Geophysical Research Letters

RESEARCH LETTER
10.1029/2019GL086588

Magnitudes and Spatial Patterns of Interdecadal Temperature Variability in CMIP6

Luke A. Parsons[1], M. Kathleen Brennan[1], Robert C.J. Wills[1], and Cristian Proistosescu[2,3,4]

[1]Department of Atmospheric Sciences, University of Washington, Seattle, WA, USA, [2]Joint Institute for the Study of the Atmosphere and Ocean, University of Washington, Seattle, WA, USA, [3]Department of Atmospheric Sciences, University of Illinois Urbana-Champaign, Champaign, IL, USA, [4]Department of Geology, University of Illinois Urbana-Champaign, Champaign, IL, USA

Key Points:
- Several CMIP6 control simulations show more interdecadal global mean temperature variability than previous model generations
- Even the most variable CMIP6 models never show century-length global mean temperature trends that exceed observed warming trends
- Unlike in control simulations, observed global mean temperature variability is coherent with variability in tropical convective regions

Supporting Information:
- Supporting Information S1

Correspondence to:
L. A. Parsons,
lakp@uw.edu

Citation:
Parsons, L. A., Brennan, M. K., Wills, R. C. J., & Proistosescu, C. (2020). Magnitudes and spatial patterns of interdecadal temperature variability in CMIP6. *Geophysical Research Letters*, *47*, e2019GL086588. https://doi.org/10.1029/2019GL086588

Received 6 DEC 2019
Accepted 25 MAR 2020

Abstract Attribution and prediction of global and regional warming requires a better understanding of the magnitude and spatial characteristics of internal global mean surface air temperature (GMST) variability. We examine interdecadal GMST variability in Coupled Modeling Intercomparison Projects, Phases 3, 5, and 6 (CMIP3, CMIP5, and CMIP6) preindustrial control (piControl), last millennium, and historical simulations and in observational data. We find that several CMIP6 simulations show more GMST interdecadal variability than the previous generations of model simulations. Nonetheless, we find that 100-year trends in CMIP6 piControl simulations never exceed the maximum observed warming trend. Furthermore, interdecadal GMST variability in the unforced piControl simulations is associated with regional variability in the high latitudes and the east Pacific, whereas interdecadal GMST variability in instrumental data and in historical simulations with external forcing is more globally coherent and is associated with variability in tropical deep convective regions.

Plain Language Summary Ongoing and future global and regional warming will progress as a combination of internal climate variability and forced climate change. Understanding the magnitude and spatial patterns associated with internal climate variability is an important aspect of being able to predict when, where, and how climate change will be felt around the globe. Here, we show that the latest climate model simulations, which will be used in the Intergovernmental Panel on Climate Change (IPCC) Assessment Report 6 (AR6), simulate a large range in magnitudes of internal global mean temperature variability. Although there are large unforced global temperature trends in some models, we find that even the most variable models never generate unforced global temperature trends equal to the recently observed global warming trends forced by greenhouse gas emissions. We examine the regions associated with internal climate variability and forced climate change in climate model simulations and find that only forced simulations show a pattern of warming consistent with instrumental data.

1. Introduction

Future global and regional temperature variability will manifest itself as a combination of unforced internal variability and externally forced changes (e.g., warming from greenhouse gases, short-lived cooling following volcanic eruptions). Understanding the magnitude and physical mechanisms that drive internal variability is critical for advancing knowledge about how future patterns of global warming will develop. If internal climate variability is weak relative to global warming, then future climate change on regional and global scales may manifest itself as a relatively smooth warming trend. By contrast, if internal decadal variability is large relative to externally forced variability, then warming hiatus decades could follow rapid warming (e.g., Dai et al., 2015; Shuman, 2012; Trenberth & Fasullo, 2013).

Correctly simulating the magnitude of internal, natural variability in preindustrial control (piControl) simulations is important for climate change detection and attribution efforts (e.g., Bindoff et al., 2013; Santer et al., 1995; Stott et al., 2010), for projections of climate change impacts, and for decadal prediction efforts (Hu & Deser, 2013; Meehl et al., 2009; Screen et al., 2014). Additionally, understanding the relative contributions of internal and externally forced climate changes to local climate variations (e.g., Goosse et al., 2005) is critical for studying how and when climate change "signals" emerge from the "noise" of internal variability

(e.g., Deser et al., 2012; Deser et al., 2014; Hawkins & Sutton, 2012; PAGES2k Consortium, 2019). Interdecadal variability in particular is important for understanding the degree to which the observed warming trend is attributable to increased greenhouse gas concentrations (e.g., Stouffer et al., 1994) and how to interpret estimates of future climate sensitivity drawn from observations of the Earth's energy budget (Dessler et al., 2018; Marvel et al., 2018; Otto et al., 2013). Interdecadal global mean surface temperature (GMST) variability is also correlated with long-term climate sensitivity in both forced (Cox et al., 2018) and unforced (Nijsse et al., 2019) model simulations.

Climate model simulations and instrumental-based data tend to agree that most GMST variability on subdecadal timescales is associated with the El Niño Southern Oscillation (e.g., Brown et al., 2015). Yet, there is less agreement among climate models and observations in terms of the magnitude and spatial patterns associated with interdecadal GMST variability (Brown et al., 2015, 2017; Parsons & Hakim, 2019), with both instrumental (Laepple & Huybers, 2014a) and paleoclimate (Dee et al., 2017; Laepple & Huybers, 2014b; Parsons et al., 2017) evidence suggesting that climate models may underestimate local, low-frequency climate variability. Here we characterize the magnitude and spatial patterns of interdecadal GMST variability in the new generation of Coupled Model Intercomparison Project, Phase 6 (CMIP6) models and compare these results to instrumental-based data and simulations from Coupled Model Intercomparison Project, Phase 3 and Phase 5 (CMIP3 and CMIP5) models.

