AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Atmospheres

RESEARCH ARTICLE
10.1002/2015JD024382

Key Points:
- Both CMIP ensembles show statistically significant warming at global and continental scales during the twentieth century
- The twentieth century temperature trend is smaller by one tenth of a degree Celsius in CMIP5 than in CMIP3
- Greater role of internal variability at decadal scales globally and for long-term trends regionally

Supporting Information:
- Supporting Information S1

Correspondence to:
S. Kumar,
sanjiv.kumar@noaa.gov

Citation:
Kumar, S., J. L. Kinter III, Z. Pan, and J. Sheffield (2016), Twentieth century temperature trends in CMIP3, CMIP5, and CESM-LE climate simulations: Spatial-temporal uncertainties, differences, and their potential sources, *J. Geophys. Res. Atmos.*, *121*, 9561–9575, doi:10.1002/2015JD024382.

Received 16 NOV 2015
Accepted 30 JUL 2016
Accepted article online 4 AUG 2016
Published online 31 AUG 2016

Twentieth century temperature trends in CMIP3, CMIP5, and CESM-LE climate simulations: Spatial-temporal uncertainties, differences, and their potential sources

Sanjiv Kumar[1], James L. Kinter III[2], Zaitao Pan[3], and Justin Sheffield[4,5]

[1]Physical Science Division, Earth System Research Laboratory, National Oceanic and Atmospheric Administration, Boulder, Colorado, USA, [2]Center for Ocean-Land-Atmosphere Studies, George Mason University, Fairfax, Virginia, USA, [3]Department of Earth and Atmospheric Sciences, Saint Louis University, St. Louis, Missouri, USA, [4]Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey, USA, [5]Geography and Environment, University of Southampton, Southampton, UK

Abstract The twentieth century climate simulations from the Coupled Model Intercomparison Project Phase 3 (CMIP3) and Phase 5 (CMIP5) are compared to assess the models' ability to capture observed near-surface air temperature trends at global, continental, and regional scales. We computed trends by using a nonparametric method and considering long-term persistence in the time series. The role of internal variability is examined by using large ensemble climate simulations from the National Center for Atmospheric Research model Community Earth System Model (CESM). We computed temperature trends for three periods: (1) the twentieth century, (2) the second half of the twentieth century, and (3) the recent hiatus period to contrast the roles of external forcing and internal variability at various spatial and temporal scales. Both CMIP ensembles show statistically significant warming at global and continental scales during the twentieth century. We found a small but statistically significant difference between CMIP3 (0.57 ± 0.07 °C/century) and CMIP5 (0.47 ± 0.06 °C/century) twentieth century temperature trends, with the CMIP3 estimate being closer to the observations. The spatial structure of long-term temperature trends, and top-of-the atmosphere net radiation trends, suggests that differences in model parameterizations and feedback processes that lead to a smaller net radiative forcing are likely contributing to the differences between CMIP3 and CMIP5. The estimate of internal variability based on the CESM large ensemble spans 24% of the uncertainty in CMIP5 for the twentieth century temperature trends, and 76% for the recent hiatus period, both at global scales, and 43% and almost 100% during the corresponding time periods at regional scales.

1. Introduction

The Coupled Model Intercomparison Project Phase 5 (CMIP5) provides an unprecedented set of data that can be used to study climate variability and change [*Taylor et al*., 2012]. The Intergovernmental Panel on Climate Change Fifth Assessment Report (IPCC AR5) was largely based on CMIP5 model results [*Intergovernmental Panel on Climate Change* (*IPCC*), 2013]. The CMIP5 data set builds on the previous CMIP3 project on which the IPCC AR4 is based [*IPCC*, 2007]. The CMIP5 models were generally run at higher spatial resolutions than those in CMIP3 and incorporated improved or new representations of climate/earth system processes, e.g., interactive carbon cycle, aerosol processes, and land use change [*Flato et al*., 2013]. Studies have compared the CMIP5 with CMIP3 focusing mostly on components over specific regions (e.g., *Stroeve et al*. [2012] for Artic ice extent, *Sheffield et al*. [2013] for North America, and *Sperber et al*. [2013] for Asian monsoon). Furthermore, there are reported mixed results suggesting no realized benefit [e.g., *Knutti and Sadlacek*, 2013] and wider spread in CMIP5 temperature changes than CMIP3 [*Jones et al*., 2013]. Similarly, *Knutson et al*. [2013a] have found a generally similar detectability of anthropogenic forcing on observed warming in both CMIP3 and CMIP5 simulations.

