Eastern Pacific) or Niño4 (Western Pacific) index (Supplementary Fig. 1), and the spatial structure of Pacific variability does not exhibit shifts to new centres of action outside these indices (Fig. 1a,b, Methods, Supplementary Table 1 and Supplementary Fig. 2). LongRunMIP simulations include scenarios of both abrupt and gradual CO_2 change with different CO_2 perturbations (Methods and Supplementary Table 2). To compare short- and long-term responses, we define a 'transient' period as years 50–150 in abrupt simulations and 1–140 in gradual, and consider the remainder of each simulation as 'equilibrium' (Methods and Fig. 1d,e). Results are not sensitive to this precise division (Methods and Supplementary Fig. 3).

In the equilibrium period, ENSO amplitude decreases in nearly all high-CO_2 simulations (Fig. 2) while ENSO frequency shows no clear changes (Supplementary Fig. 4). The only exceptions are simulations from MPI-ESM-1.2 and CNRM-CM6.1, which show slight

[1]Program in Ecology, Evolution, Environment and Society, Dartmouth College, Hanover, NH, USA. [2]Program in Environmental Sciences, Northwestern University, Evanston, IL, USA. [3]Centre for Climate Research, Singapore, Singapore. [4]Department of the Geophysical Sciences, University of Chicago, Chicago, IL, USA. [5]Department of Atmospheric Science, Colorado State University, Fort Collins, CO, USA. [6]Max Planck Institute for Meteorology, Hamburg, Germany. [7]National Centre for Atmospheric Science, University of Reading, Reading, UK. [8]Danish Meteorological Institute, Copenhagen, Denmark. ✉e-mail: chen_chen@nea.gov.sg; moyer@uchicago.edu

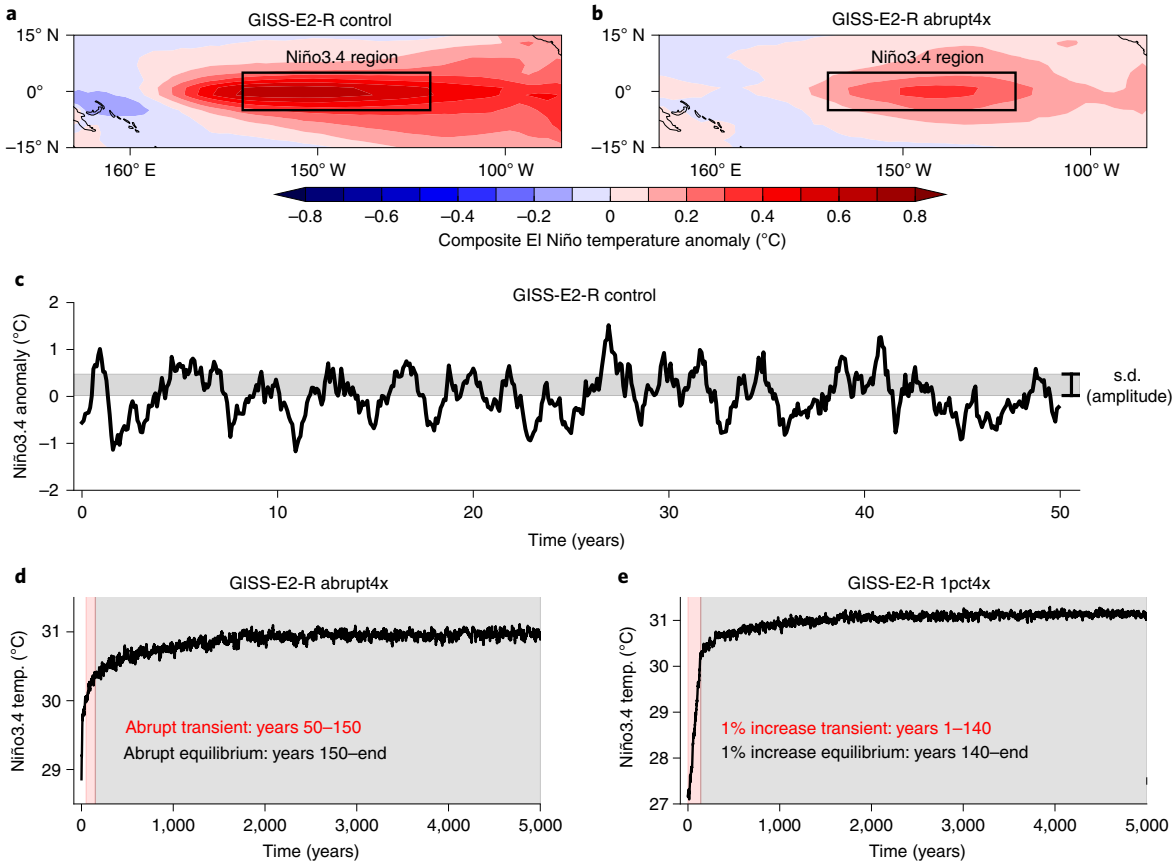

Fig. 1 | Methods used to analyse ENSO in LongRunMIP simulations. a,**b**, Composites of Pacific temperature anomalies during El Niño events, defined as the Niño3.4 index between its 75th and 90th percentiles (Methods) in the GISS-E2-R control (**a**) and GISS-E2-R abrupt4× (**b**) simulations. Pattern correlation between the two composites is 0.92. **c**, Example time series of Niño3.4 temperature anomalies; grey shading denotes s.d., which we define as ENSO amplitude. The figure shows a 50-year segment to illustrate the evolution of individual ENSO events, but in analysis we define amplitude using 100-year segments. **d**,**e**, Illustration of transient and equilibrium periods using Niño3.4 average temperature in the GISS-E2-R abrupt4× (**d**) and GISS-E2-R 1pct4× (**e**) simulations.

increases; in the eight other models, ENSO amplitude decreases. All changes are statistically significant according to a Wilcoxon rank-sum test on the distribution of amplitudes in 100-year overlapping segments (Methods). Declines in amplitude scale with CO_2 forcing, as would be expected given a causal relationship: in the three models with simulations of different forcing, larger CO_2 increases are monotonically associated with greater ENSO amplitude decline (Fig. 2). In CESM1.0.4 and HadCM3L, ENSO amplitude declines by 29 and 8% at 2×CO_2 and by 61 and 47% at 8×CO_2, respectively. Changes in equilibrium ENSO frequency, on the other hand, show no consistent pattern (Supplementary Fig. 4). We define ENSO occurrence based on threshold exceedance for the Niño3.4 temperature anomaly, with the threshold defined as the 75th–90th percentiles for control and high-CO_2 simulations separately (Methods). If frequency is defined using an absolute threshold instead[8,9], an amplitude decline must be associated with a corresponding reduction in the number of identified El Niño events (Supplementary Fig. 4). With a percentile-based definition, the robust weakening in the magnitude of El Niño events in LongRunMIP simulations is not accompanied by a clear change in their number per year (see Supplementary Table 3 for ENSO amplitude and frequency in each simulation). Note that our results are not inconsistent with previous studies that defined El Niño using precipitation thresholds and found future increases in frequency[8]. These studies used absolute rather than fractional thresholds, which are crossed more frequently when precipitation means increase (Supplementary Fig. 5). LongRunMIP models also show increases in extreme El Niño precipitation (Supplementary Fig. 5), consistent with findings that in a warmer climate, Eastern Pacific precipitation variability can increase even while SST fluctuations decrease[8,30].