2. Data and Methods

2.1. Surface Temperature Data

We examine local and global mean annual mean surface air temperature variability in CMIP3 (Meehl et al., 2007), CMIP5 (Taylor et al., 2012), and CMIP6 (Eyring et al., 2016) simulations. We use the atmospheric skin surface temperature (TS) variable from the last 330 years from 15 CMIP3 piControl simulations and the last 400 years from 30 CMIP5 and 39 CMIP6 piControl simulations (Table S1). Although 39 CMIP6 models had at least 400 years of piControl simulation data at the time of analysis, more piControl simulations are expected to be released. We have replicated our analysis on the 2-m surface air temperature (TAS) variable, and the magnitudes of TS and TAS variability are nearly identical at interdecadal timescales (see also, Manabe & Stouffer, 1996). We compare piControl GMST variability to variability from 10 CMIP5 past1000 (last millennium) simulations, which are driven by preindustrial forcing of the last millennium reconstructed from the paleoclimate record (Schmidt et al., 2011). We also compare piControl simulations to GMST variability reconstructed from paleoclimate proxy records (PAGES2k, 2019). The proxy-based GMSTs are generated using eight reconstruction methods (methods and references in PAGES2k, 2019). For each iteration (N = 1,000) of eight reconstruction methods, we calculate the standard deviation of GMST, then show the full range of these 8,000 estimates of interdecadal variability. We limit our analysis of the last millennium to the time period 1450–1849 Common Era (CE) to include a similar time series length used in the piControl calculations and to maximize proxy record availability, which tends to decrease earlier in time (e.g., Anderson et al., 2019).

We also examine interdecadal GMST variability in five instrumental-based data sources: National Aeronautics Space Administration (NASA) Goddard Institute of Space Studies (GISS) surface temperature analysis (GISTEMP; Hansen et al., 2010), the Berkeley Earth surface temperatures (BE; Rohde et al., 2013), the 20th century reanalysis v3 surface temperatures (20CRv3, Slivinski et al., 2019), the Hadley Centre Climate Research Unit global temperature data set v4 (HadCRUT4; Morice et al., 2012), and the National Oceanic and Atmospheric Administration (NOAA) Global Temperatures v5 (NOAAGlobalTempv5; Smith et al., 2008). We compare variability from instrumental-based data and the piControl simulations to variability from the CMIP6 historical (1880–2015 CE), the CMIP5 historical (1880–2005 CE), and CMIP3 historical (1900–1999 CE) simulations. The historical simulations are forced by observed concentrations of atmospheric greenhouse gases and aerosols, among other time-varying external forcings (Eyring et al., 2016; Taylor et al., 2012). Although many CMIP historical experiments started in 1850, here we use the time period after 1880 to maximize overlap with the instrumental data sets.

2.2. Calculating Global Mean Temperature Variability

In all piControl and past1000 simulations, we remove the linear TS trend from each grid point before calculating GMST to remove the long-term "drift" or spin-up trends from model output (e.g., Gupta et al., 2013).

Table S1 lists GMST trends for each CMIP6 piControl simulation before local trend removal. We also remove the linear trend (1450–1850 CE) from each paleoclimate-based GMST time series. Although linear trend removal is used to remove spin-up model "drift" in the past1000 simulations, this procedure could also remove millennial trends from external forcing (e.g., Kaufman et al., 2009; Stevenson et al., 2019). To quantify the magnitude of interdecadal variability in GMST time series, we use a third-order, 25-year low-pass, Butterworth filter (Butterworth, 1930) to smooth all GMST, then calculate the standard deviation of the low-pass filtered GMST time series. Throughout the text, when we refer to "interdecadal" variability, we are referring to variability beyond 25-year timescales (>25 years, including interdecadal to centennial timescales), but results are not particularly sensitive to the choice of low-pass filter cutoff (e.g., 10-, 20-, or 25-year filter).

2.3. Calculating Coherence of Global Mean and Local Temperature Variability

We investigate the geographic regions associated with GMST variability (e.g., Brown et al., 2016; Manabe & Stouffer, 1996; Mann & Park, 1993; Stouffer et al., 2000) using a multitaper spectral coherence approach (Mann & Park, 1993; Parsons & Hakim, 2019; Thomson, 1982) in CMIP3, CMIP5, and CMIP6 piControl simulations and CMIP6 historical simulations (1915–2014 CE; Table S1), and gridded instrumental-based surface temperature data (GISTEMP, BE, 20CRv3 1915–2014 CE). In the instrumental-based data and the CMIP6 historical simulations, we do not remove the linear trends from local TS prior to conducting coherence calculations in order to examine the differences in the patterns of coherence between forced and unforced variability.

We use the coherence and phase of local TS time series and GMST in time-frequency space to test where local variability shows a significant lead-lag relationship with global variability. At a given frequency f, coherency estimates C(f) range between 0 and 1 with increasing values when the local and global time series covary at a given frequency. We also use the phase between the local and global mean variability to test the lead-lag relationship between these time series. Regions in coherency maps are stippled where two criteria are met: (i) local versus global coherency exceeds a noise threshold at the 95% confidence level and (ii) the phase relationship between local and GMST shows that local variability either leads GMST variability (i.e., phase between −90° and 0°) or local variability lags GMST by up to a year. Coherency calculated on randomized time series shows that interdecadal coherence must exceed 0.46 to surpass the 95% noise threshold in time series 400 years in length, and interdecadal coherence must exceed 0.52 to surpass the 95% noise threshold in time series 100 years in length. Although coherence is calculated at multiple frequency ranges, here we use the average coherency and phase at interdecadal ($f < 0.04$ year^{-1}, e.g., >25 years) timescales. Coherence results are not particularly sensitive to our choice of frequency cutoff (e.g., 10, 20, or 25 years), but we find that results are most stable beyond 20-year timescales. CMIP and instrumental data are regridded to a ~2° × 2° grid before coherence is calculated.

2.4. Trends in Global Mean Surface Temperatures in Model and Instrumental Data

Given the wide range of internal GMST variability in the CMIP6 simulations (Figure 1), we test if the magnitude of internal piControl GMST variability exceeds the recent forced GMST trends. We compare GMST trends in the 400 years from the CMIP6 piControl simulations to trends from the GISTEMP (1880–2018 CE), BE (1880–2018 CE), 20CRv3 (1880–2015 CE), HadCRUT4 (1880–2018 CE), and NOAAGlobalTempv5 (1880–2018 CE). Although several of these data sets include pre-1880 data, we focus our analysis on the time period post-1880 to use similar time periods for all data sets. We use a 100-year sliding window to calculate GMST trends in the CMIP6 and instrumental-based data and show the distributions of these trends. We have tested the sensitivity of our results to the width of this trend window and find that the maximum observed trends tend to exceed maximum CMIP6 piControl trends at window widths greater than 30 years (Figure S1).

