AGU PUBLICATIONS

Geophysical Research Letters

RESEARCH LETTER
10.1002/2017GL074622

Key Points:
- Twentieth century tropical Pacific SST trends in models and observations are in opposition with implications for confidence in future projections
- Many SST trends are insignificant relative to internal variability suggesting relatively late emergence of anthropogenic changes
- Constraining observational uncertainties and process-level validation of dynamics relevant to climate change in tropical Pacific are needed

Supporting Information:
- Supporting Information S1

Correspondence to:
S. Coats,
coats@ucar.edu

Citation:
Coats, S., & Karnauskas, K. B. (2017). Are simulated and observed twentieth century tropical Pacific sea surface temperature trends significant relative to internal variability? *Geophysical Research Letters*, *44*, 9928–9937. https://doi.org/10.1002/2017GL074622

Received 27 JUN 2017
Accepted 15 SEP 2017
Accepted article online 18 SEP 2017
Published online 13 OCT 2017

Are Simulated and Observed Twentieth Century Tropical Pacific Sea Surface Temperature Trends Significant Relative to Internal Variability?

S. Coats[1,2] **and K. B. Karnauskas**[2,3]

[1]Advanced Studies Program, National Center for Atmospheric Research, Boulder, CO, USA, [2]Cooperature Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA, [3]Department of Atmospheric and Oceanic Sciences, University of Colorado, Boulder, CO, USA

Abstract Historical trends in the tropical Pacific zonal sea surface temperature gradient (SST gradient) are analyzed herein using 41 climate models (83 simulations) and 5 observational data sets. A linear inverse model is trained on each simulation and observational data set to assess if trends in the SST gradient are significant relative to the stationary statistics of internal variability, as would suggest an important role for external forcings such as anthropogenic greenhouse gasses. None of the 83 simulations have a positive trend in the SST gradient, a strengthening of the climatological SST gradient with more warming in the western than eastern tropical Pacific, as large as the mean trend across the five observational data sets. If the observed trends are anthropogenically forced, this discrepancy suggests that state-of-the-art climate models are not capturing the observed response of the tropical Pacific to anthropogenic forcing, with serious implications for confidence in future climate projections. There are caveats to this interpretation, however, as some climate models have a significant strengthening of the SST gradient between 1900 and 2013 Common Era, though smaller in magnitude than the observational data sets, and the strengthening in three out of five observational data sets is insignificant. When combined with observational uncertainties and the possibility of centennial time scale internal variability not sampled by the linear inverse model, this suggests that confident validation of anthropogenic SST gradient trends in climate models will require further emergence of anthropogenic trends. Regardless, the differences in SST gradient trends between climate models and observational data sets are concerning and motivate the need for process-level validation of the atmosphere-ocean dynamics relevant to climate change in the tropical Pacific.

1. Introduction

The tropical Pacific zonal sea surface temperature (SST) gradient (hereinafter SST gradient) plays an outsized role in the global climate system. The mean state of the tropical Pacific, for instance, modulates the El Niño–Southern Oscillation (ENSO), the dominant mode of internannual to decadal time scale global climate variability (e.g., Fedorov & Philander 2000; Sarachik & Cane 2010). Small changes in the SST gradient, in particular, drive precipitation changes in much of the extratropics in observations (e.g., Schubert et al., 2004; Seager, 2015) and in future climate projections by coupled general circulation models (CGCMs—e.g., Seager & Vecchi, 2010). Such changes also have an impact on global temperature variability and have been implicated as causing the recent so-called hiatus in global warming (Kosaka & Xie, 2013). The importance of the SST gradient to global climate necessitates an understanding of how it will respond to anthropogenic forcing, specifically whether the climatological SST gradient, with warmer SSTs in the west than east, is expected to strengthen (a positive trend) or weaken (a negative trend).

Theory offers a number of possible outcomes. First, an atmosphere-centric theory has been proposed, wherein the zonally asymmetric tropical atmospheric (Walker) circulation should weaken in response to anthropogenic forcing as a consequence of the mass and energy balances of the hydrologic cycle (Held & Soden, 2006; Vecchi & Soden 2007). Because the atmosphere and ocean are tightly coupled over the tropical Pacific, particularly on interannual time scales, the weakening Walker circulation should lead to a weakening SST gradient by reducing the strength of upwelling in the eastern equatorial Pacific. Second, an ocean-centric theory proposes that the SST gradient should strengthen in response to anthropogenic forcing, because the vigorous mean upwelling and poleward divergence of heat flux in the eastern equatorial Pacific should dilute surface warming, resulting in a stronger SST contrast with the western tropical Pacific that is then amplified

by a Bjerknes-type feedback (Clement et al., 1996; Seager & Murtugudde, 1997; Sun & Liu, 1996). Finally, when the thermodynamic coupling of the atmosphere and ocean is considered, it has been proposed that the SST gradient will weaken because of more effective surface evaporative cooling in the western than eastern tropical Pacific (Xie et al., 2010).

Motivated by these theories, efforts have been made to relate observed changes in the tropical Pacific to anthropogenic forcing, leading to a number of opposing conclusions (Cane et al., 1997; Compo & Sardeshmukh, 2010; Compo et al., 2013; Deser et al., 2010; Karnauskas et al., 2009; Kumar et al., 2010; L'Heureux, Collins, et al., 2013; Solomon & Newman, 2012; Tokinaga, Xie, Deser, et al., 2012; Tokinaga, Xie, Timmermann, et al., 2012; Vecchi et al., 2006; Zhang et al., 2010). For instance, observations of the atmosphere suggest that the Walker circulation is weakening as a consequence of the mass and energy balances of the hydrologic cycle and that this response, at least in CGCMs, leads to a weakening SST gradient (Vecchi et al., 2006; Vecchi & Soden, 2007). SST reconstructions based on observations, however, suggest that the SST gradient has actually strengthened over the twentieth century (Cane et al., 1997; Karnauskas et al., 2009). While this discrepancy could indicate that CGCMs have a different forced response than the real world, atmospheric GCMs (AGCMs) forced with observations of SSTs suggest that in order to reconcile observations of the Walker circulation and SST gradient, SST reconstructions must be biased and, further, that observed atmospheric changes are actually a response to a weakening SST gradient (Tokinaga, Xie, Deser, et al., 2012; Tokinaga, Xie, Timmermann, et al., 2012). To further complicate matters, recent observations of the Walker circulation suggest that it is actually strengthening, albeit over a shorter time period (Kociuba & Power, 2015; L'Heureux, Lee, et al., 2013; Sandeep et al., 2014; Sohn et al., 2016), although the opposite has also been suggested (Bellomo & Clement, 2015).