While a number of these studies have focused on the consistency between CMIP3 and CMIP5 simulations, it is also important to document differences between these two generation of climate models for the following reasons: (1) to find specific differences because of improvements and/or additional model processes, e.g., aerosol processes and land use change and (2) to document changes across generations of climate models. For example, *Knutti et al*. [2013] have found improvements in temperature and precipitation simulations in CMIP5 models compared to CMIP3 and CMIP2 models and (3) to provide a guidance or benchmark for future

improvements. For example, there is some evidence suggesting an overestimation of non-greenhouse gas (GHG) anthropogenic forcing response in CMIP5 [*Bindoff et al*., 2013; *Jones et al*., 2013], but the response has not been comprehensively analyzed.

Since the release of the IPCC AR4, several agencies involved in the management of natural resources, e.g., the U.S. Geological Survey and the Bureau of Reclamation, have prepared climate change impact assessment and adaptation plans by using CMIP3 data [e.g., *Brekke et al*., 2009; *Hay et al*., 2011; *Sale et al*., 2012]. The U.S. Global Change Research Program produced its third National Climate Assessment based on CMIP3 data [*Walsh et al*., 2014]. It is therefore useful to assess the differences between CMIP3 and CMIP5 so that the user community can be better informed about the uncertainties and robustness in the two generations of climate simulations.

In CMIP simulations, the role of external forcing is mixed with parameterization uncertainties in the models and the role of internal variability. Recently, available climate simulations from the Community Earth System Model Large Ensemble (CESM-LE) developed at the National Center for Atmospheric Research provide an opportunity to investigate the role of internal variability [*Kay et al*., 2015]. Here we assess the role of internal variability at various spatial and temporal scales ranging from global to local and century to decadal time scales. This analysis is limited by the representation of internal variability in only one climate model.

The first objective of this study is to compare the twentieth century temperature trends derived from CMIP3, CMIP5, and CESM-LE climate simulations and observations. Considering the observational uncertainties, we use three observational data sets: the Hadley Center and Climate Research Unit analysis version 4 (HadCRUT4 [*Morice et al*., 2012]), the National Oceanic and Atmospheric Administration Merged Land-Ocean Surface Temperature Analysis version 3.5.3 (NOAA-MLOST [*Smith et al*., 2008]), and the Goddard Institute for Space Studies Surface Temperature Analysis (GISSTEMP [*Hansen et al*., 2010]; see section S1 in the supporting information).

Over the last 15 years (1999 to 2013) there has been a perceived slowdown in global warming rate, also known as the "hiatus" period [*Meehl et al*., 2011]. While a number of studies have emphasized the role of internal variability, particularly the role of the central eastern Pacific and the transfer of heat to the deep ocean layers [e.g., *Meehl et al*., 2011, 2014a, *Kosaka and Xie*, 2013; *Trenberth and Fasullo*, 2013], other studies have also emphasized the role of external forcing [e.g., *Fyfe et al*., 2013; *Schmidt et al*., 2014] and data uncertainty [*Karl et al*., 2015]. *Karl et al*. [2015] suggested that the observations do not support the notion of a global warming hiatus. The second objective of this study is to contribute to the discussion on the role of internal variability versus external forcing at various spatial and temporal scales. Studying the role of internal variability is an active research area [e.g., *Deser et al*., 2012a; *Swart et al*., 2015; *Dai et al*., 2015], which is particularly important for detection and attribution studies at regional scales [e.g., *Wan et al*., 2014; *Kumar et al*., 2015; *Najafi et al*., 2015].