In the transient period, changes in ENSO amplitude differ substantially in magnitude and even in sign but, in nearly all simulations, the transient change is smaller in magnitude than the equilibrium change (Figs. 2 and 3). The ensemble behaviour appears well explained by a combination of (1) partial evolution towards an equilibrium response and (2) unforced variation in the initial transient-period amplitude (Fig. 3 and Supplementary Figs. 6 and 7). CNRM-CM6.1 and GFDL-CM3 are the exceptions, with transient increases outside the control simulation range. Internal variation in ENSO amplitude can be substantial in millennial-scale simulations. We quantify these long-timescale changes by calculation of ENSO amplitudes in overlapping 100-year segments of each control run, and computing the 5th–95th percentile range of their distribution. All LongRunMIP models exhibit substantial fluctuation in ENSO amplitude over time: individual 100-year amplitudes in the average model can be as low as 15% below its long-term average or as high as 15% above it (Fig. 3 and Supplementary Figs. 6 and 7). These results are consistent with estimates of ENSO internal variability derived from initial-condition ensembles of shorter simulations[15–17].

The diversity in transient ENSO responses therefore appears to result from both variations in the pace of model equilibration

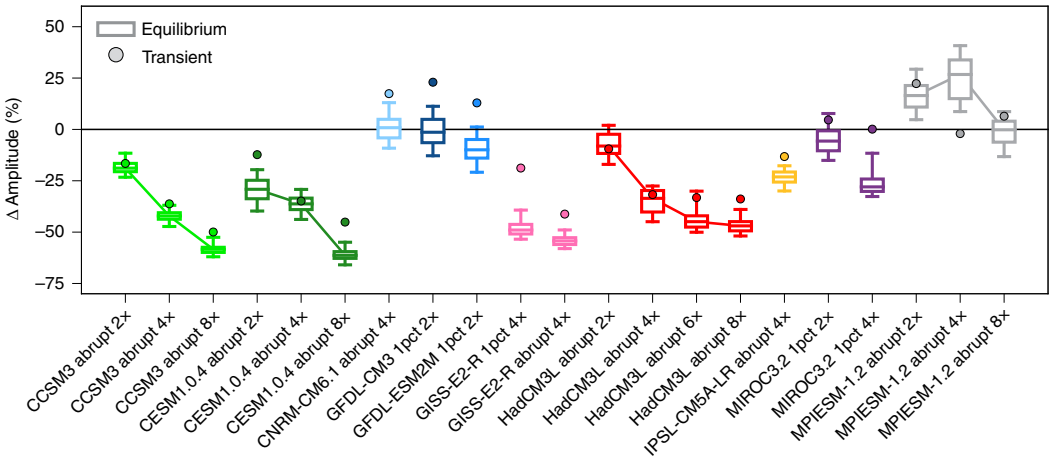

Fig. 2 | Percentage changes in ENSO amplitude after a sustained increase in CO_2. Simulations from the same model are shown in the same colour. Box plots show distributions of changes in 100-year running segments in each high-CO_2 equilibrium period, as compared to the complete control run. Middle line denotes the median, box boundaries show the interquartile range (25th–75th percentiles) and whiskers extend to the fifth and 95th percentiles of the data. Circles indicate changes in the transient period. All changes are statistically significant according to a Wilcoxon rank-sum test (Methods), and all models other than MPI-ESM-1.2 and CNRM-CM6.1 show declines in ENSO amplitude in all scenarios.