3. Results

3.1. Magnitude of Interdecadal Variability and Trends

The standard deviation of interdecadal GMST variability in some CMIP6 models (e.g., EC-Earth3, BCC-CSM2-MR, CNRM-CM6-1) is 0.14 to 0.23 °C, or about three to six times larger than the GMST interdecadal variability in the "least variable" models (e.g., CESM-FV2, SAM0-UNICON, CAMS-CSM1), which

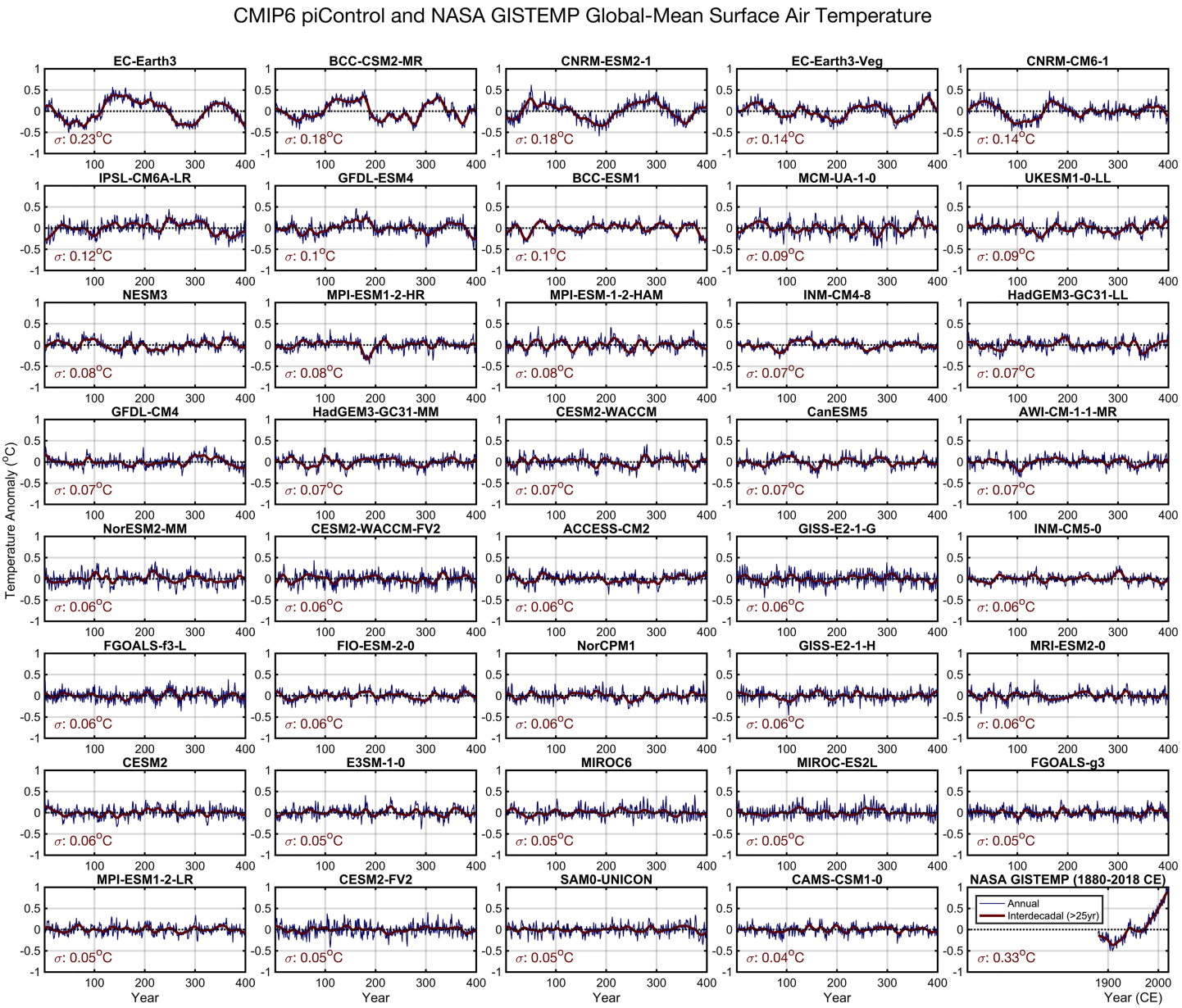

Figure 1. Time series of global mean surface air temperature (GMST) anomalies in 39 Coupled Model Intercomparison Project, Phase 6 (CMIP6) preindustrial control (piControl) simulations and the observation-based NASA GISTEMP data set. Blue lines show unfiltered, annual mean GMST, and red lines show 25-year low-pass filtered GMST (section 2.2). Anomalies are with respect to the long-term mean GMST in CMIP6 simulations and the 1951–1980 CE mean in the GISTEMP data. Red numbers in lower left of each panel show standard deviations of interdecadal GMST.

show interdecadal variability of ~0.04 °C to 0.05 °C (Figure 1). Although several models simulate unforced GMST interdecadal variability greater than 0.10 °C, most models show less variability, with a median value of 0.07 °C. We contextualize this range of CMIP6 variability using CMIP3 and CMIP5 piControl simulations, transiently forced CMIP5 and CMIP6 simulations, the instrumental record, and paleoclimate-based reconstructions of GMST (Figure 2a). The distributions of CMIP5 and CMIP6 GMST variability are similar, but several CMIP6 piControl simulations show more GMST variability than previous generations of model simulations. Furthermore, the most variable CMIP6 models show more