2. Data and Methodology

We analyzed 22 CMIP3 and 41 CMIP5 climate models including all available ensemble members from the all-forcings, historical climate simulations. A total of 66 historical climate simulations from CMIP3 and 138 from CMIP5 were analyzed (Table S1 in the supporting information). We extended these simulations until 2013 by using climate projections from business-as-usual scenarios: Special Report on Emission Scenarios-A1B in CMIP3 and 8.5 W m^{-2} Representative Concentration Pathway in CMIP5 (RCP8.5). Monthly outputs from climate simulations include near-surface air temperature (tas), top of the atmosphere (TOA) incoming shortwave radiation (rsdt), outgoing shortwave radiation (rsut), and outgoing longwave radiation (rlut). Model outputs were regridded to a common resolution ($2.5° \times 2.5°$) by using a method that preserves area averages. We considered the 1901–1998 period as the twentieth century and 1950–1998 as the second half of the twentieth century to ensure consistency between CMIP3 and CMIP5. The recent hiatus is analyzed for 1999 to 2013, which follows a strong El Niño event during 1997/1998 (http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ensoyears.shtml), thereby providing a positive temperature anomaly at the start of the trend period in the observation.

The role of internal variability is investigated by using CESM-LE climate simulations (30 members [*Kay et al*., 2015]). Different realizations (ensemble members) of the same climate model under the same forcing were produced by slightly perturbing the initial conditions. The CESM-LE historical climate simulations are available for the period of 1920 to 2005 and extended to the future by using the RCP8.5 scenario.