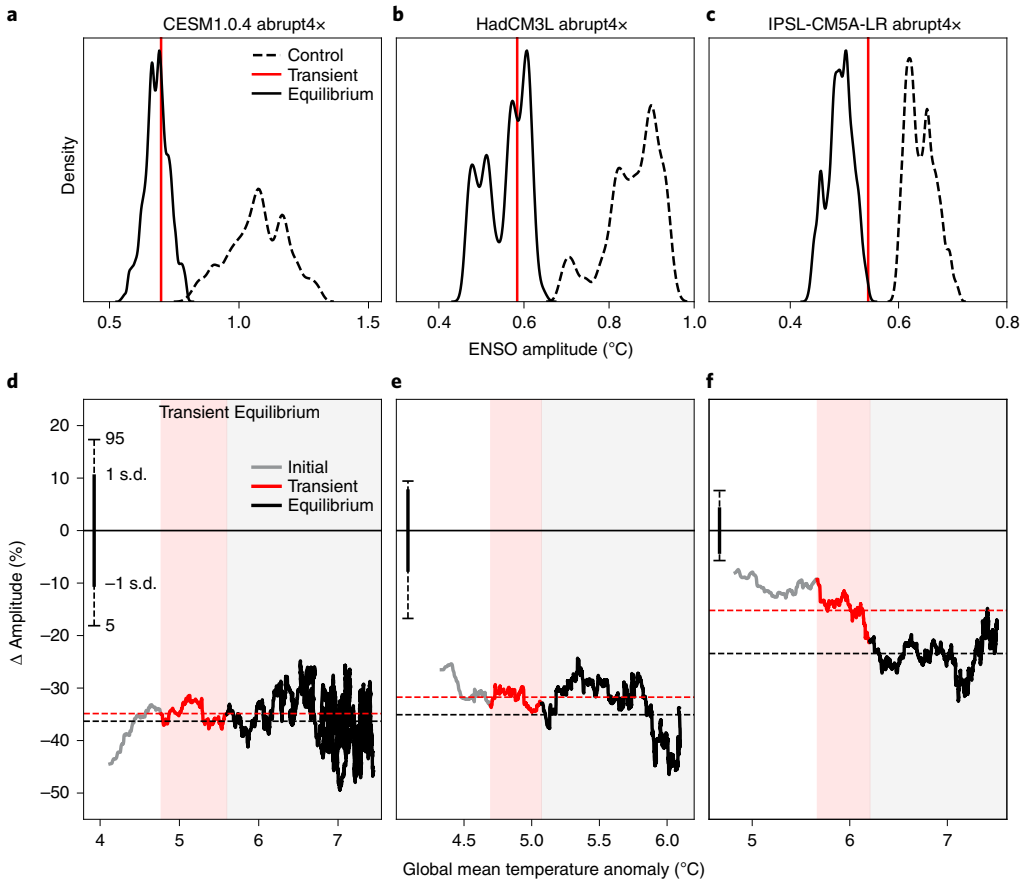

Fig. 3 | Evolution of ENSO amplitude under increased CO_2 for three representative LongRunMIP models. a–c, Distributions of 100-year overlapping ENSO amplitudes in control simulation (dashed black) and 4×CO_2 equilibrium periods (solid black) in the CESM1.0.4 abrupt4× (**a**), HadCM3L abrupt4× (**b**) and IPSL-CM5A-LR abrupt4× (**c**) simulations. Solid red vertical line denotes the 4×CO_2 transient period. **d–f**, Changes in ENSO amplitude plotted against the evolving global mean surface temperature anomaly (Methods) for the CESM1.0.4 abrupt4× (**d**), HadCM3L abrupt4× (**e**) and IPSL-CM5A-LR abrupt4× (**f**) simulations. Dashed horizontal lines denote the mean amplitude change for transient (red) and equilibrium (black) periods. Bars on the left show the distribution of 100-year overlapping ENSO amplitudes in the control simulation, with the 5th–95th percentiles and ±1 s.d. marked. Scales on the *x* axes reflect the differing climate sensitivities of the models, and temperature anomalies begin above zero because the first 100-year mean encompasses several degrees of temperature change. Transient and equilibrium responses are similar when ENSO equilibration is rapid relative to warming (**a**,**d**) but otherwise diverge (**b**,**e**,**c**,**f**). Supplementary Figs. 6 and 7 show all simulations.

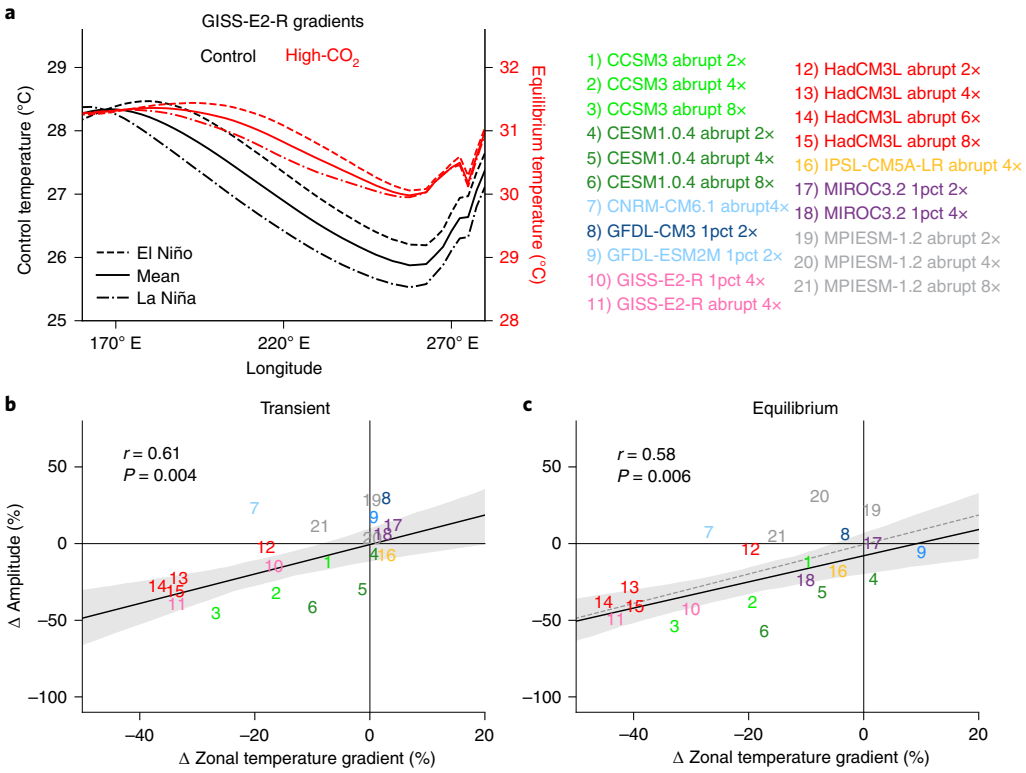

Fig. 4 | Relationship between zonal temperature gradient change and ENSO amplitude change. a, Representative longitudinal temperature profiles from an abrupt4× experiment. Black and red solid lines denote the control and abrupt4× mean gradient, respectively; dashed lines denote the gradient averaged over all El Niño events; and dashed-dotted lines denote the gradient averaged over all La Niña events (Methods). Note that, while the *y* axes have different absolute temperatures, they use the same relative temperature scale to enable comparison between the two gradients. High-CO_2 mean zonal temperature gradient weakens even more strongly than the El Niño-like gradient in the control simulations. **b**,**c**, Relationship between zonal temperature gradient change and ENSO amplitude change for all simulations in the transient (**b**) and equilibrium (**c**) periods. Middle lines represent the least-squares linear regression lines, with outer shading showing 95% confidence intervals. Dashed line in **c** is the regression line from **b** overlaid on the regression in **c** for visual comparison. Supplementary Fig.13 shows profiles from **a** for all simulations.

and internal variability in ENSO amplitude on centennial timescales. In around half of the LongRunMIP simulations, the transient ENSO amplitude response lies within the 5th–95th percentile envelope of the millennial-scale unforced variability defined above. In the other half, the transient response is outside the envelope of variability but in the same direction as the equilibrium response. Differences in these models can be explained largely by the differing rates with which ENSO responses equilibrate after an increase in CO_2. In models where ENSO equilibrates rapidly relative to global mean warming (for example, CESM1.0.4; Fig. 3a,d), the transient and equilibrium responses are similar; in models where equilibration is more gradual (for example, HadCM3L and IPSL-CM5A-LR; Fig. 3b,c,e,f), ENSO evolution continues beyond the transient period and transient and equilibrium responses differ.

Explaining this diversity helps reconcile disparate findings from previous studies, by demonstrating that disparities may result from differences in equilibration and internal variability rather than from any fundamental difference in ENSO-related physical processes[13,15,16]. For example, the LongRunMIP GFDL-ESM2M simulation shows a transient increase in ENSO amplitude but an equilibrium period decrease (Fig. 2), consistent with a previous paper that found an amplitude decrease in several hundred years of an abrupt CO_2 change scenario[12]. In addition, the transient increase in our GFDL-ESM2M simulation lies within the control simulation variability, so it may be a manifestation of unforced variability; large amounts of unforced variability can confound identification of transient changes (Supplementary Figs. 6 and 7). More broadly, other recent papers have used SST-based metrics to project an increase in ENSO variability[10] but substantial intermodel disagreement remains[31], at least partly due to internal variability in twenty-first-century simulations[32]. These findings are consistent with our results showing model disagreement in transient changes in ENSO amplitude (Fig. 2), given the substantial multidecadal fluctuations in ENSO amplitude (Fig. 3).