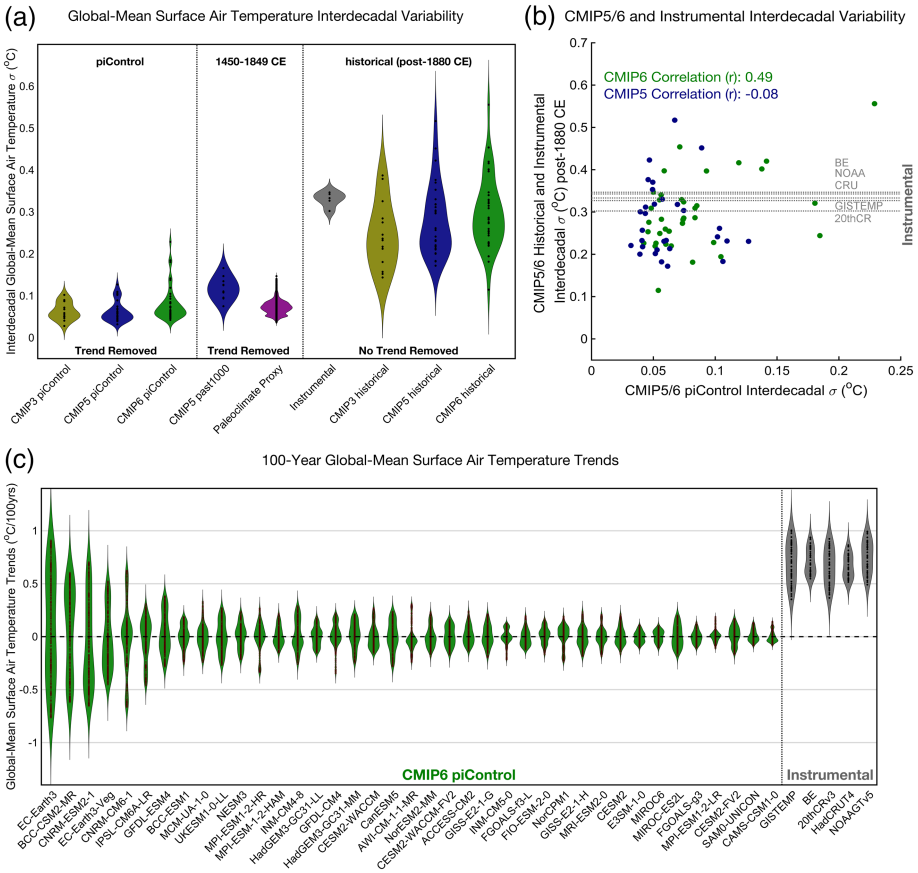

Figure 2. Violin plots showing the distribution of standard deviations of interdecadal global mean surface air temperatures (GMST) (a), scatter plot showing relationship between piControl and historical simulation GMST variability (b), and violin plots showing 100-year running GMST trends (c). In (b), blue dots show CMIP5, green dots show CMIP6, and horizontal, grey lines show instrumental-based estimates of GMST interdecadal variability. Dots in violin plots mark values used to make violin plots, which show the probability density of the data at different values and are smoothed by a kernel density estimator. Note in (a) that the spread in climate model results represents uncertainty in model physics related to internal variability, whereas the spread in instrumental-based results shows sampling uncertainty, observational uncertainty, and uncertainty in the interpolation methodology.

GMST variability than paleoclimate-based GMST reconstructions and CMIP5 last millennium simulations (1450–1849 CE; Figure 2a).

When CMIP6 models are forced with observed greenhouse gases and aerosols of the industrial era (post-1880 CE), they are able to simulate interdecadal variability that is similar to the instrumentally observed magnitude of forced GMST variability (Figures 2a and 2b). Furthermore, there is little to no relationship among the magnitude of piControl interdecadal GMST variability and the magnitude of interdecadal variability in the historical simulations (Figure 2b), indicating that forcing in historical simulations has a larger influence on GMST variability than the internal variability of the model.

The CMIP6 models showing the most interdecadal GMST variability appear to simulate internal GMST trends that could approach the recent forced warming observed in the industrial era (Figure 1). We compare centennial GMST trends from CMIP6 piControl to GMST trends from the instrumental-based data post-1880 CE (e.g., Manabe & Stouffer, 1996). In the CMIP6 models that simulate the most interdecadal variability (left violin plots in Figure 2c), isolated GMST trends approach the recently observed warming trends. However, most models tend to show no piControl trends that even approach recent warming (Figure 2c). Out of 11,700 overlapping centuries from the CMIP6 piControl simulations analyzed here, 3.3% of simulated 100-year trends in these simulations exceed the minimum observed 100-year trend (0.35 °C/100 years, 1880–1979 CE in GISTEMP), 0.5% of simulated trends exceed the observed median

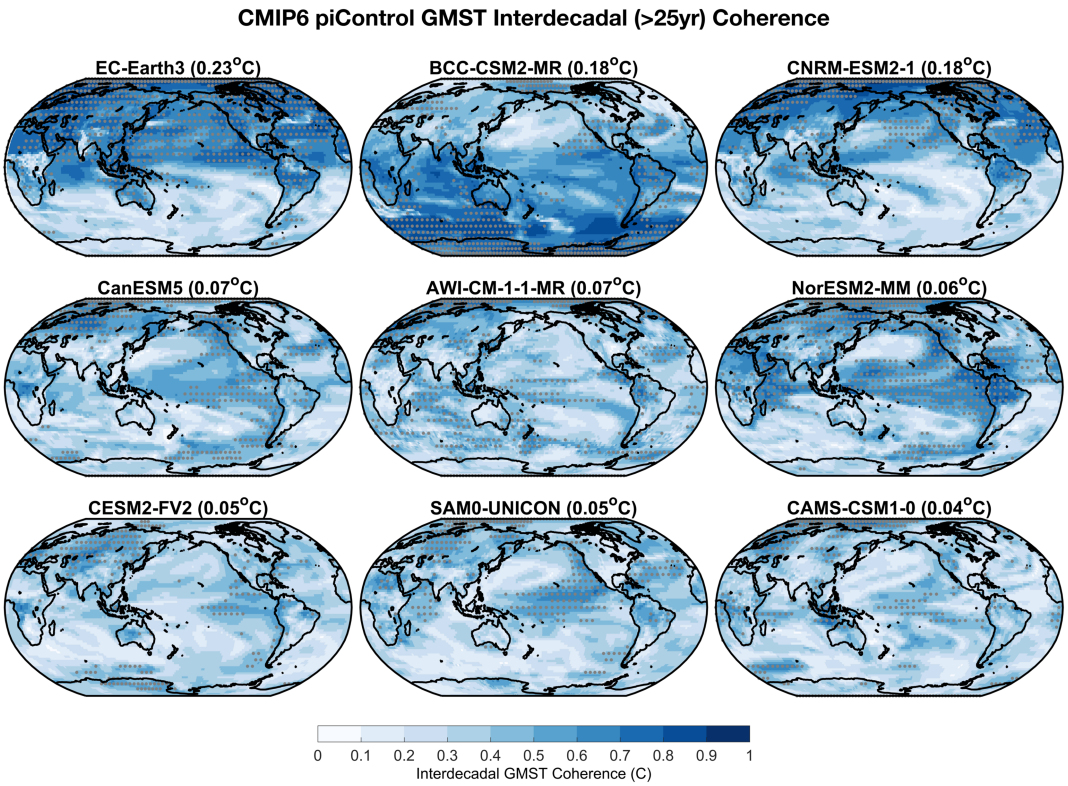

Figure 3. Spectral coherence C(f) of local temperature and global mean surface air temperature (GMST) at interdecadal ($f < 0.04$ year^{-1}) timescales in the last 400 years of nine Coupled Model Intercomparison Project, Phase 6 (CMIP6) preindustrial control (piControl) simulations. Stippling denotes local geographic regions where the global versus local relationship exceeds the noise threshold and local variability leads global variability at interdecadal timescales (section 2.3). Note that these panels show a subselection of CMIP6 models, including the three most, three median, and three least variable models as measured by standard deviation of GMST. Figure S2 shows coherence from all 39 CMIP6 piControl simulations.