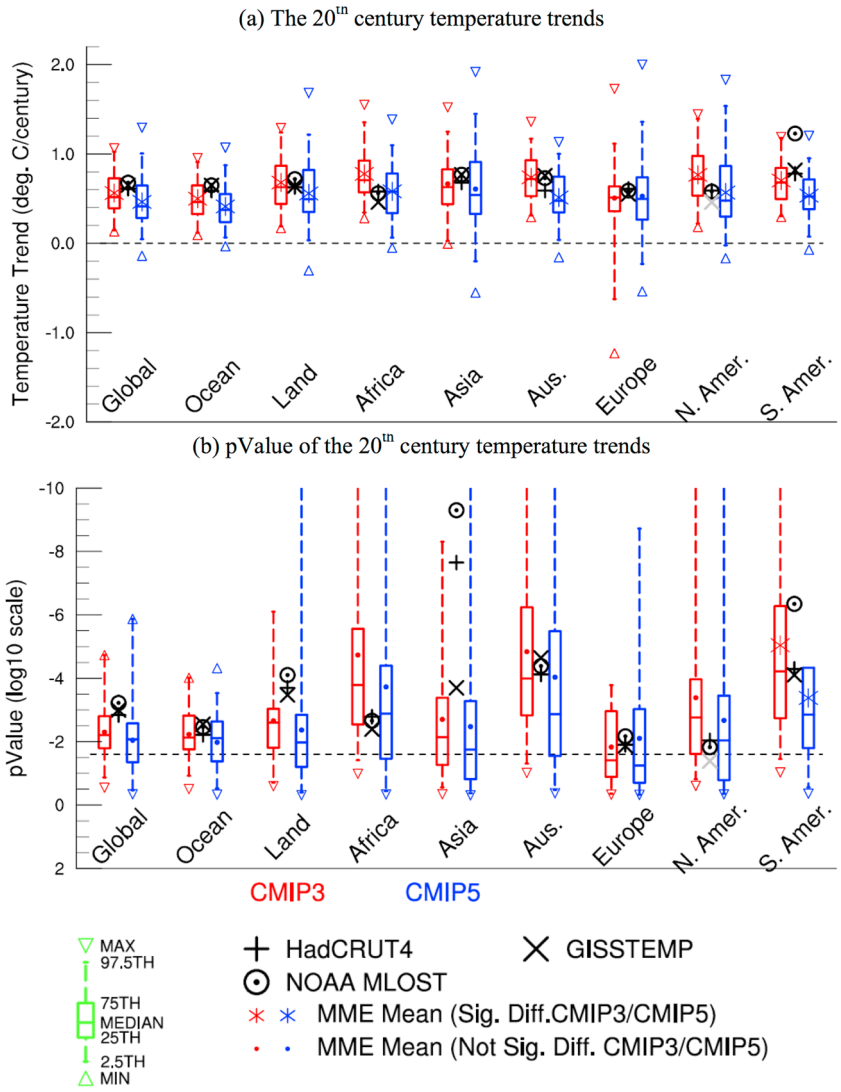

Figure 1. (a and b) The twentieth century temperature trends in CMIP3, CMIP5, and observations at global and continental scales. In Figure 1a observed temperature trends are shown in black if they are statistically significant (P value < 0.05), otherwise in gray. P values of trends significance are shown in Figure 1b, where the dashed line indicates the threshold value for rejection of the null hypothesis of no trend. Note the reversed y axis in Figure 1b; i.e., for a value above the dashed line, the null hypothesis is rejected at the 5% significance level.

We employed a nonparametric trend detection technique, which has the following advantages: no a priori assumption of a linear trend, distribution free, robust against outliers, and higher power for nonnormally distributed data [*Onoz and Bayazit*, 2003; *Yue et al*., 2002; *Kumar et al*., 2009, 2013a]. The magnitude of trends is determined by using the Theil-Sen approach (TSA) [*Sen*, 1968; *Thiel*, 1950]. If $x_1, x_2, \ldots, x_n$ is an annual time series (X_t) of length n then the TSA slope is given by

$$\beta = \text{median}\left[\frac{x_j - x_i}{j - i}\right] \text{for all } i < j. \quad (1)$$

The statistical significance of trends at less than 5% level is determined by using the Mann-Kendall test considering long-term persistence (LTP) in the time series as described in *Kumar et al*. [2009]. LTP represents low-frequency climate variability (decadal to multidecadal), which is a major source of uncertainty in trend analysis [*Koutsoyiannis and Montanari*, 2007; *DelSole et al*., 2011; *Kumar et al*., 2013b]. The presence of LTP leads to underestimation of variance and thereby a false identification of statistically significant trends by using traditional trend analysis method.