Previous studies suggest a relationship between changes in ENSO and those in the Pacific zonal temperature gradient, although the nature of this relationship is debated[5–7,24] and may be state dependent[33]. In all but one LongRunMIP simulation, warming under high-CO_2 conditions is larger in the Eastern Pacific than in the Western Pacific, resulting in a reduced zonal temperature gradient, sometimes termed 'El Niño-like warming'[18] (Fig. 4 and Supplementary Figs. 8 and 9) (the exception is GFDL-ESM2M, as noted previously[34]). Pacific gradients can decline to less than half of their control values to the point that, in many cases, the future Pacific actually exceeds a permanent El Niño: that is, the mean gradient in the forced simulation is weaker even than El Niño conditions in the control run (Supplementary Fig. 10). Potential drivers of this weakened gradient include changes in cloud albedo and amount[18], greater evaporative cooling in the Western Pacific[35] and a slowdown in zonal atmospheric circulation[19]. In almost all LongMIP simulations the Walker circulation does weaken (Supplementary Fig. 8), consistent with a mechanism in which decreased circulation weakens the trade winds and depresses upwelling, reducing the zonal temperature gradient[7]. The zonal temperature gradient

response is correlated with the ENSO amplitude response across simulations, and in both transient and equilibrium periods: stronger decreases in mean temperature gradient are associated with stronger decreases in amplitude (Fig. 4b,c). The relationships in the transient ($P=0.004$) and equilibrium ($P=0.006$) periods are also similar: the regression line from the transient period lies within the uncertainty bands of the equilibrium period. These results remain qualitatively unchanged if we use only one simulation from each model (Supplementary Fig. 11).

Observations have suggested that the zonal temperature gradient has been strengthening, which raises some concerns about the models[26]. Model representations of ENSO have been shown to be biased in studies of, for example, relative humidity[26], interbasin warming contrasts[36] or the equatorial undercurrent[37]. However, observed trends in the zonal temperature gradient are well within the range of unforced variability from LongRunMIP control simulations (Methods and Supplementary Fig. 12). It remains under debate whether the observed trend is due to internal variability[27,38], to a forced response not captured by the models[26] or even to observational error[39]. Conclusively settling the debate is outside the scope of this work, but LongRunMIP control simulations show that it is at least plausible that the observed trend is a manifestation of internal variability and not inconsistent with the forced response.

Our results provide insights into the mechanisms controlling ENSO changes. The oscillation of ENSO temperature anomalies results from a balance between a positive dynamical feedback (the Bjerknes feedback, in which higher Eastern Pacific SSTs reduce easterly wind stress and further allow heat to build up) and a negative thermodynamic feedback (in which SST anomalies in turn increase heat flux from the surface)[5,40,41]. Although no solid theory predicts the net effect of changes in these feedbacks in a warmer world[6], previous studies have suggested that a decreased zonal temperature gradient should be associated with a stronger Bjerknes feedback that amplifies ENSO: increased upper-ocean stratification and a shoaling of the thermocline would increase dynamical wind–ocean coupling[5,7,10,42]. The LongRunMIP behaviour might seem in conflict with this mechanism but we find that, in all LongRunMIP models, the thermocline does indeed shoal, flatten latitudinally and sharpen vertically as predicted (Supplementary Fig. 14). The issue is therefore not a conflict of mechanism but rather that some additional process outweighs this dynamical effect, producing a net reduction in ENSO amplitude.

In the LongRunMIP models, the negative thermodynamic feedback does indeed strengthen. That is, the forced runs show a stronger additional net export of energy from the ocean surface per degree Niño3.4 temperature anomaly, with about half of the effect coming from increased latent heat export and half from increased cloudiness that reduces incoming short-wave radiation (consistent with refs. [40,41,43,44]). For the abrupt4× CO_2 forcing ensemble mean, the relationship between additional net heat flux and temperature anomalies is 6.2 $Wm^{-2}K^{-1}$, >70% stronger than in the control runs (Supplementary Fig. 15). Reduced short-wave radiation and increased latent heat flux make nearly equal contributions at 3.3 and 3.2 $Wm^{-2}K^{-1}$, respectively, 72 and 59% stronger than in control runs (Supplementary Figs. 16 and 17). The thermodynamic feedback strengthens in all LongRunMIP models except for GISS-E2-R, which exhibits other anomalies such as a reversed Walker circulation and biased local radiative feedbacks[45]. Increased heat export therefore appears to compensate for strengthened dynamical feedbacks to dampen ENSO amplitude. However, because LongRunMIP data are not sufficient for rigorous analysis of the strength of wind–ocean coupling in our simulations, it is difficult to quantitatively balance these two feedbacks. We leave further analysis of these feedbacks, as well as more recent proposed processes such as ENSO 'self-modulation'[46], for future work.

The combination of a weaker ENSO amplitude and a more homogenous tropical mean state seen in LongRunMIP models is supported by palaeoclimate evidence. These conditions are analogous to the permanent El Niño that may have characterized the Pliocene epoch[21,22,25]. The LongRunMIP experiments do involve CO_2 concentrations higher than those likely in the Pliocene[47], but our results demonstrate that Pliocene-like conditions are at least physically plausible in a warmer world.

These results demonstrate the necessity of millennial-scale simulations of coupled climate models for understanding changes in low-frequency phenomena such as ENSO. Previous studies have argued or shown that unforced variability in ENSO characteristics confounds studies of its change in relatively short simulations[13,15,16,48]. We confirm here that centennial-scale unforced variability is comparable to projected changes under a wide range of CO_2 forcing, providing a major complicating factor to short-term analyses. Millennial-scale climate simulations allow us to diagnose the long-term forced effect of CO_2 changes: a robust long-term decline in ENSO amplitude in tandem with a flatter mean zonal temperature gradient, potentially due to enhanced thermodynamic damping. For understanding the human impacts of climate change, the most relevant changes are those in the twenty-first century under relatively modest CO_2 increases, and it remains unclear whether models can guide reliable forecasts of ENSO changes on decadal to centennial timescales[14]. Our results do suggest that the observed strengthening of the Pacific zonal temperature gradient is consistent with internal variability[27,38], meaning that these millennial-scale simulations can usefully inform short-term interpretation of the observational record. Regardless, our results demonstrate a consistent long-term weakening of ENSO and lend support to the possibility of a permanent El Niño in a high-CO_2 world.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41558-021-01099-2.

Received: 21 October 2019; Accepted: 11 June 2021;
Published online: 30 August 2021

References

1. Cayan, D. R., Redmond, K. T. & Riddle, L. G. ENSO and hydrologic extremes in the Western United States. *J. Clim.* **12**, 2881–2893 (1999).
2. Lehodey, P., Bertignac, M., Hampton, J., Lewis, A. & Picaut, J. El Niño Southern Oscillation and tuna in the western Pacific. *Nature* **389**, 715–718 (1997).
3. Harger, J. R. E. ENSO variations and drought occurrence in Indonesia and the Philippines. *Atmos. Environ.* **29**, 1943–1955 (1995).
4. McPhaden, M. J., Zebiak, S. E. & Glantz, M. H. ENSO as an integrating concept in Earth science. *Science* **314**, 1740–1745 (2006).
5. Philip, S. & Oldenborgh, G. J. V. Shifts in ENSO coupling processes under global warming. *Geophys. Res. Lett.* **33**, L11704 (2006).
6. Cai, W. et al. ENSO and greenhouse warming. *Nat. Clim. Change* **5**, 849–859 (2015).
7. Collins, M. et al. The impact of global warming on the tropical Pacific Ocean and El Nino. *Nat. Geosci.* **3**, 391–397 (2010).
8. Cai, W. et al. Increasing frequency of extreme El Niño events due to greenhouse warming. *Nat. Clim. Change* **4**, 111–116 (2014).
9. Cai, W. et al. Increased frequency of extreme La Niña events under greenhouse warming. *Nat. Clim. Change* **5**, 132–137 (2015).
10. Cai, W. et al. Increased variability of eastern Pacific El Niño under greenhouse warming. *Nature* **564**, 201–206 (2018).
11. Rashid, H. A., Hirst, A. C. & Marsland, S. J. An atmospheric mechanism for ENSO amplitude changes under an abrupt quadrupling of CO_2 concentration in CMIP5 models. *Geophys. Res. Lett.* **43**, 1687–1694 (2016).
12. Kohyama, T. & Hartmann, D. L. Nonlinear ENSO warming suppression. *J. Clim.* **30**, 4227–4251 (2017).