Of particular interest is whether CGCMs will correctly simulate changes to the SST gradient over the coming century. To begin validating CGCMs, SST gradient trends in observational data sets and CGCMs can be assessed for consistency, and herein, we quantitatively analyze these trends in five SST reconstructions and one set of uninterpolated observations of SSTs (hereinafter observational data sets) as compared to an ensemble of 83 simulations from 41 state-of-the-art CGCMs completed as part of the Coupled Model Intercomparison Project phase 5 (CMIP5—Table S1—Taylor et al., 2012). This work represents an advance in two regards: First, SST gradient trends in all available observational data sets and CGCM simulations are compared within a single consistent framework, and second, we use a linear inverse model (LIM) to assess the significance of these trends relative to the stationary statistics of internal variability. The latter development is important because the magnitude of interannual and decadal internal variability in the tropical Pacific (along with observational uncertainties—e.g., Deser et al., 2010; Tokinaga, Xie, Deser, et al., 2012; Tokinaga, Xie, Timmermann, et al., 2012) has largely precluded a confident attribution of long-term changes in the SST gradient to anthropogenic forcing—see, for instance, the many opposing conclusions on observed anthropogenic changes referenced above. For climate models "large" ensembles of simulations can be used to separate anthropogenic and internal changes (e.g., Kay et al., 2015; Rodgers et al., 2015), but these ensembles are computationally expensive and we still only have a single trajectory of the observed climate system against which to compare the CGCMs. LIMs are a computationally efficient approach that can be applied to both observational data sets and CGCMs to produce ensembles of trajectories that have a demonstrated ability to reproduce the internal characteristics of the climate system (e.g., Newman et al., 2016, 2011; Penland & Matrasova, 1994). Solomon and Newman (2012), for instance, use a LIM-based approach to assess the consistency of SST trends in observational data sets, finding that any discrepancies between data sets can be explained by their representation of the ENSO. Herein, we apply the same LIM to both observational data sets and CGCMs to answer two questions: Are trends in the SST gradient significant relative to the stationary statistics of internal variability, as would suggest that they may be anthropogenic in origin? If so, are SST gradient trends in the CGCMs and observational data sets consistent, as would suggest that CGCMs realistically simulate this response to anthropogenic forcing?

2. Methods

All data sets based on instrumental observations of SST are listed in Table S2 in the supporting information. Five of the data sets (Hirahara et al., 2014; Huang et al., 2015; Kaplan et al., 1998; Liu et al., 2015; Rayner et al., 2003; Reynolds et al., 2007; Smith et al., 2008) are reconstructions of SST that are based on observations of

SSTs that have been bias corrected and then interpolated onto an even grid, while one (Kennedy et al., 2012a, 2012b) is observations of SST that are bias corrected and binned. All CGCM data are from the CMIP5 (Taylor et al., 2012) and the simulations employed herein are listed in Table S1. All data used for calculating trends were smoothed using a 3 month running average filter to prevent aliasing the highest frequencies in the long-term trends and for use in the linear inverse model (LIM)-based filtering and significance testing (outlined below). The historical analysis period is 1900–2013 Common Era (C.E.), unless otherwise noted. Because the CMIP5 historical simulations end in 2005 C.E., these are appended to the 2006–2013 C.E. period in the representative concentration pathway (RCP) 8.5 simulations to analyze historical SST trends. Analyses of future SST trends employ the RCP 8.5 simulations for the period 2006–2100 C.E. A LIM is a simple and computationally inexpensive method to emulate the statistics of dynamical systems for which the decorellation time scale of nonlinear processes is short relative to linear processes (Penland & Matrasova, 1994; Penland & Sardeshmukh, 1995; Solomon & Newman, 2012). This is demonstrably valid for internal atmosphere-ocean variability in the tropical Pacific on monthly-to-decadal time scales (Newman et al., 2016, 2011; Penland & Matrasova, 1994); thus, the evolution of tropical Pacific SST anomalies can be approximated by a linear dynamical system forced by white noise:

$$\frac{d\mathbf{X}}{dt} = \mathbf{L}\mathbf{X} + \zeta$$

where **X** is the SST anomaly state vector between 1900 and 2013 C.E. with a monthly time step (but smoothed with a 3 month running average), **L** is the deterministic feedback matrix, and ζ is the white noise forcing term. With analogy to univariate red noise, this can be described as multivariate red noise when **X** is an m dimensional vector (in this case representing the m spatial locations of SST anomaly data over the tropical Pacific—40°E to 290°E; 30°S to 30°N), **L** is an m by m matrix and ζ is also m dimensional and allows for spatial structure across the m spatial locations. As noted above, SST anomalies are smoothed with a 3 month running average before constructing the SST anomaly state vector. This is done to emphasize the linear sources of SST variability in the tropical Pacific, effectively averaging multiple nonlinear processes (which have inherently short time scales) into each time step. **L** can be estimated from the SST anomaly state vector:

$$\mathbf{L} = n^{-1} \ln\left(\mathbf{C_n}\mathbf{C_0^{-1}}\right)$$

where $\mathbf{C_n}$ is the covariance matrix of the SST anomalies at lag n and $\mathbf{C_0}$ is the covariance matrix with no lag. Following Penland and Matrasova (1994), n is set at 3 months; however, **L** is not sensitive to this choice if the dynamical system being approximated is effectively linear (e.g., Penland & Matrasova, 1994; Penland & Sardeshmukh, 1995). The covariance matrix of the noise term ($\mathbf{Q} = \rangle\zeta\zeta^T\langle dt$) is estimated from the fluctuation dissipation relationship (with $d\mathbf{C_0}/dt = 0$, because $\mathbf{C_0}$ is assumed to be unchanging with time):

$$0 = \mathbf{L}\mathbf{C_0} + \mathbf{C_0}\mathbf{L^T} + \mathbf{Q}$$

Together these equations can be used to produce random realizations of tropical Pacific SST anomalies ($\boldsymbol{X}(t)$) in a two-step process following Penland and Sardeshmukh (1995):

$$\mathbf{Y}(t + \Delta t) = (\mathbf{I} + \Delta t)\mathbf{Y}(t) + \sqrt{\Delta t}\mathbf{V}\mathbf{D}\mathbf{r};$$
$$\mathbf{X}(t + \Delta t) = (\mathbf{Y}(t + \Delta t) + \mathbf{Y}(t))/2$$

where **I** is the identity matrix, **r** is zero-mean unit variance Gaussian random numbers generated at each time step, Δt is the time step, and **V** and **D** are the eigenvectors and the square root of the eigenvalues of **Q**. The 114 year (the length of the historical analysis period—1900–2013 C.E.) or 95 year (the length of the future analysis period—2006–2100 C.E.) realizations of the stochastically forced LIM are produced using a 6 h time step and a 10 year spin-up time. For computational efficiency all calculations use the first 20 empirical orthogonal functions (EOFs) of tropical Pacific SST anomalies over the historical analysis period (this represents a range of 71% to 91% of the variance across the 83 model simulations and 5 observational data sets—average of 83%); $\mathbf{X}(t)$ is thus 20 principal components that must be projected onto the 20 EOFs to produce full tropical Pacific SST anomaly fields. To assess significance of SST trends relative to internal variability, 100 random realizations of tropical Pacific SST anomalies are produced and linear trends are calculated from each. A trend is significant if it is larger than 95 or smaller than all but 5 of the realizations. Hereinafter, all use of the terms significant or insignificant is relative to the stationary statistics of internal variability derived from the LIM realizations.