GMST trend (0.69 °C/100 years, ~1899–1998 CE), and no simulated piControl 100-year trends equal or exceed the maximum observed global warming trend of the last 100 years (1.0 °C/100 years, 1919–2018 CE in GISTEMP). These results are not particularly sensitive to trend window width; the maximum observed warming trends consistently exceed almost all CMIP6 piControl trends at window widths greater than 30 years (Figure S1).

3.2. Geographic Regions Associated With Interdecadal Variability

Simulating a realistic magnitude of GMST variability across a range of timescales is an important target for climate models (e.g., Zhu et al., 2019). However, it is also important to simulate the appropriate mechanisms and geographic regions associated with this GMST variability. We examine the geographic distribution of interdecadal GMST variability using a spectral coherence method (section 2.3). In the CMIP6 piControl simulations showing the most interdecadal GMST variance (top row, Figure 3), GMST is most coherent with high-latitude temperature variability. By contrast, in models that simulate less interdecadal variability, GMST is most coherent with the eastern and central tropical Pacific as well as regions associated with the Pacific Decadal Oscillation (PDO) or Interdecadal Pacific Oscillation (IPO) (e.g., Deser et al., 2010). These "lower variability" models also tend to show weaker local coherence with GMST at interdecadal timescales (middle, bottom rows, Figure 3).

Although no CMIP6 piControl simulations show internal GMST trends that exceed observed GMST trends 1919–2018, some CMIP6 piControl simulations show GMST trends that approach or equal the observed trends in the 19th and 20th centuries (Figure 2). Here, we investigate if unforced, internal variability is as globally coherent as recent global warming. The maps in Figure 4 show the mean coherence of GMST interdecadal variability from the CMIP3/5/6 piControl simulations (Figures S2–S4 show results for individual

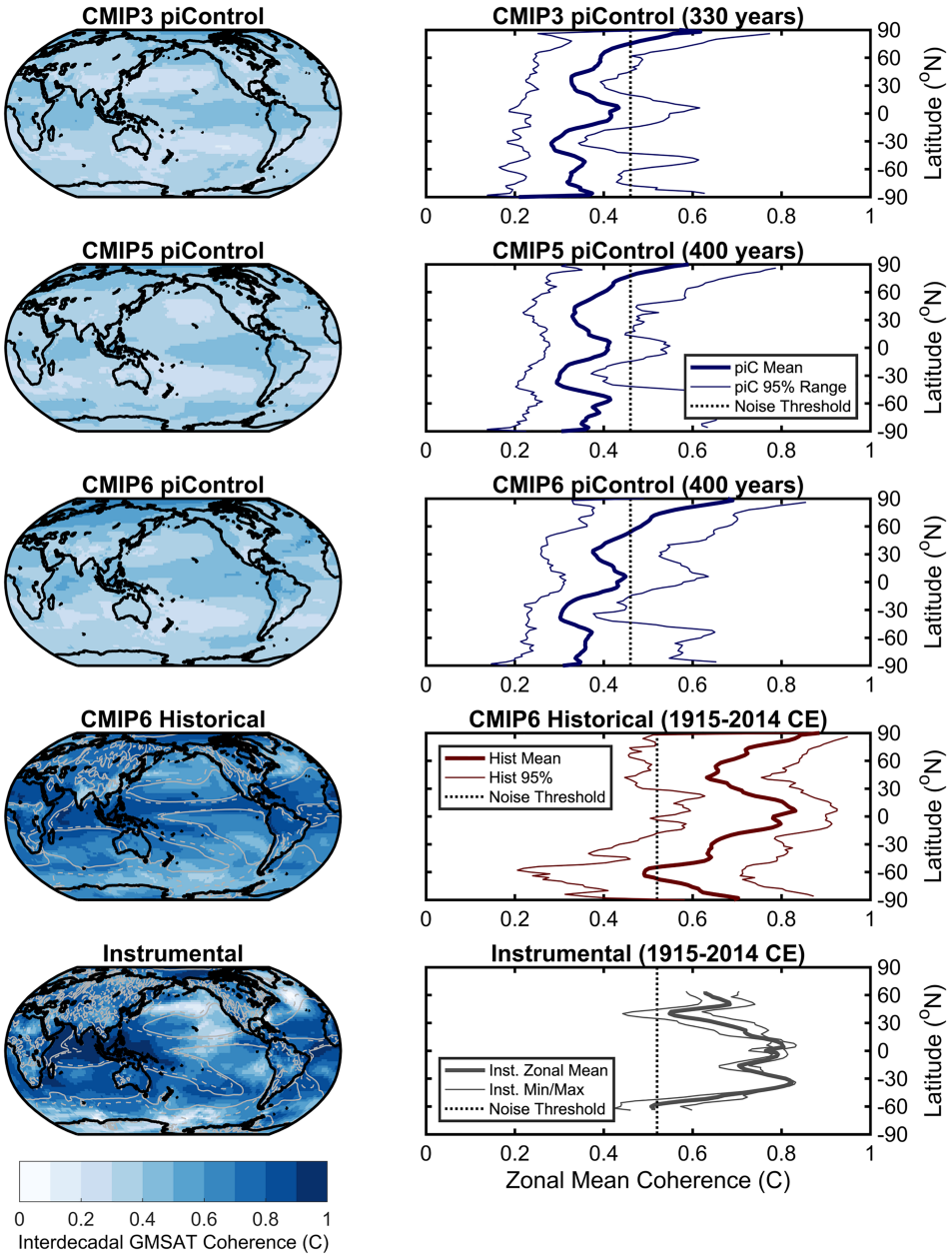

Figure 4. Spectral coherence as in Figure 2 (left) and zonal mean coherence (right) from Coupled Model Intercomparison Project, Phases 3, 5, and 6 (CMIP3, CMIP5, CMIP6) preindustrial control (piControl) simulations (top three rows), CMIP6 historical simulations (second from bottom), and instrumental-based data (GISTEMP, BE, 20CRv3, bottom). Vertical dashed line in zonal mean panels (right) show spectral coherence 95% noise threshold. Instrumental zonal mean plots do not show data poleward of 65° due to a lack of data. Dashed and solid gray contours outline regions of upward and, respectively, downward vertical motion (ω) at 500mb in the last 30 years of the CMIP6 historical simulations (second row from bottom) and the 20th century reanalysis (bottom row).

models across model generations), 37 CMIP6 historical simulations (1915–2014 CE; Figure S5 shows results from individual CMIP6 historical simulations), and the instrumental-based data (1915–2014 CE).