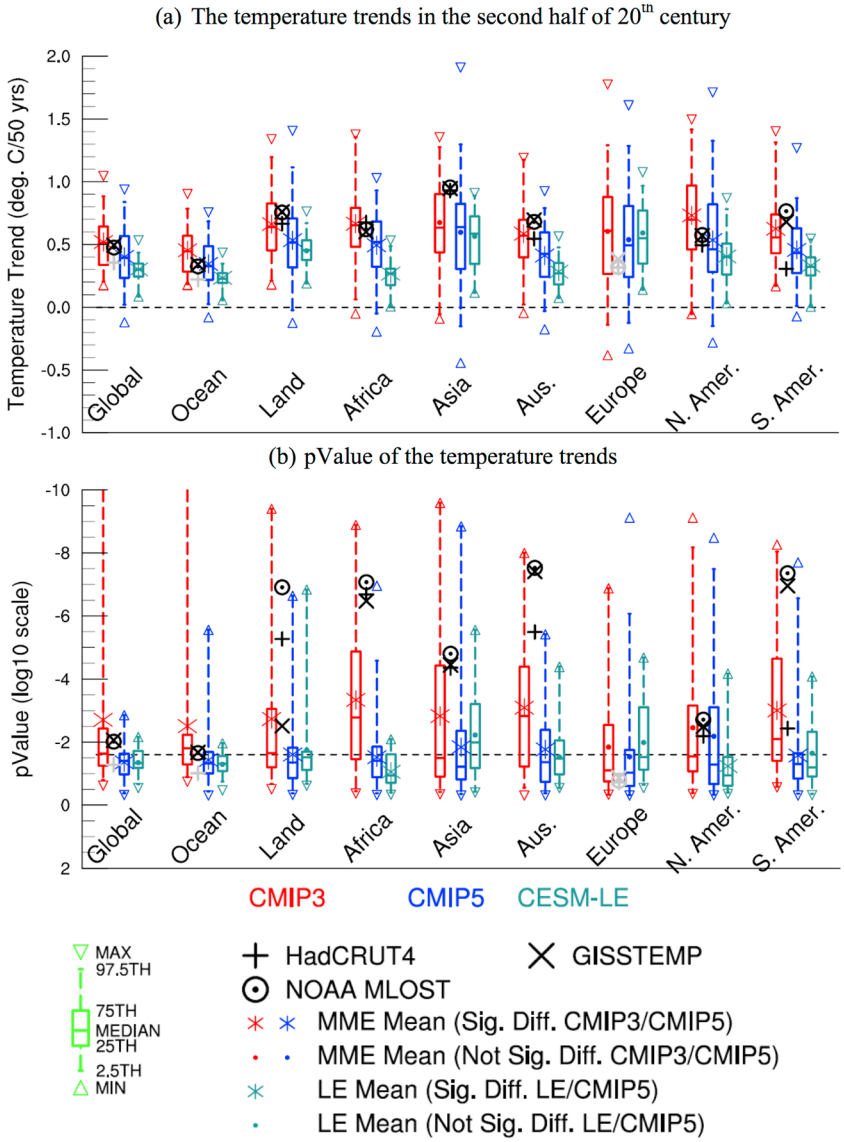

Figure 2. (a and b) Temperature trends in the second half of the twentieth century. In Figure 2a observed temperature trends are shown in black if they are statistically significant (P value < 0.05), otherwise in gray. P values of trends significance are shown in Figure 2b, where the dashed line indicates the threshold value for rejection of the null hypothesis of no trend. Note the reversed y axis in Figure 2b; i.e., for a value above the dashed line, the null hypothesis is rejected at the 5% significance level.

The model data were masked for the availability of all three observational data sets: 8 months or more data in a year and 90% or more data in the annual time series for the trend period. The global coverage of the observational record that meets the above criteria increases from 33% for the whole twentieth century, to 67% for the second half, and to 75% for the recent hiatus period (Figures S1–S3 in the supporting information). We did not change the number of grid points (spatial coverage) from one year to another; we only changed the spatial coverage from one trend period to another. Twenty-two land regions were defined based on *Giorgi* [2002] (Figure S4). We also define two oceanic regions, one in the North Atlantic and another in the North Pacific (Figure S4). We apply the observational masking when model results are compared with observations (e.g., Figures 1, 2, 3, 4, and 7); the observational masking is not applied for the remaining figures (Figures 5, 6, 8, 9, and 10). We compute the statistical significance of trends by using global or regional average annual temperature anomaly time series. Significance of local trends is discussed elsewhere [*Kumar et al.*, 2013a].