13. Wittenberg, A. T. Are historical records sufficient to constrain ENSO simulations? *Geophys. Res. Lett.* **36**, L12702 (2009).
14. Chen, C., Cane, M. A., Wittenberg, A. T. & Chen, D. ENSO in the CMIP5 simulations: life cycle, diversity, and responses to climate change. *J. Clim.* **30**, 775–801 (2017).
15. Maher, N., Matei, D., Milinski, S. & Marotzke, J. ENSO change in climate projections: forced response or internal variability? *Geophys. Res. Lett.* **45**, 11390–11398 (2018).
16. Sun, C. et al. Uncertainties in simulated El Niño–Southern Oscillation arising from internal climate variability. *Atmos. Sci. Lett.* **19**, e805 (2018).
17. Zheng, X.-T., Hui, C. & Yeh, S.-W. Response of ENSO amplitude to global warming in CESM large ensemble: uncertainty due to internal variability. *Clim. Dyn.* **50**, 4019–4035 (2018).
18. Meehl, G. A. & Washington, W. M. El Niño-like climate change in a model with increased atmospheric CO_2 concentrations. *Nature* **382**, 56–60 (1996).
19. Vecchi, G. A. & Soden, B. J. Global warming and the weakening of the tropical circulation. *J. Clim.* **20**, 4316–4340 (2007).
20. An, S.-I., Kug, J.-S., Ham, Y.-G. & Kang, I.-S. Successive modulation of ENSO to the future greenhouse warming. *J. Clim.* **21**, 3–21 (2008).
21. Wara, M. W., Ravelo, A. C. & Delaney, M. L. Permanent El Niño-like conditions during the Pliocene warm period. *Science* **309**, 758–761 (2005).
22. Fedorov, A. V. et al. The Pliocene paradox (mechanisms for a permanent El Niño). *Science* **312**, 1485–1489 (2006).
23. Zhang, Y. G., Pagani, M. & Liu, Z. A 12-million-year temperature history of the tropical Pacific Ocean. *Science* **344**, 84–87 (2014).
24. Yang, H. & Zhang, Q. Anatomizing the ocean's role in ENSO changes under global warming. *J. Clim.* **21**, 6539–6555 (2008).
25. Ivany, L. C., Brey, T., Huber, M., Buick, D. P. & Schone, B. R. El Niño in the Eocene greenhouse recorded by fossil bivalves and wood from Antarctica. *Geophys. Res. Lett.* **38**, L16709 (2011).
26. Seager, R. et al. Strengthening tropical Pacific zonal sea surface temperature gradient consistent with rising greenhouse gases. *Nat. Clim. Change* **9**, 517–522 (2019).
27. Chung, E.-S. et al. Reconciling opposing Walker circulation trends in observations and model projections. *Nat. Clim. Change* **9**, 405–412 (2019).
28. Rugenstein, M. et al. LongRunMIP: motivation and design for a large collection of millennial-length AOGCM simulations. *Bull. Am. Meteorol. Soc.* **100**, 2551–2570 (2019).
29. McPhaden, M. J., Lee, T. & McClurg, D. El Niño and its relationship to changing background conditions in the tropical Pacific Ocean. *Geophys. Res. Lett.* **38**, L15709 (2011).
30. Stevenson, S., Wittenberg, A. T., Fasullo, J., Coats, S. & Otto-Bliesner, B. Understanding diverse model projections of future extreme El Niño. *J. Clim.* **34**, 449–464 (2020).
31. Wang, B. et al. Historical change of El Niño properties sheds light on future changes of extreme El Niño. *Proc. Natl Acad. Sci. USA* **116**, 22512–22517 (2019).
32. Ng, B., Cai, W., Cowan, T. & Bi, D. Impacts of low-frequency internal climate variability and greenhouse warming on the El Niño-Southern Oscillation. *J. Clim.* **34**, 2205–2218 (2020).
33. Wyman, D. A., Conroy, J. L. & Karamperidou, C. The tropical pacific ENSO–mean state relationship in climate models over the last millennium. *J. Clim.* **33**, 7539–7551 (2020).
34. Kohyama, T., Hartmann, D. L. & Battisti, D. S. La Niña-like mean-state response to global warming and potential oceanic roles. *J. Clim.* **30**, 4207–4225 (2017).
35. Knutson, T. R. & Manabe, S. Time-mean response over the tropical Pacific to increased CO_2 in a coupled ocean–atmosphere model. *J. Clim.* **8**, 2181–2199 (1995).
36. Zhang, L. & Karnauskas, K. B. The role of tropical interbasin SST gradients in forcing Walker circulation trends. *J. Clim.* **30**, 499–508 (2017).
37. Coats, S. & Karnauskas, K. A role for the equatorial undercurrent in the ocean dynamical thermostat. *J. Clim.* **31**, 6245–6261 (2018).
38. Olonscheck, D., Rugenstein, M. & Marotzke, J. Broad consistency between observed and simulated trends in sea surface temperature patterns. *Geophys. Res. Lett.* **47**, e2019GL086773 (2020).
39. Chemke, R. & Polvani, L. M. Opposite tropical circulation trends in climate models and in reanalyses. *Nat. Geosci.* **12**, 528–532 (2019).
40. Lloyd, J., Guilyardi, E., Weller, H. & Slingo, J. The role of atmosphere feedbacks during ENSO in the CMIP3 models. *Atmos. Sci. Lett.* **10**, 170–176 (2009).
41. Lloyd, J., Guilyardi, E. & Weller, H. The role of atmosphere feedbacks during ENSO in the CMIP3 models. Part III: the shortwave flux feedback. *J. Clim.* **25**, 4275–4293 (2012).
42. Wang, G. et al. Continued increase of extreme El Niño frequency long after 1.5 C warming stabilization. *Nat. Clim. Change* **7**, 568–572 (2017).
43. Guilyardi, E. et al. Atmosphere feedbacks during ENSO in a coupled GCM with a modified atmospheric convection scheme. *J. Clim.* **22**, 5698–5718 (2009).
44. Middlemas, E. A., Clement, A. C., Medeiros, B. & Kirtman, B. Cloud radiative feedbacks and El Niño–Southern Oscillation. *J. Clim.* **32**, 4661–4680 (2019).
45. Bloch-Johnson, J., Rugenstein, M. & Abbot, D. S. Spatial radiative feedbacks from internal variability using multiple regression. *J. Clim.* **33**, 4121–4140 (2020).
46. Cai, W. et al. Butterfly effect and a self-modulating El Niño response to global warming. *Nature* **585**, 68–73 (2020).
47. Tierney, J. E., Haywood, A. M., Feng, R., Battacharya, T. & Otto-Bleisner, B. L. Pliocene warmth consistent with greenhouse gas forcing. *Geophys. Res. Lett.* **46**, 9136–9144 (2019).
48. Stevenson, S., Fox-Kemper, B., Jochum, M., Rajagopalan, B. & Yeager, S. G. ENSO model validation using wavelet probability analysis. *J. Clim.* **23**, 5540–5547 (2010).