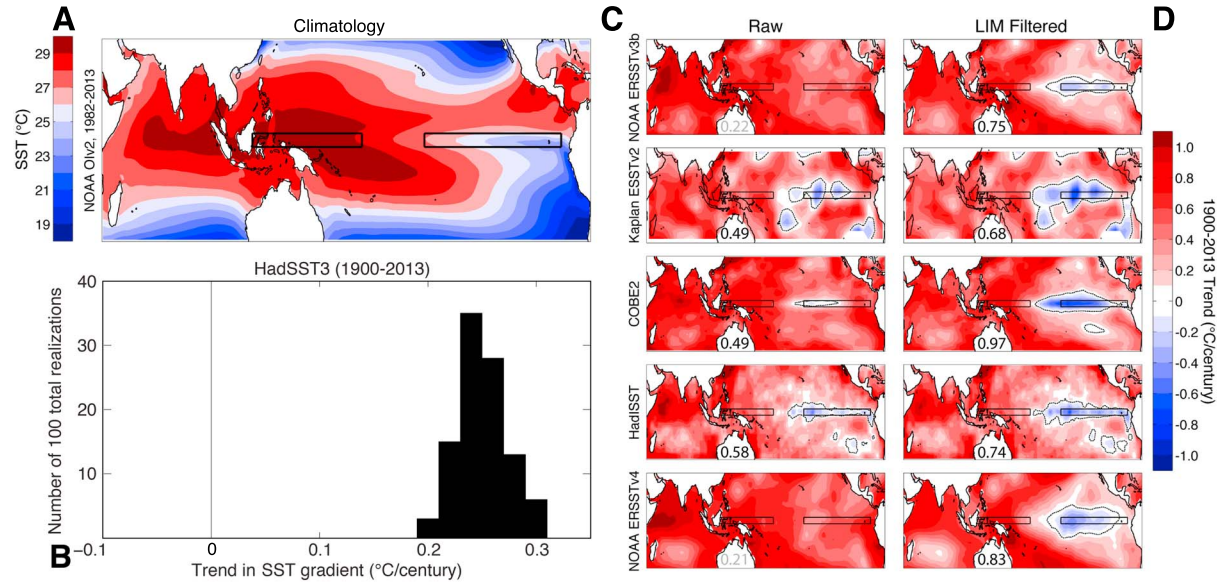

Figure 1. (a) SST climatology (°C) between 1982 and 2013 C.E., blue colors are cooler than the regional mean (26°C) and red colors are warmer. (b) Trend (°C/century) in the SST gradient (SST difference between 2.5°N-S, 117°E–173°E and 2.5°N-S, 205°E–275°E) between 1900 and 2013 C.E. in all 100 realizations of HadSST3 (Kennedy et al., 2012a, 2012b). (c) SST trend (°C/century) between 1900 and 2013 C.E. The SST reconstructions used in each panel are listed on the *y* axis and the references can be found in Table S1. (d) SST trend (°C/century) between 1900 and 2013 C.E. after filtering ENSO using an optimal perturbation filter (Solomon & Newman, 2012). The boxes used to define the SST gradient are plotted and the value of the trend in the SST gradient is labeled in the map inset of Figures 1c and 1d (black trend values are significant at the 95% level). All trends are estimated from a linear least squares fit to the 3 month running average SST time series.

The filter used for removing ENSO from the observational SST data sets is the LIM-based optimal perturbation filter of Solomon and Newman (2012). Following Solomon and Newman (2012), the maximum anomaly growth $\widetilde{\mathbf{X}}(t+t_n)$ over a time period (t_n) is initiated by the optimal initial condition $\widetilde{\mathbf{X}}(t)=\Phi_1$ where is Φ_1 is the first right singular vector of $\mathbf{L}(t_n)$. It has been shown that in the real world (i.e., observational SST data sets) this maximum anomaly growth is realized by ENSO events for t_n of 3 to 9 months. ENSO events can thus be filtered from these observational SST data sets by removing the evolution (over some time period t_1) of the projection of Φ_1 on tropical Pacific SST anomalies at each time step. In Solomon and Newman (2012) t_1 is set at 21 months and t_n at 3 months. The same parameter choices and filter setup are used herein with the exception of the time period over which it was implemented (1900–2013 C.E. as opposed to 1900–2010 C.E.). Using the optimal perturbation filter on the CGCMs would allow for an ideal comparison of anthropogenic trends in observational data sets and CGCMs. As noted in Solomon and Newman (2012), however, the optimal perturbation filter will erroneously remove long-term trends if the SST pattern of those trends projects on the SST pattern associated with the ENSO. While Solomon and Newman (2012) demonstrate that this is not the case for the observational data sets, it is the case for many CGCMs. Adding additional variables to the LIM, as suggested therein, was not sufficient to prevent the aliasing of short-term variability and long-term trends by the optimal perturbation filter (not shown). Therefore, we instead test the significance of trends relative to the stationary statistics of internal variability using the stochastically forced LIM, as noted above.

LIMs of the tropical Pacific often include other variables, such as wind stress, thermocline depth, and sea surface height. In order to maximize the number of CGCMs (many of which do not make these variables available) and the length of the analysis period (instrumental records of these variables tend to be shorter), the LIM used herein includes only SSTs. Nevertheless, for CGCMs that make these other variables available sensitivity testing suggests that including them in the LIM-based significance testing does not appreciably change results (not shown). Likewise, LIMs can be constructed after detrending the input fields (in this case SSTs) to remove the influence of anthropogenic forcing. To maintain consistency with Solomon and Newman (2012), and further encouraged by the fact that LIMs are not highly sensitive to long-term trends in their input fields (e.g., Newman et al., 2013), the LIMs used herein do not employ detrended SSTs. Nevertheless, the same analyses were also completed with LIMs that employ detrended SSTs and the results are largely consistent.

The SST gradient is defined as the difference in average SSTs between the western and eastern tropical Pacific (SST difference between 2.5°N-S, 117°E–173°E and 2.5°N-S, 205°E–275°E). This definition is based on the climatology in Figure 1a, and it is positive in the climatological mean, with warmer SSTs in the west than east. A strengthening SST gradient or positive trend in the SST gradient thus indicates an enhancement of the climatological SST gradient. The metric of Walker circulation strength is the sea level pressure (SLP) difference between those same two boxes but in the opposite sense (east minus west—SLP difference between 2.5°N-S, 205°E–275°E and 2.5°N-S, 117°E–173°E). This SLP gradient is likewise positive in the climatological mean, with higher SLPs in the east than west. A strengthening Walker circulation or positive trend in the Walker circulation thus indicates an enhancement of the climatological SLP gradient.