We find that the magnitude of GMST interdecadal coherence with local variability in the control simulations is, on average, about half the magnitude of GMST coherence in the forced CMIP6 historical runs and in the instrumental-based data. CMIP6 piControl and historical simulations both show elevated coherence of

GMST and local temperatures north of ~60°N, with a maximum in the high-latitude Arctic. However, the zonal mean coherence plots (Figure 4, right) show that for piControl variability, most of the low latitudes and Southern Hemisphere do not cross the 95% noise threshold for the coherence calculations in the multimodel mean. By contrast, the CMIP6 historical simulations and instrumental-based data show interdecadal GMST coherence that consistently exceeds the noise threshold across most of the tropics and into the midlatitudes, with coherence maxima near the equator and poles in the multimodel mean. Therefore, although certain CMIP6 piControl simulations may show what appears to be "global" temperature trends that approach the instrumentally observed trends of the early 20th century, this internal variability is not as globally coherent as the externally forced global warming trends simulated by these CMIP6 models or observed in the instrumental-based data.

The spatial patterns of coherence with GMST also show striking differences between unforced and forced simulations. In the ensemble mean of control simulations, GMST is most coherent with variability over the Arctic, North Atlantic, extratropical Pacific (PDO region), and regions associated with atmospheric teleconnections from the tropical Pacific. Historical simulations also show significant coherence between Arctic temperatures and global temperatures. However, forced simulations show the highest coherence between GMST and low-latitude regions of ascent and deep convection in the atmosphere, such as the Indo-Pacific Warm Pool and the Inter-Tropical and South Pacific Convergence Zones (ITCZ, SPCZ). Forced CMIP5 last millennium simulations also show geographically similar increases in coherence in the tropics relative to the control runs (Figure S7). The coherence from the instrumental record is more consistent with this forced (historical, last millennium) coherence pattern in models, with strong coherence in the Indo-Pacific Warm Pool and tropical Atlantic, and a notable lack of coherence with the extratropical North Pacific and North Atlantic and with the central and eastern Pacific ITCZ region (Figure 4).

The CMIP6 coherence patterns in both unforced (Figure 3) and forced (Figure 4) simulations are consistent across CMIP generations; forced simulations show the strongest interdecadal coherence in the tropics, and unforced simulations show coherence in the IPO region or at high latitudes. Indeed, piControl simulations from individual models show qualitatively similar spatial coherence patterns across model generations (e.g., GFDL CM2.0, CM3, ESM 4; BCC-CSM1-1, BCC-CSM2-MR; CNRM-CM5, CNRM-CM6-1; CanESM2, CanESM5), although a few models show distinct coherence patterns (e.g., GISS-E2-H, GISS-E2-1-H). Most CMIP6 piControl simulations (>60%, or >23/39) agree that the Arctic and North Atlantic show a significant relationship with GMST variability at interdecadal timescales (stippling in last panel in Figure S2). However, agreement across models should be interpreted with caution because these CMIP6 piControl simulations include several similar models from the same modeling groups, and models from different groups are not necessarily independent from one another (e.g., Knutti et al., 2013).

To understand the spatial pattern of temperature coherence with interdecadal GMST variability in the CMIP5/6 historical and piControl simulations, we compare the standard deviations of local interdecadal TS variability in these simulations (Figure S8). CMIP5 and CMIP6 piControl simulations show the most internal interdecadal TS variability over high latitudes and the least internal interdecadal variability over the tropics. Although forcing in historical simulations causes the largest increases in interdecadal TS variability over high latitudes, the largest relative (percent) increase in forced interdecadal variability occurs over regions of deep convection, where background TS variability in the piControl simulations is near zero (Figure S8). Therefore, almost all of the interdecadal variability in these regions of deep convection is forced in historical and last millennium simulations.

4. Discussion and Conclusions

Here we examine the magnitudes of interdecadal GMST variability and the geographic regions associated with this variability in CMIP3, CMIP5, and CMIP6 simulations. Many of the most variable CMIP6 models (e.g., BCC-CSM2-MR, CNRM-ESM 2-1, GFDL-ESM 4) are updated versions of some of the most variable CMIP5 models, and several of the least variable CMIP6 models (e.g., FGOALS-g3, MIROC-ES2L) similarly are updated versions of some of the least variable CMIP5 models. Figures S2–S4 show the magnitudes of interdecadal variability in CMIP piControl simulations, as well as the geographic regions associated with this variability across model generations. Although most CMIP6 models show interdecadal variability similar to previous generations of models, several CMIP6 models (e.g., EC-Earth3, BCC-CSM2-MR, CNRM-ESM

2-1, EC-Earth3-Veg, CNRM-CM6-1) simulate more internal interdecadal GMST variability than any of the CMIP5 piControl simulations, the transiently forced CMIP5 last millennium simulations, or paleoclimate reconstructions of GMST of the last millennium (linear trends removed 1450–1850 CE), suggesting that several CMIP6 models may simulate unrealistically high internal interdecadal variability (Figure 2). It is also important to note that the magnitude of GMST interdecadal variability can vary substantially (~5–24%) from century to century in CMIP6 piControl simulations (Figure S6). Nonetheless, when CMIP6 models are forced with observed transient forcing from the industrial era ("historical" simulations), they generate interdecadal variability similar to the instrumental-based data (Figures 2a and 2b).

Although the magnitude of internal interdecadal GMST variability in the CMIP6 piControl simulations may approach observed warming trends in the late 19th and early 20th centuries (Figures 1 and 2c), the spatial patterns associated with this variability are quite different in the forced and unforced climate system. We examine the geographic regions associated with interdecadal GMST variability in the piControl simulations, historical simulations, and instrumental-based data. Most CMIP piControl simulations show the strongest local coherence with GMST interdecadal variability over high-latitude regions (Figures 3 and 4; Figures S2–S4). By contrast, forced interdecadal variability (e.g., warming from greenhouse gases post-1900 CE) tends to show high coherence across the tropics (Figure S5). Indeed, when we compare coherence in the CMIP5/6 historical simulations to coherence in the piControl simulations from these models, we find large increases in coherence over most of the globe, with the exception of the Southern Ocean, northeastern Pacific, and extratropical North Atlantic (Figure S7). Similarly, transient forcing of the last millennium 850–1850 CE (e.g., short-lived cooling from volcanic eruptions) shows large increases in coherence over the tropics (Figure S7; Parsons & Hakim, 2019).