Methods

The LongRunMIP archive. The LongRunMIP is a new archive of millennial-length, fully coupled climate model simulations collected to characterize the response of the climate system to CO_2 over the course of thousands of years[28]. The following LongRunMIP models are included in our analysis: CCSM3[49,50], CESM1.0.4 (refs. [51–53]), CNRM-CM6-1 (ref. [54]), GFDL-CM3 (refs. [55,56]), GFDL-ESM2M[56,57], GISS-E2-R[58–61], HadCM3L[62,63], IPSL-CM5A-LR[64], MIROC3.2 (refs. [65–67]) and MPI-ESM-1.2 (refs. [68,69]). While other models are included in the LongRunMIP archive, we select the ten models for which monthly mean surface air temperature and sea level pressure are available for at least 1,000 years after an increase in CO_2 concentrations.

The most common forcing experiment in our sample of the LongRunMIP is an instantaneous quadrupling of CO_2 (abrupt4×), but there are also several abrupt2×, abrupt6× and abrupt8× experiments and several ramps of an annual increase of 1% CO_2 (ppm) to the 2× or 4× CO_2 level. Each model has a control simulation and at least one high-CO_2 simulation, for a total of 21 high-CO_2 simulations and ten control runs; see Supplementary Table 2 for the models, experiments and lengths that we used. Each simulation in our subset of the LongRunMIP is at least 1,000 years in length, with the exception of the MIROC3.2 control run (681 years).

The goal of this analysis of LongRunMIP is not direct comparison of our results with those from studies using the CMIP5 archive, since we include a range of forcings (for example, abrupt8×CO_2) that are not included in CMIP5, and the model versions are slightly different in some cases. In addition, these experiments are not meant to be compared with experiments using representative concentration pathways (RCPs), because most of our experiments use substantially higher CO_2 concentrations than those of the RCPs. The purpose of the LongRunMIP is to distinguish the short- and long-term responses of the Earth system to increased CO_2, and to draw inferences about climate projections based on this unique ability to understand millennial-scale responses and variability.

The models we chose are not independent. However, they sample well from the CMIP5 archive and include models from multiple countries and modelling centres (for example, the National Center for Atmospheric Research, Goddard Institute for Space Studies, Hadley Centre for Climate Prediction and Research and Max Planck Institute). In addition, the inclusion of multiple types of experiment from the same model allows us to examine the sensitivity of our results to the forcing type. In particular, GISS-E2-R abrupt4× and 1pct4× have very similar equilibrium amplitude changes but very different transient amplitude changes.

Spatial indices. We analyse monthly mean surface air temperature (TAS) from each model as a proxy for sea surface temperature (SST), because TAS is more consistently available on monthly timescales in the LongRunMIP archive. Thermodynamic exchange between the atmosphere and surface ocean is sufficiently efficient to render surface air temperature an effective representation of SST[20]. We use the Niño3.4 index (Fig. 1a,b), area-averaged temperature anomalies from 5° S–5° N and 120° W–170° W, to measure ENSO amplitude. We calculate anomalies by removing the long-term monthly mean from each value within each simulation, thus eliminating mean shifts and the seasonal cycle, and we quadratically detrend each time series[9].

We also use surface air temperature to describe the zonal temperature gradient in the Equatorial Pacific, which we define as the difference between the area-averaged (non-anomalized) temperature in the Niño4 region (5° S–5° N, 160° E–150° W) and the area-averaged temperature in the Niño3 region (5° S–5° N, 150° W–90° W). An increase/positive change in the zonal temperature gradient means enhanced warming in the Western Pacific warm pool (La Niña-like warming), while a decrease/negative change in the temperature gradient means enhanced warming in the Eastern Pacific cold tongue (El Niño-like warming). We follow ref. [70] in describing the zonal Pacific atmospheric circulation (Walker circulation) by using the difference in sea level pressure (SLP) between the Eastern Pacific (5° S–5° N, 160° W–80° W) and Western Pacific (5° S–5° N, 80° E–160° E). An increase in the SLP gradient implies increased convection in the Western Pacific relative to the Eastern Pacific, and therefore a strengthened Walker circulation, while a decrease in the SLP gradient implies a weakened Walker circulation.

Spatial correlation analysis. To understand how increases in CO_2 affect the spatial structure of variability in the Pacific, we map the standard deviation of temperature across the Pacific (15° S–15° N, 150° E–80° W) after smoothing the temperature time series at each location with a 12-month running mean to reduce the influence of the seasonal cycle and other non-ENSO variability. We then compute pattern correlations between the standard deviation map in each forced simulation relative to its corresponding control simulation (Supplementary Table 1) using centred pattern correlations, meaning that the spatial average is removed within each simulation and only deviations from the spatial average are compared. Large values of this correlation indicate similar spatial patterns of variability under control and high-CO_2 conditions.

Indeed, many LongRunMIP simulations have high (>0.9) pattern correlation values (Supplementary Table 1). However, some (for example, GISS-E2-R and HadCM3L) do not. In Supplementary Fig. 2 we show two examples of simulations with high pattern correlations and four with smaller correlations. Importantly, even in simulations with relatively smaller pattern correlations (<0.9), variability decreases across the Pacific under increased CO_2 and there is no centre of action for El Niño outside the Niño3, Niño3.4 and Niño4 regions (a location of small variability appears south of these regions in HadCM3L abrupt8×, but the magnitude of this variability is small compared to the control simulation). These results support our use of these spatial indices to track changes in ENSO variability.

ENSO amplitude. We define ENSO amplitude as the standard deviation of the detrended time series of Niño3.4 temperature anomalies[16] (Fig. 1c). To define changes in amplitude between each control and high-CO_2 simulation, we take the 100-year running standard deviation of Niño3.4 anomalies in both experiments and then calculate the percentage difference between each value in the high-CO_2 experiment and the mean amplitude value from the control.