3. Results

There are positive trends in the SST gradient as large as 0.58°C per century between 1900 and 2013 C.E. in some observational data sets (i.e., HadISST—Rayner et al., 2003), and a positive trend in all five (Figure 1c). However, not all observational data sets have a positive trend in the SST gradient that is significant relative to the stationary statistics of internal variability (Figure 1c) and there are inconsistent patterns of SST trends across the full tropical Pacific basin in each (e.g., Figure 1c, top and bottom). Nevertheless, when ENSO is filtered from these observational data sets following the methodology of Solomon and Newman (2012), all indicate a consistent pattern and magnitude of SST trends (Figure 1d). This suggests that any differences in the SST trends between the observational data sets can be explained by different estimates of ENSO variability (Solomon & Newman, 2012). Uninterpolated observations of SSTs also indicate a strengthening SST gradient between 1900 and 2013 C.E., with a positive trend in all 100 realizations of HadSST3 (Kennedy et al., 2012a, 2012b)—each with a different treatment of bias correction and associated uncertainties (Figure 1b). This suggests that the interpolation methodologies used to produce the SST reconstructions in Figure 1c are not likely to explain the positive SST gradient trends. Since part of the positive SST gradient trend could be related to a shift in Pacific Decadal Variability (PDV) toward colder conditions after the 1980s (e.g., Mantua et al., 1997—see also the characteristic extratropical horseshoe pattern of SSTs in Figure S1, right), we confirm that the positive trend is present without the modern shift toward cold PDV conditions (Figure S1, left). Together these results suggest that there has been an increase in the SST gradient between 1900 and 2013 C.E.

In response to projected future anthropogenic forcing (RCP8.5—2006–2100 C.E.), CMIP5 CGCMs overwhelmingly (74% of the 83 simulations, with 53% being significant) project a weakening SST gradient (negative trends in Figures 2a and 2c). There are 11 simulations (13% of the total—10 of the 11 simulations are from the CSIRO-Mk3-6-0) with significant positive trends in the SST gradient. It is interesting to note that these significant positive trends are largely manifest as an interhemispheric pattern of temperature change, with greater warming in the Northern than Southern Hemispheres (Figures 2b and 4—with the exception of F-Goals, which is addressed below)—a response to anthropogenic forcing that has been noted elsewhere (Xie et al., 2010; Zhang & Li, 2014) and demonstrated to be reproduced without ocean dynamics (Cai et al., 2015; Zhang & Li, 2014).

The CMIP5 CGCMs, on average (63% of the 83 simulations), also simulate a weakening SST gradient over the period in common with the observational data sets (1900–2013 C.E.—Figures 3a and 3c) and this weakening appears to be related to a weakening Walker circulation (Figure S2—e.g., Vecchi & Soden, 2007; Tokinaga, Xie, Timmermann, et al., 2012). The patterns comprising Figure 3a, however, are less consistent than for the future projections (not shown), likely suggesting a weaker response to what are weaker forcing conditions over the observational interval and, subsequently, a more important role for internal variability. This is also manifest in the significance of trends in the SST gradient relative to internal variability, with only 11% being significant (5% positive and 6% negative) as compared to 66% in the future projections (13% positive and 53% negative). Not one of the CMIP5 CGCMs, however, simulates a positive trend in the SST gradient over the period 1900–2013 C.E. as large as the mean trend of the unfiltered observational data sets in Figure 1c (Figure 3a), and only two simulations have a positive trend as large as the smallest trend among the observational data sets. This result is largely insensitive to the use of only a single ensemble member from each of the 41 models (Figure S3), the trend start date (Figure S4), or the location of the boxes used to define the SST gradient (Figure S5).

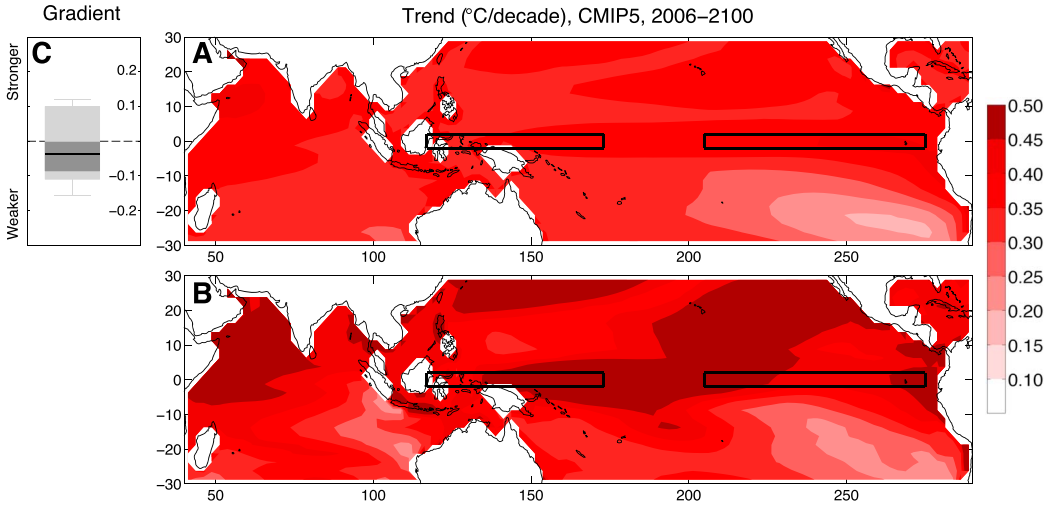

Figure 2. SST trends (°C/decade) between 2006 and 2100 C.E. for 83 CMIP5 simulations from 41 CGCMs (Table S2). (a) Ensemble mean of the SST trend (°C/decade) between 2006 and 2100 C.E. for the CMIP5 simulations. (b) Mean of the SST trend (°C/decade) between 2006 and 2100 C.E. for the CMIP5 simulations with positive trends in the SST gradient (SST difference between 2.5°N-S, 117°E–173°E and 2.5°N-S, 205°E–275°E) that are significant at the 95% level (11 simulations, Figure 4). (c) Range in the trend (°C/decade) in the SST gradient between 2006 and 2100 C.E. for the CMIP5 simulations. Black line is the median, dark grey the 25th to 75th percentiles, light grey the 5th to 95th percentiles and whiskers are the full range. All trends are estimated from a linear least squares fit to 3 month running average SST time series. The 83 CMIP5 simulations are using the RCP8.5 emissions scenario.