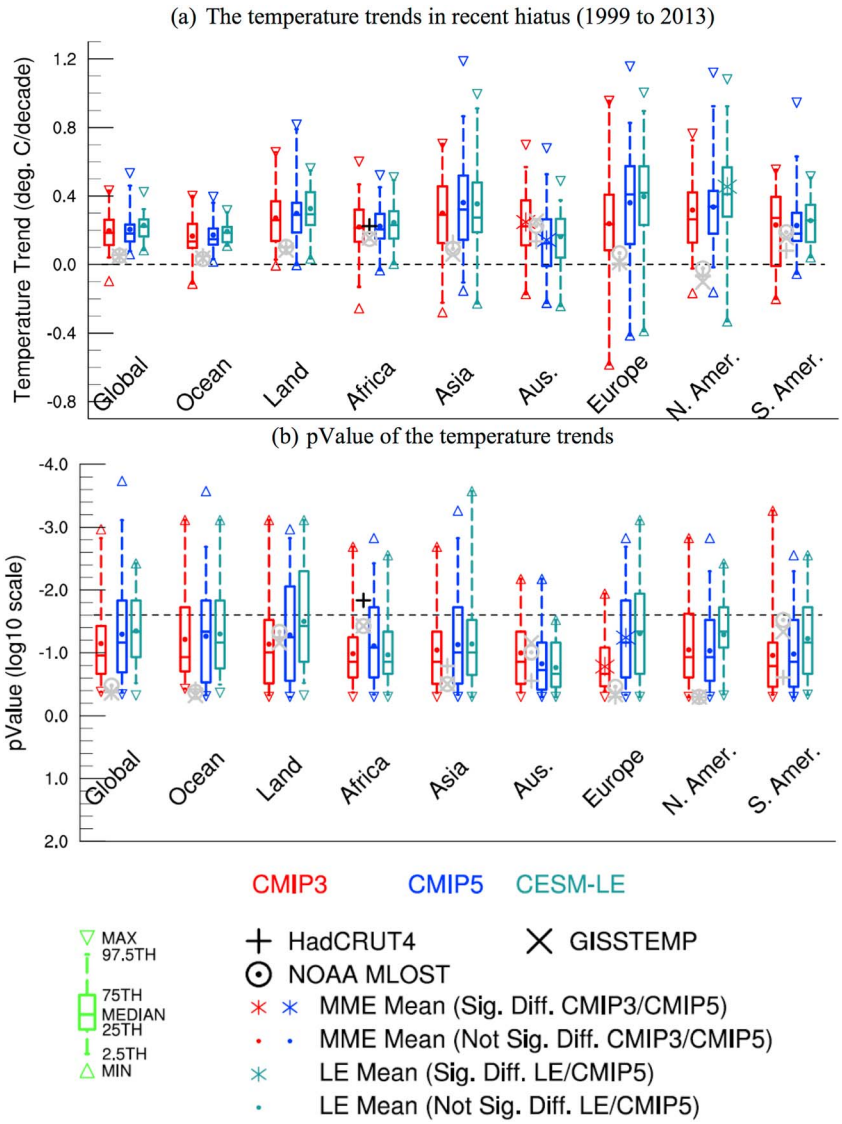

Figure 3. (a and b) Temperature trends in the recent hiatus period (1999 to 2013). In Figure 3a observed temperature trends are shown in black if they are statistically significant (P value < 0.05), otherwise in gray. P values of trends significance are shown in Figure 3b, where the dashed line indicates the threshold value for rejection of the null hypothesis of no trend. Note the reversed y axis in Figure 3b; i.e., for a value above the dashed line, the null hypothesis is rejected at the 5% significance level.

To avoid biases due to having more ensemble members from one climate model than another, we used a multimodel ensemble (MME) weighted average approach, thus ensuring a "one model one vote" policy [*Santer et al.*, 2007; *Jones et al.*, 2013]. The trend estimate from each climate simulation was given a weight (w):

$$w = \frac{1}{ME_m}, \tag{2}$$

where M (=22 for CMIP3 and 41 for CMIP5) is the total number of climate models in a given CMIP simulation and E_m is the total number of ensemble members for the given climate model. The statistical significance of the weighted MME mean difference (referred to as the MME mean, hereafter) is determined by using a Student's t test (see section S2). We also compared only those climate models that have generational representativeness in both CMIP3 and CMIP5 ensembles.