El Niño events and ENSO frequency. We quantify changes in ENSO frequency by constructing a measure of El Niño event occurrence in four steps: (1) we define the threshold for a 'typical' El Niño event as the mean of the 75th–90th percentile Niño3.4 temperature anomalies in each simulation (although our results are not sensitive to this definition); (2) we smooth the data with a 12-month running mean to avoid multiple transitions between El Niño and non-El Niño states within a single event; (3) we calculate the number of times in each century that the Niño3.4 anomaly transitions from below the threshold to greater than or equal to it, and we define each transition as one 'event'; and (4) we normalize this value by calculating a percentage change in frequency for each century of the forced simulation relative to the number of events per century in the control.

We perform this calculation twice: first by defining the temperature anomaly threshold separately for the control and high-CO_2 simulations (Supplementary Fig. 4), and second by defining it solely based on the control simulation and applying that control definition to the high-CO_2 simulation (Supplementary Fig.4). The first metric fully decouples shifts in ENSO frequency from changes in the overall probability distribution of temperature anomalies, and therefore provides an independent metric to determine changes in the number of El Niño events. The second metric conflates changes in ENSO amplitude and those in ENSO frequency: a decrease in standard deviation by definition implies a narrowing of the probability distribution of temperature anomalies, and therefore a reduced likelihood of crossing thresholds defined solely by the control simulation. Results from this metric confirm our findings of a decreased amplitude and narrower distribution of temperature anomalies.

Statistical analysis. We determine whether differences between control and high-CO_2 simulations are statistically significant by comparison of two distributions: 100-year running standard deviations in the case of ENSO amplitude and 100-year running means in the case of the zonal temperature and SLP gradients. We test this difference with a non-parametric Wilcoxon rank-sum test (also known as a Mann–Whitney U-test), which tests the null hypothesis that the two samples are from the same underlying distribution by comparison of the relative ranks of the two samples when combined into one distribution. We use a non-parametric test since the distributions of ENSO amplitudes are often non-normal (Supplementary Fig. 6), although we find similar results with a Student's t-test. Significance is assessed at the 5% level ($P<0.05$) in all cases.

We use Pearson's r as the correlation coefficient in Fig. 4 ($n=21$, degrees of freedom = 19), again assessing significance at the 5% level. However, because Pearson's r assumes that the data are normally distributed, we recompute the correlations in Fig. 4 using Spearman's rank correlation (ρ, a non-parametric test) and we again find that the two relationships are statistically significant at the 5% level (transient $P=0.005$, equilibrium $P=0.008$).

Distinguishing short- and long-term responses. We divide each high-CO_2 simulation into a transient and an equilibrium period to quantify the time dependence of changes due to increased CO_2 (Fig. 1d,e). To best compare changes in ENSO measured from LongRunMIP with those measured by studies using the CMIP5 archive, we define the transient period according to the length of the typical CMIP5 abrupt and 1% ramp experiments. Thus, in simulations with a 1% ramp of CO_2, we define the transient period as the first 140 years of the high-CO_2 simulation. In simulations with an abrupt increase in CO_2, we define the transient period as years 50–150 (we exclude the first 50 years to control for rapid adjustments to CO_2)[11]. In both cases we define the equilibrium period as the remainder of the simulation after the transient period. We also test the sensitivity of our results to these definitions by, for example, defining the transient period as various percentages (for example, 75%) of the total global mean surface temperature anomaly in each model, or shifting the break between transient and equilibrium periods. The results are qualitatively unchanged (Supplementary Fig. 1).

Differentiation between transient and equilibrium periods is somewhat arbitrary. Full equilibration requires millennia[28], but our classification is an approximation of the periods in a time series in which trends may shift from an

immediate to a longer-term response. See ref. [71] for further discussion of climate equilibrium and its relation to transient timescales.

We process Niño3.4 data for each simulation by subtracting the long-term monthly mean from each monthly value and removing the long-term trend using a quadratic fit[9]. We remove the trend from the transient and equilibrium periods separately to account for trend shifts after the initial adjustment to forcing.

Pacific mean climate. We quantify changes in the Pacific mean climate by taking 100-year running means of the zonal temperature gradient and SLP gradient from each simulation and calculating percentage changes between each value from the high-CO_2 simulation and the mean value from the control. To calculate average zonal temperature gradients during El Niño events (dashed lines in Fig. 4a), we define a typical El Niño event as the 75th–90th percentile Niño3.4 temperature anomalies as in our frequency calculation, and we calculate the mean zonal temperature gradient only during those months whose temperature anomalies fall within that boundary. Our results are not sensitive to these definitions—for example, we find similar results if we expand the definition of El Niño events to encompass the 75th–100th percentiles. We define La Niña events similarly, as the months in which Niño3.4 temperature anomalies fall within the tenth–25th percentiles, with results again not sensitive to this precise definition.

Global mean temperature anomalies. To use global mean temperature anomalies as a baseline in Fig. 3, we define global mean temperature as the weighted average of TAS, with weights determined by the square root of the cosine of latitude. Global mean surface temperature anomalies for each high-CO_2 simulation are defined as changes relative to the average global mean surface temperature from the entire control simulation. One-hundred-year rolling means are used in Fig. 3 for an apples-to-apples comparison with 100-year rolling standard deviations (that is, ENSO amplitude).

Comparison of observed and modelled zonal temperature gradient trends. LongRunMIP control simulations provide useful information regarding millennial-scale unforced variability in the tropical Pacific. To contribute to the ongoing debate over whether observed trends in the Pacific zonal temperature gradient are forced responses[26] or internal variability[27], we compare these observed trends to distributions of trends from the ten control simulations we analyse here (Supplementary Fig. 12). Observed SST data are drawn from the datasets ERSSTv.5 (ref. [72]) and HadISST [73].

We calculate rolling 25- and 35-year least-squares linear trends in the zonal temperature gradient, as defined in Spatial indices, from each control simulation over the entire simulation and from the observations over the years 1958–2017 (ref. [26]). Supplementary Fig. 12 plots these distributions of trends.

Thermodynamic feedback changes. To understand the thermodynamics of El Niño events under high-CO_2 conditions, we consider the negative heat flux feedback that acts in the direction opposite to the positive Bjerknes feedback[40,41,43,44]. An increase in this feedback would tend to dampen ENSO amplitude. This negative feedback can be approximated as linear with the following formulation: $F' = \alpha T'$, with feedback parameter α, atmosphere-to-ocean heat flux anomalies F' and temperature anomalies T'. α is therefore a negative feedback when positive SST anomalies reduce net heat flux into the ocean, dampening El Niño events.

Net surface fluxes are defined as the sum of long-wave, short-wave, latent heat and sensible heat fluxes. We therefore measure the net size of the thermodynamic feedback, α_{NET}, by regressing net surface flux anomalies in the Niño3.4 region onto temperature anomalies from the same region[40,41]. Previous literature[40,41,43,44] has highlighted the dominant contributions of the short-wave feedback, α_{SW} and the latent heat flux feedback, α_{LH} to net flux feedback. As such, we focus on these two specific feedbacks. We calculate the short-wave feedback by regressing short-wave downwelling radiation anomalies (the short-wave upwelling feedback is negligible) over the Niño3.4 region onto Niño3.4 temperature anomalies, and we calculate the latent heat flux feedback by regressing latent heat flux anomalies onto temperature anomalies in the same fashion. We use 100-year rolling regressions, estimated with ordinary least squares, to estimate distributions of these feedbacks in each simulation analogous to the distributions of ENSO amplitude.