There are CGCMs that while weaker than the observational data sets, have a significant increase in the SST gradient between 1900 and 2013 C.E. (Figure S6). It is interesting to analyze the climate changes that correspond to these SST trends in the simulations for which ocean variables are available (no ocean variables are available for CSIRO-MK3-6-0). One consistent feature is a strengthening and shoaling of the equatorial undercurrent (EUC) on the order of 25% per century (consistent with Drenkard & Karnauskas, 2014; Karnauskas et al., 2016) and spatially concurrent moderation of warming of the equatorial eastern Pacific Ocean over the top few hundred meters including the surface (Figure 4). Interestingly, the mechanism of EUC strengthening in these CGCMs is not consistent. Within the IPSL-CM5A-LR simulation the strengthening EUC is weaker

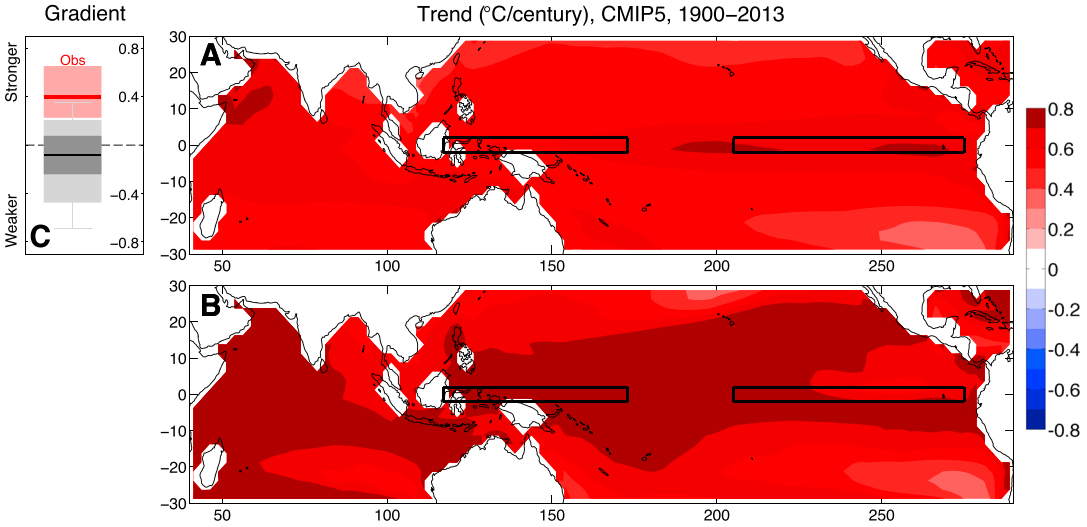

Figure 3. SST trends (°C/century) between 1900 and 2013 C.E. for 83 CMIP5 simulations from 41 models (Table S2). (a) Ensemble mean of the SST trend (°C/century) between 1900 and 2013 C.E. for the CMIP5 simulations. (b) Mean of the SST trend (°C/century) between 1900 and 2013 C.E. for the CMIP5 simulations with positive trends in the SST gradient (SST difference between 2.5°N-S, 117°E–173°E and 2.5°N-S, 205°E–275°E) that are significant at the 95% level (four simulations, Figure 4). (c) Range in the trend (°C/century) in the SST gradient between 1900 and 2013 C.E. for the CMIP5 simulations. Black line is the median, dark grey the 25th to 75th percentiles, light grey the 5th to 95th percentiles and whiskers are the full range. The range in trends from the raw observational data sets in Figure 1c is plotted as the lightly shaded region with the red line indicating the mean. All trends are estimated from a linear least squares fit to the 3 month running average SST time series. The 83 CMIP5 simulations are using the historical simulations appended to simulations using the RCP8.5 emissions scenario.

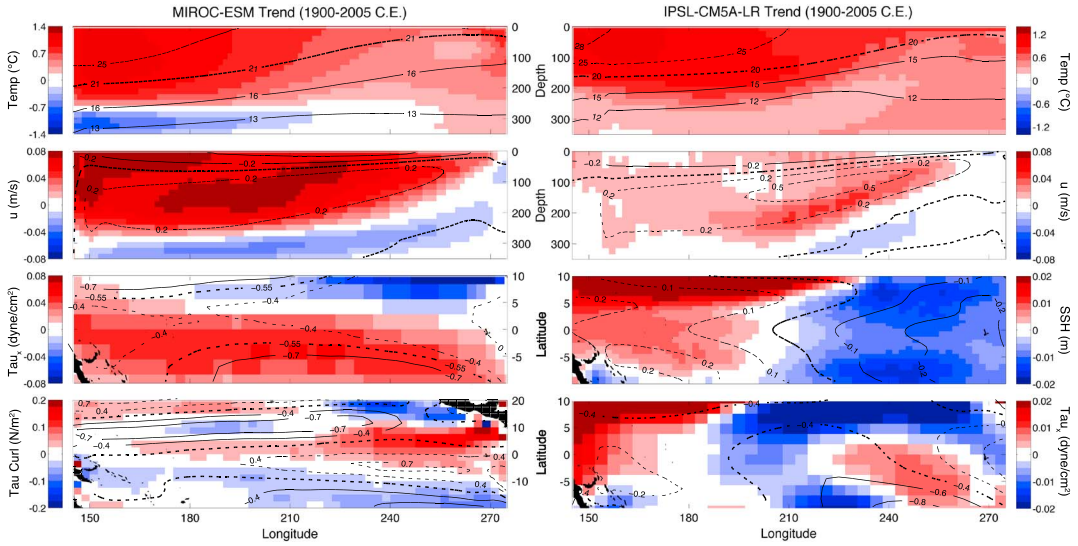

Figure 4. Trends (colors) and climatology (contours) between 1900 and 2005 C.E. for the first ensemble member of the IPSL-CM5A-LR and MIROC-ESM models, both of which have a significant positive trend in the SST gradient (SST difference between 2.5°N-S, 117°E–173°E and 2.5°N-S, 205°E–275°E) over the historical interval (1900–2013 C.E.—Figures 3 and S6). (a and b) The three-dimensional fields are averaged between 2°S and 2°N. For the climatology, the thick dashed line is either the regional mean (Temp and Tau_x) or the zero contour (u, SSH, and Tau Curl). Thin dashed lines are values greater than the regional mean or zero contour, and solid lines are values less than the regional mean or zero contour. All trends are estimated from a linear least squares fit to the monthly time series at each depth by longitude or latitude by longitude grid point.

and is part of a canonically La Niña-like strengthening of zonal wind in the central equatorial Pacific, associated increase in wind stress divergence driven upwelling and strengthening of the sea surface height gradient and pressure gradient force that drives the EUC (Figure 4). This mechanism is largely consistent with the ocean dynamical thermostat mechanism as originally posed (Clement et al., 1996), although the EUC in this model also strengthens via a strengthening Northern Hemisphere western boundary current (not shown but with a mechanism similar to the Southern Hemisphere wind stress curl mechanism in CMIP3 CGCMs—Sen Gupta et al., 2012). In the MIROC-ESM simulation, by contrast, the zonal winds weaken over the tropical Pacific (Figure 4). While this should weaken both equatorial upwelling and the EUC, the off equatorial wind stress curl trends project strongly on the climatological wind stress curl pattern and thus provide a compensating increase in pycnocline convergence at both the interior and western boundaries (i.e., strengthening subtropical cells—Figure 4). SST trends in the F-Goals model, which has a significant increase in the SST gradient in the future but not over the historical interval, are driven by a mechanism similar to that in the IPSL-CM5A-LR simulation with an increase in wind stress divergence driving upwelling in the central equatorial Pacific Ocean (not shown).