The interdecadal GMST coherence pattern in simulations with external forcing (i.e., historical and last millennium) shows a strong fingerprint of external forcing, marked by high coherence in regions of deep convection (Figure 4). This pattern is generally not present in unforced simulations, with CMIP5/6 piControl simulations showing minimal internal variability over regions of deep convection (Figure S8). This fingerprint of forced interdecadal coherence could present a useful avenue for further analysis of forced and unforced variability in the climate system. The lack of internal variability in regions of deep convection may be due to the fact that surface warming in these regions leads to very strong radiative damping (Dong et al., 2019; Zhou et al., 2017), and the magnitude of variability is inversely proportional to the strength of radiative damping (e.g., Roe, 2009). These results suggest that sustained interdecadal temperature anomalies in these regions of deep convection lead to large changes in Earth's energy budget, which can only be supported by a sustained external forcing. However, over the ITCZ in the eastern tropical Pacific, models forced with greenhouse gases indicate that local TS variability should be coherent with global temperatures and, by extension, with west Pacific temperatures, but this pattern is not found in the instrumental-based data. This instrumental-model mismatch is likely related to the inability of models to reproduce changes in the east-west tropical Pacific temperature gradients (e.g., Coats & Karnauskas, 2017).

In summary, CMIP6 piControl simulations show a wide range of magnitudes of interdecadal GMST variability. Nonetheless, even the most variable CMIP6 piControl simulations never show century-length GMST trends that exceed observed warming trends. Furthermore, the geographic coherence of forced global mean variability is distinct from the geographic coherence of unforced variability. In order to determine if the large magnitude of interdecadal GMST variability in some CMIP6 models is realistic, future works should examine the mechanisms of this variability and the causes of the spread in its magnitude across CMIP5/6 piControl simulations.

Acknowledgments

L.A.P. thanks the Washington Research Foundation (WRF) for funding support. M.K.B. was supported by an NSF graduate research fellowship (Grant 1650114). C.P. was supported by a JISAO postdoctoral fellowship and the National Science Foundation (Grant AGS-1752796). R.C.J.W was supported by the National Science Foundation (Grant AGS-1929775). We also thank R. Stouffer, G. Hakim, D. Frierson, R. Wood, and S. Sanchez for valuable input, and P Huybers for the code to calculate multitaper coherence. *Instrumental-based surface temperature data provided by the Climate Research Unit, University of East Anglia (*https://crudata.uea.ac.uk/cru/data/temperature/*) and the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA (*https://www.esrl.noaa.gov/psd/). We acknowledge the World Climate Research Program's Working Group on Coupled Modeling, which is responsible for the CMIP, and we thank the climate modeling groups 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 Portals. CMIP data can be found at https://esgf-node.llnl.gov/projects/esgf-llnl/.

References

Anderson, D., Tardif, R., Horlick, K., Erb, M., Hakim, G., Noone, D., et al. (2019). Additions to the last millennium reanalysis multi-proxy database. *Data Science Journal*, *18*.

Bindoff, N. L., Stott, P. A., AchutaRao, K. M., Allen, M. R., Gillett, N., Gutzler, D., et al. (2013). *Detection and attribution of climate change: from global to regional*.

Brown, P. T., Li, W., Jiang, J. H., & Su, H. (2016). Spread in the magnitude of climate model interdecadal global temperature variability traced to disagreements over high-latitude oceans. *Geophysical Research Letters*, *43*.

Butterworth, S. (1930). On the theory of filter amplifiers. *Wireless Engineer*, *7*, 536–541.

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

Cox, P. M., Huntingford, C., & Williamson, M. S. (2018). Emergent constraint on equilibrium climate sensitivity from global temperature variability. *Nature*, *553*, 319–322. https://doi.org/10.1038/nature25450

Dai, A., Fyfe, J. C., Xie, S., & Dai, X. (2015). Decadal modulation of global surface temperature by internal climate variability. *Nature Climate Change*, *5*, 555.

Dee, S. G., Parsons, L. A., Loope, G. R., Overpeck, J. T., Ault, T. R., & Emile Geay, J. (2017). Improved spectral comparisons of paleoclimate models and observations via proxy system modeling: Implications for multi-decadal variability. *Earth and Planetary Science Letters*, *476*, 34–46. http://doi.org/10.1016/j.epsl.2017.07.036

Deser, C., Alexander, M. A., Xie, S.-P., & Philips, A. S. (2010). Sea surface temperature variability: Patterns and mechanism. *Annual Review of Marine Science*, *2*, 115–143. https://doi.org/10.1146/annurev-marine-120408-151453

Deser, C., Phillips, A., Bourdette, V., & Teng, H. (2012). Uncertainty in climate change projections: The role of internal variability. *Climate Dynamics*, *38*, 527–546.

Deser, C., Phillips, A. S., Alexander, M. A., & Smoliak, B. V. (2014). Projecting north American climate over the next 50 years: Uncertainty due to internal variability. *Journal of Climate*, *27*, 2271–2296.

Dessler, A. E., Mauritsen, T., & Stevens, B. (2018). The influence of internal variability on Earth's energy balance framework and implications for estimating climate sensitivity. *Atmospheric Chemistry and Physics*, *18*, 5147–5155.

Dong, Y., Proistosescu, C., Armour, K. C., & Battisti, D. S. (2019). Attributing historical and future evolution of radiative feedbacks to regional warming patterns using a Green's function approach: The preeminence of the Western Pacific. *Journal of Climate*, *32*(17), 5471–5491.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the coupled model Intercomparison project phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development (Online)*, *9*.

Goosse, H., Renssen, H., Timmermann, A., & Bradley, R. S. (2005). Internal and forced climate variability during the last millennium: A model-data comparison using ensemble simulations. *Quaternary Science Reviews*, *24*, 1345–1360.

Gupta, A. S., Jourdain, N. C., Brown, J. N., & Monselesan, D. (2013). Climate drift in the CMIP5 models. *Journal of Climate*, *26*, 8597–8615.

Hansen, J., Ruedy, R., Sato, M., & Lo, K. (2010). Global surface temperature change. *Reviews of Geophysics*, *48*.

Hawkins, E., & Sutton, R. (2012). Time of emergence of climate signals. *Geophysical Research Letters*, *39*.

Hu, A., & Deser, C. (2013). Uncertainty in future regional sea level rise due to internal climate variability. *Geophysical Research Letters*, *40*, 2768–2772.

Kaufman, D. S., Schneider, D. P., McKay, N. P., Ammann, C. M., Bradley, R. S., Briffa, K. R., et al. (2009). Recent warming reverses long-term Arctic cooling. *Science*, *325*(5945), 1236–1239. https://doi.org/10.1126/science.1173983

Knutti, R., Masson, D., & Gettelman, A. (2013). Climate model genealogy: Generation CMIP5 and how we got there. *Geophysical Research Letters*, *40*(6), 1194–1199.