These calculations use anomalies where monthly values from each simulation are referenced to their own long-term monthly means, so that mean shifts (and therefore climatological feedbacks) are removed along with the seasonal cycle. All quantities are detrended in the same manner as the ENSO indices (Distinguishing short- and long-term responses).

Data availability

All processed data required to reproduce the results of this study are available at https://github.com/ccallahan45/Callahan-et-al_NCC_2021/, archived on Zenodo at https://doi.org/10.5281/zenodo.4718010 (ref. [74]). Raw LongRunMIP data are not provided due to large file sizes, but these data are publicly available at https://data.iac.ethz.ch/longrunmip/, with further information available at http://www.longrunmip.org.

Code availability

Analysis code required to reproduce the results of this study is available at https://github.com/ccallahan45/Callahan-et-al_NCC_2021/, archived on Zenodo at https://doi.org/10.5281/zenodo.4718010 (ref. [74]).

References

49. Yeager, S. G., Shields, C. A., Large, W. G. & Hack, J. J. The low-resolution CCSM3. *J. Clim.* **19**, 2545–2566 (2006).
50. Danabasoglu, G. & Gent, P. R. Equilibrium climate sensitivity: is it accurate to use a slab ocean model? *J. Clim.* **22**, 2494–2499 (2009).
51. Gent, P. R. et al. The Community Climate System Model version 4. *J. Clim.* **24**, 4973–4991 (2011).
52. Danabasoglu, G. et al. The CCSM4 ocean component. *J. Clim.* **25**, 1361–1389 (2012).
53. Rugenstein, M. A. A., Sedlacek, J. & Knutti, R. Nonlinearities in patterns of long-term ocean warming. *Geophys. Res. Lett.* **43**, 3380–3388 (2016).
54. Voldoire, A. et al. Evaluation of CMIP6 DECK experiments with CNRM-CM6-1. *J. Adv. Model. Earth Syst.* **11**, 2177–2213 (2019).
55. Donner, L. J., Wyman, B. L. & Hemler, R. S. et al. The dynamical core, physical parameterizations, and basic simulation characteristics of the atmospheric component AM3 of the GFDL Global Coupled Model CM3. *J. Clim.* **24**, 3484–3519 (2011).
56. Paynter, D., Frolicher, T. L., Horowitz, L. W. & Silvers, L. G. Equilibrium climate sensitivity obtained from multimillennial runs of two GFDL climate models. *J. Geophys. Res. Atmos.* **123**, 1921–1941 (2018).
57. Dunne, J. P. et al. GFDL's ESM2 global coupled climate-carbon earth system models. Part I: physical formulation and baseline simulation characteristics. *J. Clim.* **25**, 6646–6665 (2012).
58. Miller, R. L. et al. CMIP5 historical simulations (1850–2012) with GISS ModelE2. *J. Adv. Model. Earth Syst.* **6**, 441–478 (2014).
59. Nazarenko, L. et al. Future climate change under RCP emission scenarios with GISS ModelE2. *J. Adv. Model. Earth Syst.* **7**, 244–267 (2015).
60. Rind, D. et al. Multicentury instability of the Atlantic meridional circulation in rapid warming simulations with GISS ModelE2. *J. Geophys. Res. Atmos.* **123**, 6331–6355 (2018).
61. Schmidt, G. A. et al. Configuration and assessment of the GISS ModelE2 contributions to the CMIP5 archive. *J. Adv. Model. Earth Syst.* **6**, 141–184 (2014).
62. Cox, P. M., Betts, R. A., Jones, C. D., Spall, S. A. & Totterdell, I. J. Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature* **408**, 184–187 (2000).
63. Cao, L., Duan, L., Bala, G. & Caldeira, K. Simulated long-term climate response to idealized solar geoengineering. *Geophys. Res. Lett.* **43**, 2209–2217 (2016).
64. Dufresne, J.-L., Foujols, M.-A. & Denvil, S. et al. Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5. *Clim. Dyn.* **40**, 2123–2165 (2013).
65. Hasumi, H. & Emori, S. K-1 coupled GCM (MIROC) description. https://ccsr.aori.u-tokyo.ac.jp/~hasumi/miroc_description.pdf (2004).
66. Yamamoto, A. et al. Global deep-ocean oxygenation by enhanced ventilation in the Southern Ocean under long-term global warming. *Glob. Biogeochem. Cycles* **29**, 1801–1815 (2015).
67. Yoshimori, M. et al. A review of progress towards understanding the transient global mean surface temperature response to radiative perturbation. *Prog. Earth Planet. Sci.* **3**, 21 (2016).
68. Mauritsen, T. et al. Developments in the MPI-M Earth System Model version 1.2 MPI-ESM1.2 and its response to increasing CO_2. *J. Adv. Model. Earth Syst.* **11**, 998–1038 (2019).
69. Rohrschneider, T., Stevens, B. & Mauritsen, T. On simple representations of the climate response to external radiative forcing. *Clim. Dyn.* **53**, 3131–3145 (2019).
70. Vecchi, G. A. et al. Weakening of tropical Pacific atmospheric circulation due to anthropogenic forcing. *Nature* **441**, 73–76 (2006).
71. Rugenstein, M. et al. Equilibrium climate sensitivity estimated by equilibrating climate models. *Geophys. Res. Lett.* **47**, e2019GL083898 (2020).
72. Huang, P. W. et al. *NOAA Extended Reconstruction Sea Surface Temperature (ERSST), Version 5*. NOAA National Centers for Environmental Information https://doi.org/10.7289/V5T72FNM (2017).
73. Rayner, N. A. et al. Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.* **108**, 4407 (2003).
74. Callahan, C. W. Robust decrease in ENSO amplitude under long-term warming. *Zenodo* https://zenodo.org/record/4718010 (2021).

Acknowledgements

We thank M. Jansen, N. Maher, J. Franke and K. Schwarzwald for helpful discussions and insights. This research was performed as part of the Center for Robust Decision-making on Climate and Energy Policy at the University of Chicago, funded by NSF through the

Decision Making Under Uncertainty programme (grant no. SES-1463644 to E.J.M.). Computing resources were provided by the University of Chicago Research Computing Center and Dartmouth College Research Computing. M.R. is funded by the Alexander von Humboldt foundation. This project received funding from the European Research Council under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 786427, project Couplet) to J.B.-J. Finally, this work would not have been possible without the efforts of the contributors to the LongRunMIP project.

Author contributions

C.W.C., C.C. and E.J.M. conceptualized the study. C.W.C. and C.C. performed the analyses. J.B.-J., M.R., S.Y. and E.J.M. provided data, analysis tools and feedback on results. C.W.C., C.C. and E.J.M. wrote the manuscript, with all authors providing feedback.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41558-021-01099-2.

Correspondence and requests for materials should be addressed to C.C. or E.J.M.

Peer review information *Nature Climate Change* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.