4. Discussions and Conclusions

Not one of 83 analyzed CMIP5 simulations have a positive trend in the SST gradient between 1900 and 2013 C.E. as large as the mean trend across five observational data sets (Figure 2). While this discrepancy suggests that state-of-the-art CGCMs are not capturing the observed response of the tropical Pacific to anthropogenic forcing, there are caveats to this interpretation. For instance, there are multiple simulations (four total from three CGCMs) with significant (relative to the stationary statistics of internal variability) positive trends in the SST gradient between 1900 and 2013 C.E., although the underlying physical mechanisms are not consistent across these models (Figure 4). Furthermore, the majority of SST gradient trends in CGCMs (11%, relative to 66% in the future projections) and SST gradient trends in three of the five observational data sets are not significant relative to the stationary statistics of internal variability. This is perhaps unsurprising given the large magnitude internal variability on interannual-to-decadal (and potentially longer time scales—e.g., Karnauskas, Smerdon, et al., 2012) in the tropical Pacific. By consequence, the "signal" of anthropogenic forcing is likely to emerge relatively late from the "noise" of internal variability. Techniques to remove internal variability from observational data sets have proved successful (Compo & Sardeshmukh, 2010; Solomon &

Newman, 2012) but have remained elusive in CGCMs. A more confident validation of anthropogenic SST gradient trends in CGCMs is thus likely to require further time for emergence of anthropogenic changes.

Despite these caveats, there is still sufficient evidence to suggest that the SST trends in the CMIP5 CGCMs are inconsistent with the increase in the SST gradient in the observational data sets. For instance, this discrepancy is unlikely to result from having too few simulations to fully sample internal variability, as 83 simulations from 41 CGCMs is a large and diverse ensemble, although CGCMs may still be underestimating centennial time scale internal variability that is also not sampled by the LIM. The inconsistency between the observational data sets and CGCM simulations is instead likely to result from model biases or observational uncertainties. If model biases are the issue, one possibility is that the CGCMs simulate too strong of a response of the atmosphere to anthropogenic forcing (i.e., weakening Walker circulation—although biases in simulated Walker circulation variability can impact trends, e.g., Sohn et al., 2016) and/or that such changes unrealistically dominate any attempt by ocean dynamics to mitigate surface warming (e.g., Clement et al., 1996). A further possibility is that the models are lacking or underrepresenting such ocean dynamics, for instance the vigorous mean upwelling and poleward divergence of heat in the eastern equatorial Pacific that underlies the ocean dynamical thermostat mechanism as originally posed (Clement et al., 1996). These ocean dynamics, however, are simulated reasonably well by CGCMs—or at least within observational error (e.g., Karnauskas, Johnson, et al., 2012—see also the seasonal cycle in SST trends in CMIP5 CGCMs noted by Xie et al., 2010). Other possible biases include, but are not limited to, representation of interannual and decadal variability and their rectification on the mean state (Cai et al., 2015; Johnson, 2013; Kohyama et al., 2017) and other ocean dynamics like the equatorial undercurrent (EUC—Drenkard & Karnauskas, 2015; Karnauskas, Johnson, et al., 2012). A full treatment of potential model biases and their impact on trends in the SST gradient is outside the scope of this paper but is a critical area of previous and ongoing research. If observational uncertainties are the issue, these could include changes in the spatial distribution and frequency of observations (Deser et al., 2010; Giese et al., 2010; Kennedy, 2014; Yasunaka & Hanawa, 2011), observation techniques (e.g., Thompson et al., 2008; Tokinaga, Xie, Timmermann, et al., 2012), and issues with the statistical methodologies used to reconstruct the SST field from sparse inputs (e.g., Deser et al., 2010). It is important to note, however, that uninterpolated observations of SSTs with a quantification of the uncertainties associated with bias correction indicate a positive trend in the SST gradient between 1900 and 2013 C.E. (Figure 1b), which at least partially addresses the latter two observational uncertainties.

Process-level validation of the atmosphere-ocean dynamics relevant to climate change in the tropical Pacific will help to better understand the inconsistency of SST gradient trends in observational data sets and CGCM simulations. Such endeavors are critical because if CGCMs do not capture the response of the tropical Pacific to anthropogenic forcing, then they will struggle to project the many climate impacts that are tightly coupled to this region. This would be reflected most obviously in Pacific-centric climate impacts including the magnitude and frequency of ENSO, tropical rainfall, Hadley cell width, and the position of the intertropical convergence zone. Remote regions would also be affected via atmospheric teleconnections. Particularly sensitive is the American West, where small increases in the SST gradient are known to have caused drying in observations (Seager, 2015) and where CMIP5 CGCMs already project future drying well in excess of any period over the last 1,000 years (Cook et al., 2015; Seager et al., 2013). Such projections are of utmost importance to society's ability to plan for and adapt to climate change, rendering any potential mismatch between the observed and simulated response of the SST gradient to anthropogenic forcing critically important to resolve before further confidence can be placed therein.

Acknowledgments

We acknowledge the World Climate Research Programme's Working Group on Coupled Modeling, which is responsible for CMIP, and we thank the climate modeling groups (listed in Table S1) for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portal. We, further, thank Haibo Liu and Naomi Henderson for their considerable computational support, Toby Ault for his helpful discussion on the LIM, and Peter Molnar, Gokan Danabasoglu, and Richard Seager for helpful discussions and comments on earlier versions of the manuscript. The CMIP5 model output and observational data sets are available at the following links: CMIP5 (https://esgf-node.llnl.gov/search/cmip5/), NOAA OIv2 (https://www.esrl.noaa.gov/psd/data/gridded/data.noaa.oisst.v2.html), HadISST (http://www.metoffice.gov.uk/hadobs/hadisst/), COBE2 (https://www.esrl.noaa.gov/psd/data/gridded/data.cobe2.html), Kaplan SSTv2 (https://iridl.ldeo.columbia.edu/SOURCES/.KAPLAN/.EXTENDED/.v2/), ERSSTv3b (https://www.ncdc.noaa.gov/data-access/marineocean-data/extended-reconstructed-sea-surface-temperature-ersst-v3b), ERSSTv4 (https://www.ncdc.noaa.gov/data-access/marineocean-data/extended-reconstructed-sea-surface-temperature-ersst-v4), and HadSST3 (http://hadobs.metoffice.com/hadsst3/).