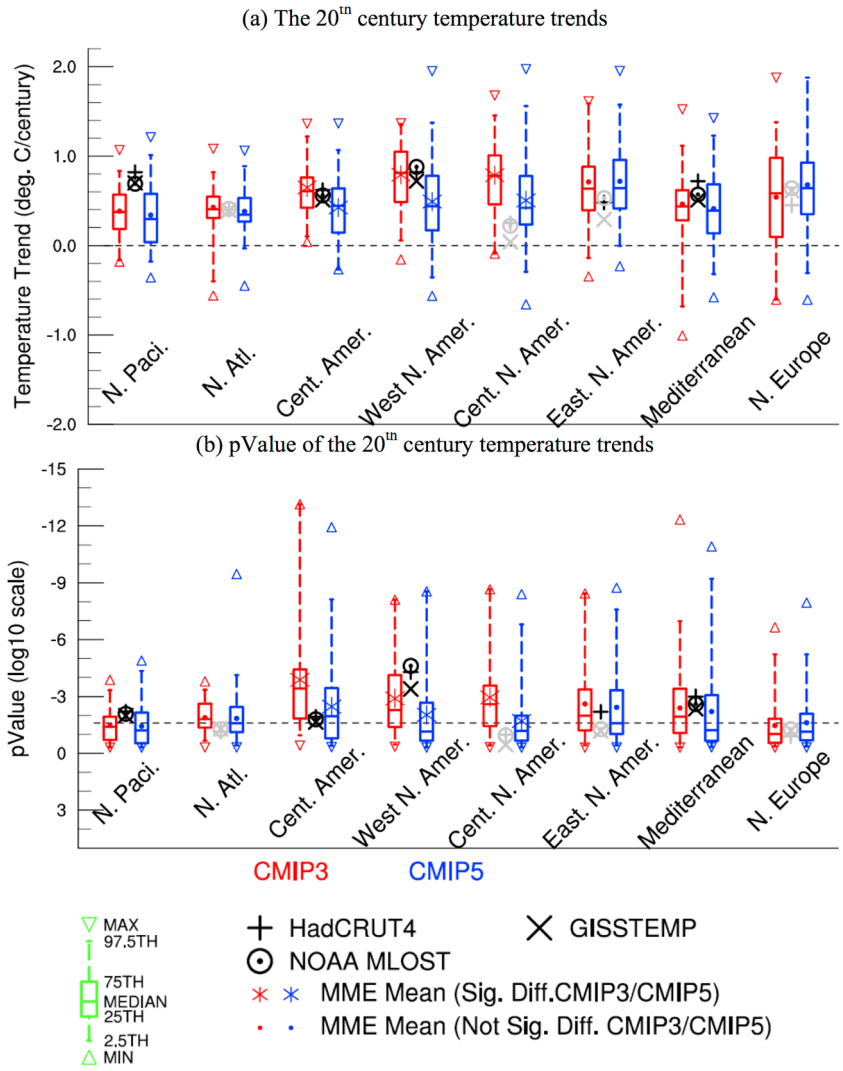

Figure 4. (a and b) The twentieth century temperature trends at regional scales. In Figure 4a observed temperature trends are shown in black if they are statistically significant (P value < 0.05), otherwise in gray. P values of trends significance are shown in Figure 4b, where the dashed line indicates the threshold value for rejection of the null hypothesis of no trend. Note the reversed y axis in Figure 4b i.e., for a value above the dashed line, the null hypothesis is rejected at the 5% significance level.

3. Results

3.1. Trend Analysis at Global, Continental, and Regional Scales

Figure 1 shows the twentieth century temperature trends in CMIP3, CMIP5, and observations at global and continental scales. The MME mean, median, interquartile ranges (25th to 75th percentiles), 95 percentile ranges (2.5th to 97.5th percentiles), and minimum and maximum values are shown by using a box plot. All three observations show statistically significant warming at global and continental scales except for North America where two of three observations show statistically significant warming, highlighting importance of the observational uncertainty. Both CMIP simulations generally capture the observed warming within their interquartile ranges; however, the multimodel mean trends are smaller than observed trends at global scales (Table 1), mostly contributed by the oceans (Figure 1a). The majority of climate simulations also show statistically significant warming (P value < 0.05; Figure 1b) except for Europe.

It is also evident in Figure 1a that the CMIP5 temperature trends are smaller compared to CMIP3. The difference between CMIP3 and CMIP5 twentieth century temperature trends are statistically significant (P value < 0.05) for global, as well as for ocean only, land only, and 4 out of 6 continents. The CMIP5 median warming rate