Laepple, T., & Huybers, P. (2014a). Global and regional variability in marine surface temperatures. *Geophysical Research Letters*, *41*, 2528–2534. https://doi.org/10.1002/2014GL059345

Laepple, T., & Huybers, P. (2014b). Ocean surface temperature variability: Large model–data differences at decadal and longer periods. *Proceedings of the National Academy of Sciences*, *111*(47), 16,682–16,687. http://doi.org/10.1073/pnas.1412077111

Manabe, S., & Stouffer, R. (1996). Low-frequency variability of surface air temperature in a 1000-year integration of a coupled atmosphere-ocean-land surface model. *Journal of Climate*, *9*, 376–393.

Mann, M. E., & Park, J. (1993). Spatial correlations of interdecadal variation in global surface temperatures. *Geophysical Research Letters*, *20*, 1055–1058.

Marvel, K. D., Pincus, R., Schmidt, G. A., & Miller, R. L. (2018). Internal variability and disequilibrium confound estimates of climate sensitivity from observations. *Geophysical Research Letters*, *45*(3), 1595–1601. https://doi.org/10.1002/2017GL076468

Meehl, G. A., Covey, C., Delworth, T., Latif, M., McAvaney, B., Mitchell, J. F., et al. (2007). The WCRP CMIP3 multimodel dataset: A new era in climate change research. *Bulletin of the American Meteorological Society*, *88*, 1383–1394.

Meehl, G. A., Goddard, L., Murphy, J., Stouffer, R. J., Boer, G., Danabasoglu, G., et al. (2009). Decadal prediction: Can it be skillful? *Bulletin of the American Meteorological Society*, *90*, 1467–1486.

Morice, C. P., Kennedy, J. J., Rayner, N. A., & Jones, P. D. (2012). Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 data set. *Journal of Geophysical Research*, *117*, D08101. https://doi.org/10.1029/2011JD017187

Nijsse, F. J., Cox, P. M., Huntingford, C., & Williamson, M. S. (2019). Decadal global temperature variability increases strongly with climate sensitivity. *Nature Climate Change*, *9*, 598–601.

Otto, A., Otto, F. E., Boucher, O., Church, J., Hegerl, G., Forster, P. M., et al. (2013). Energy budget constraints on climate response. *Nature Geoscience*, *6*, 415.

PAGES 2k Consortium (2019). Consistent multi-decadal variability in global temperature reconstructions and simulations over the common era. *Nature Geoscience*, *12*, 643–649.

Parsons, L., & Hakim, G. (2019). Local regions associated with interdecadal global temperature variability in the last millennium reanalysis and CMIP5 models. *Journal of Geophysical Research: Atmospheres*, *124*(17-18), 9905–9917.

Parsons, L. A., Loope, G. R., Overpeck, J. T., Ault, T. R., Stouffer, R., & Cole, J. E. (2017). Temperature and precipitation variance in CMIP5 simulations and paleoclimate records of the last millennium. *Journal of Climate*, *30*(22), 8885–8912.

Roe, G. (2009). Feedbacks, timescales, and seeing red. *Annual Review of Earth and Planetary Sciences*, *37*, 93–115.

Rohde, R., Muller, R., Jacobsen, R., Muller, E., Perlmutter, S., Rosenfeld, A., et al. (2013). A new estimate of the average Earth surface land temperature spanning 1753 to 2011. *Geoinfor Geostat: An Overview*, *1*(1).

Santer, B., Taylor, K., Wigley, T., Penner, J., Jones, P., & Cubasch, U. (1995). Towards the detection and attribution of an anthropogenic effect on climate. *Climate Dynamics*, *12*, 77–100.

Schmidt, G. A., Jungclaus, J., Ammann, C., Bard, E., Braconnot, P., Crowley, T., et al. (2011). Climate forcing reconstructions for use in PMIP simulations of the last millennium (v1. 0). *Geoscientific Model Development*, *4*, 33–45.

Screen, J. A., Deser, C., Simmonds, I., & Tomas, R. (2014). Atmospheric impacts of Arctic Sea-ice loss, 1979–2009: Separating forced change from atmospheric internal variability. *Climate Dynamics*, *43*, 333–344.

Shuman, B. (2012). Patterns, processes, and impacts of abrupt climate change in a warm world: The past 11,700 years. *Wiley Interdisciplinary Reviews-Climate Change*, *3*, 19–43.

Slivinski, L. C., Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Giese, B. S., McColl, C., et al. (2019). Towards a more reliable historical reanalysis: Improvements for version 3 of the Twentieth Century Reanalysis system. *Quarterly Journal of the Royal Meteorological Society*, *145*(724), 2876–2908.

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.

Stevenson, S., Capotondi, A., Fasullo, J., & Otto-Bliesner, B. (2019). Forced changes to twentieth century ENSO diversity in a last millennium context. *Climate Dynamics*, *52*(12), 7359–7374.

Stott, P. A., Gillett, N. P., Hegerl, G. C., Karoly, D. J., Stone, D. A., Zhang, X., & Zwiers, F. (2010). Detection and attribution of climate change: A regional perspective. *Wiley Interdisciplinary Reviews: Climate Change*, *1*, 192–211.

Stouffer, R., Hegerl, G., & Tett, S. (2000). A comparison of surface air temperature variability in three 1000-yr coupled ocean-atmosphere model integrations. *Journal of Climate*, *13*, 513–537.

Stouffer, R., Manabe, S., & Vinnikov, K. (1994). Model assessment of the role of natural variability in recent global warming. *Nature*, *367*, 634–636.

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.

Thomson, D. (1982). Spectrum estimation and harmonic-analysis. *Proceedings of the IEEE*, *70*, 1055–1096.

Trenberth, K. E., & Fasullo, J. T. (2013). An apparent hiatus in global warming? *Earth Future*, *1*, 19–32.

Zhou, C., Zelinka, M. D., & Klein, S. A. (2017). Analyzing the dependence of global cloud feedback on the spatial pattern of sea surface temperature change with a Green's function approach. *Journal of Advances in Modeling Earth Systems*, *9*, 2174–2189. https://doi.org/10.1002/2017MS001096

Zhu, F., Emile-Geay, J., McKay, N. P., Hakim, G. J., Khider, D., Ault, T. R., et al. (2019). Climate models can correctly simulate the continuum of global-average temperature variability. *Proceedings of the National Academy of Sciences of the United States of America*, *116* (8728–8733), 18. https://doi.org/10.1073/pnas.1809959116