References

Bellomo, K., & Clement, A.-C. (2015). Evidence for weakening of the Walker circulation from cloud observations. *Geophysical Research Letters*, *42*, 7758–7766. https://doi.org/10.1002/2015GL065463

Cai, W., Santoso, A., Wang, G., Yeh, S.-W., An, S., Cobb, K. M., … Wu, L. (2015). ENSO and greenhouse warming. *Nature Climate Change*, *5*, 849–859.

Cane, M.-A., Clement, A. C., Kaplan, L., Kushnir, Y., Pozdnyakov, D., Seager, R., … Murtugudde, R. (1997). Twentieth-century sea surface temperature trends. *Science*, *275*, 957–960.

Clement, A.-C., Seager, R., Cane, M.-A., & Zebiak, S.-E. (1996). An ocean dynamical thermostat. *Journal of Climate*, *9*, 2190–2196.

Compo, G.-P., & Sardeshmukh, P.-D. (2010). Removing ENSO-related variations from the climate record. *Journal of Climate*, *23*(8), 1957–1978.

Compo, G.-P., Sardeshmukh, P. D., Whitaker, J. S., Brohan, P., Jones, P. D., & McCol, C. (2013). Independent confirmation of global land warming without the use of station temperatures. *Geophysical Research Letters*, *40*, 3170–3174, https://doi.org/10.1002/grl.50425

Cook, B.-I., Ault, T.-R., & Smerdon, J.-E. (2015). Unprecedented 21st century drought risk in the American Southwest and Central Plains. *Science Advances*, *1*. https://doi.org/10.1126/sciadv.1400082

Deser, C., Phillips, A.-S., & Alexander, M.-A. (2010). Twentieth century tropical sea surface temperature trends revisited. *Geophysical Research Letters*, *37*, L10701. https://doi.org/10.1029/2010GL043321

Drenkard, E.-J., & Karnauskas, K.-B. (2014). Strengthening of the Pacific Equatorial Undercurrent in the SODA reanalysis: Mechanisms, ocean dynamics, and implications. *Journal of Climate*, *27*, 2405–2416.

Fedorov, A.-V., & Philander, S.-F. (2000). Is El Niño changing? *Science*, *288*, 1997–2002.

Giese, B.-S., Slowey, N.-C., Ray, S., Compo, G.-P., Sardeshmukh, P.-D., Carton, J.-A., & Whitaker, J.-S. (2010). The 1918/19 El Niño. *Bulletin of the American Meteorological Society*, *91*(2), 177–183.

Held, I.-M., & Soden, B.-J. (2006). Robust responses of the hydrological cycle to global warming. *Journal of Climate*, *19*, 5686–5699.

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

Huang, B., Banzon, V. F., Freeman, E., Lawrimore, J., Liu, W., Peterson, T. C., … Zhang, H.-M. (2015). Extended Reconstructed Sea Surface Temperature version 4 (ERSST. v4). Part I: Upgrades and intercomparisons. *Journal of Climate*, *28*, 911–930.

Johnson, N.-C. (2013). How many ENSO flavors can we distinguish? *Journal of Climate*, *26*(13), 4816–4827.

Kaplan, A., Cane, M. A., Kushnir, Y., Clement, A. C., Blumenthal, M. B., & Rajagopalan, B. (1998). Analyses of global sea surface temperature 1856–1991. *Journal of Geophysical Research*, *103*, 18,567–18,589. https://doi.org/10.1029/97JC01736

Karnauskas, K.-B., Cohen, A.-L., & Gove, J.-M. (2016). Mitigation of Coral Reef warming across the Central Pacific by the equatorial undercurrent: A past and future divide. *Nature Scientific Reports*, *6*. https://doi.org/10.1038/srep21213

Karnauskas, K.-B., Johnson, G.-C., & Murtugudde, R. (2012). An equatorial ocean bottleneck in global climate models. *Journal of Climate*, *25*, 343–349.

Karnauskas, K.-B., Seager, R., Kaplan, A., Kushnir, Y., & Cane, M.-A. (2009). Observed strengthening of the zonal sea surface temperature gradient across the equatorial Pacific Ocean. *Journal of Climate*, *22*, 4316–4321.

Karnauskas, K.-B., Smerdon, J.-E., Seager, R., & González-Rouco, J.-F. (2012). A Pacific centennial oscillation predicted by coupled GCMs. *Journal of Climate*, *25*, 5943–5961.

Kay, J. E., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., … Holland, M. (2015). The Community Earth System Model (CESM) large ensemble project: A community resource for studying climate change in the presence of internal climate variability. *Bulletin of the American Meteorological Society*, *96*(8), 1333–1349.

Kennedy, J.-J. (2014). A review of uncertainty in in situ measurements and data sets of sea surface temperature. *Reviews of Geophysics*, *52*, 1–32. https://doi.org/10.1002/2013RG000434

Kennedy, J.-J., Rayner, N., Smith, R.-O., Parker, D.-E., & Saunby, M. (2012a). Reassessing biases and other uncertainties in sea surface temperature observations measured in situ since 1850: 1. Measurement and sampling uncertainties. *Journal of Geophysical Research: Atmospheres*, *116*, D14103. https://doi.org/10.1029/2010JD015218

Kennedy, J.-J., Rayner, N., Smith, R.-O., Parker, D.-E., & Saunby, M. (2012b). Reassessing biases and other uncertainties in sea surface temperature observations measured in situ since 1850: 2. Biases and homogenization. *Journal of Geophysical Research: Atmospheres*, *116*, D14104. https://doi.org/10.1029/2010JD015220

Kociuba, G., & Power, S.-B. (2015). Inability of CMIP5 models to simulate recent strengthening of the Walker Circulation: Implications for projections. *Journal of Climate*, *28*, 20–35.

Kohyama, T., Hartmann, D.-L., & Battisti, D.-S. (2017). La Niña-like mean-state response to global warming and potential oceanic roles. *Journal of Climate*. https://doi.org/10.1175/JCLI-D-16-0441.1

Kosaka, Y., & Xie, S.-P. (2013). Recent global-warming hiatus tied to equatorial Pacific surface cooling. *Nature*, *501*, 403–407.

Kumar, A., Bhaskar, J., & L'Heureux, M.-L. (2010). Are tropical SST trends changing the global teleconnection during La Niña? *Geophysical Research Letters*, *37*, L12702. https://doi.org/10.1029/2010GL043394

L'Heureux, M.-L., Collins, D.-C., & Zeng-Zhen, H. (2013). Linear trends in sea surface temperature of the tropical Pacific Ocean and implications for the El Niño–Southern Oscillation. *Climate Dynamics*, *40*, 1223–1236.

L'Heureux, M. L., Lee, S., & Lyon, B. (2013). Recent multidecadal strengthening of the Walker circulation across the tropical Pacific. *Nature Climate Change*, *3*(6), 571–576.

Liu, W., Huang, B., Thorne, P. W., Banzon, V. F., Zhang, H.-M., Freeman, E., … Woodruff, S. D. (2015). Extended Reconstructed Sea Surface Temperature version 4 (ERSST. v4): Part II. Parametric and structural uncertainty estimations. *Journal of Climate*, *28*, 931–951.

Mantua, N.-J., Hare, S.-R., Zhang, Y., Wallace, J.-M., & Francis, R.-C. (1997). A Pacific interdecadal climate oscillation with impacts on salmon production. *Bulletin of the American Meteorological Society*, *78*, 1069–1079.

Newman, M., Alexander, M. A., Ault, T. R., Cobb, K. M., Deser, C., Di Lorenzo, E., … Smith, C. A. (2016). The Pacific decadal oscillation, revisited. *Journal of Climate*, *29*(12), 4399–4427.

Newman, M., Shin, S.-I., & Alexander, M.-A. (2011). Natural variation in ENSO flavors. *Geophysical Research Letters*, *38*, L14705. https://doi.org/10.1029/2011GL047658

Penland, C., & Matrasova, L. (1994). A balance condition for stochastic numerical models with application to El Niño–Southern Oscillation. *Journal of Climate*, *7*, 1352–1372.

Penland, C., & Sardeshmukh, P.-D. (1995). The optimal growth of tropical sea surface temperature anomalies. *Journal of Climate*, *8*, 1999–2024.

Rayner, N.-A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., … A. Kaplan (2003). Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research: Atmospheres*, *108*(D14), 4407. https://doi.org/10.1029/2002JD002670

Reynolds, R.-W., Smith, T. M., Liu, C., Chelton, D. B., Casey, K. S., & Schlax, M. G. (2007). Daily high-resolution-blended analyses for sea surface temperature. *Journal of Climate*, *20*, 5473–5496.

Rodgers, K.-B., Lin, J., & Frölicher, T.-L. (2015). Emergence of multiple ocean ecosystem drivers in a large ensemble suite with an Earth system model. *Biogeosciences*, *12*(11), 3301.

Sandeep, S., Stordal, F., Sardeshmukh, P.-D., & Compo, G. (2014). Pacific Walker circulation variability in coupled and uncoupled climate models. *Climate Dynamics*, *43*, 103–117.

Sarachik, E. S., & Cane, M. A. (2010). *The El Niño-southern oscillation phenomenon*. Cambridge, UK: Cambridge University Press.

Schubert, S.-D., Suarez, M.-J., Pegion, P.-J., Koster, R.-D., & Bacmeister, J.-T. (2004). On the cause of the 1930s Dust Bowl. *Science*, *303*, 1855–1859.

Seager, R. (2015). Decadal hydroclimate variability across the Americas. In *M.a.B. Climate Change* (Editor), C.-P. Chang, M. Ghil, M. Latif, and J. M. Wallace (pp. 235–253). Singapore: World Sci.

Seager, R., & Murtugudde, R. (1997). Ocean dynamics, thermocline adjustment, and regulation of tropical SST. *Journal of Climate*, *10*, 521–534.
Seager, R., & Vecchi, G.-A. (2010). Greenhouse warming and the 21st century hydroclimate of southwestern North America. *Proceedings of the National Academy of Sciences*, *107*, 21,277–21,282.
Seager, R., Ting, M., Li, C., Naik, N., Cook, B., Nakamura, J., & Liu, H. (2013). Projections of declining surface-water availability for the southwestern United States. *Nature Climate Change*, *3*(5), 482.
Sen Gupta, A., Ganachaud, A., McGregor, S., Brown, J.-N., & Muir, L. (2012). Drivers of the projected changes to the Pacific Ocean equatorial circulation. *Geophysical Research Letters*, *39*, L09605. https://doi.org/10.1029/2012GL051447
Smith, T.-M., Reynolds, R.-W., Peterson, T.-C., & Lawrimore, J. (2008). Improvements to NOAA's historical merged land-ocean surface temperature analysis (1880–2006). *Journal of Climate*, *21*, 2283–2296.
Sohn, B.-J., Lee, S., Chung, E.-S., & Song, H.-J. (2016). The role of the dry static stability for the recent change in the Pacific walker circulation. *Journal of Climate*, *29*, 2765–2779.
Solomon, A., & Newman, M. (2012). Reconciling disparate twentieth-century Indo-Pacific ocean temperature trends in the instrumental record. *Nature Climate Change*, *2*, 691–699.
Sun, D.-Z., & Liu, Z. (1996). Dynamic ocean-atmosphere coupling: A thermostat for the tropics. *Science*, *272*, 1148–1150.
Taylor, K.-E., Stouffer, R.-J., & Meehl, G.-A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, *93*, 485–498.
Thompson, D.-W., Kennedy, J.-J., Wallace, J.-M., & Jones, P.-D. (2008). A large discontinuity in the mid-twentieth century in observed global-mean surface temperature. *Nature*, *453*(7195), 646–649.
Tokinaga, H., Xie, S.-P., Deser, C., Kosaka, Y., & Okumura, Y.-M. (2012). Slowdown of the Walker circulation driven by tropical Indo-Pacific warming. *Nature*, *491*, 439–443.
Tokinaga, H., Xie, S.-P., Timmermann, A., McGregor, S., Ogata, T., Kubota, H., & Okumura, Y.-M. (2012). Regional patterns of tropical Indo-Pacific climate change: evidence of the Walker circulation weakening. *Journal of Climate*, *25*, 1689–1710.
Vecchi, G.-A., & Soden, B.-J. (2007). Global warming and the weakening of the tropical circulation. *Journal of Climate*, *20*(17), 4316–4340.
Vecchi, G.-A., Soden, B. J., Wittenberg, A. T., Held, I. M., Leetmaa, A., & Harrison, M. J. (2006). Weakening of tropical Pacific atmospheric circulation due to anthropogenic forcing. *Nature*, *441*, 73–76.
Xie, S.-P., Deser, C., Vecchi, G.-A., Ma, J., Teng, H., & Wittenberg, A.-T. (2010). Global warming pattern formation: Sea surface temperature and rainfall. *Journal of Climate*, *23*, 966–986.
Yasunaka, S., & Hanawa, K. (2011). Intercomparison of historical sea surface temperature datasets. *International Journal of Climatology*, *31*(7), 1056–1073.
Zhang, L., & Li, T. (2014). A simple analytical model for understanding the formation of sea surface temperature patterns under global warming. *Journal of Climate*, *27*, 8313–8321.
Zhang, W., Li, J., & Zhao, X. (2010). Sea surface temperature cooling mode in the pacific cold tongue. *Journal of Geophysical Research*, *115*, C12042. https://doi.org/10.1029/2010JC006501