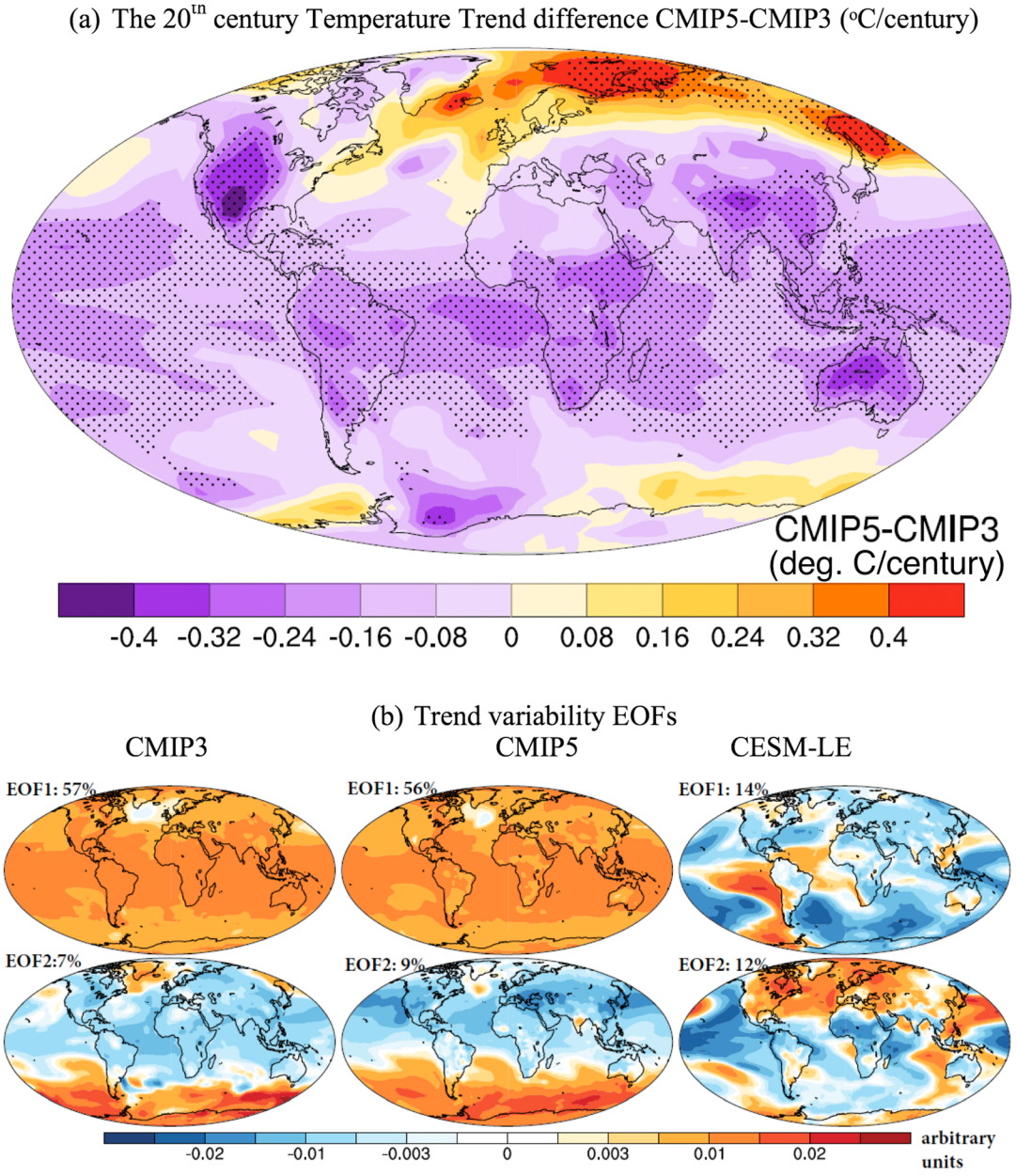

Figure 5. The twentieth century temperature trend mean difference between (a) CMIP5 and CMIP3 and (b) EOF analysis of temperature trend spread. Stippling in Figure 5a shows statistically significant difference.

and its interquartile range are also lower than CMIP3 except for Asia, and Europe. The global mean warming rates are 0.57 ± 0.07 and 0.47 ± 0.06 °C/century in CMIP3 and CMIP5, respectively; the error bars denote ±2 standard error estimate of the mean. Note that the difference between the CMIP3 and CMIP5 ensembles is almost one sixth to one fifth of the total simulated warming, and regionally, these differences are even larger (~0.3 °C/century; shown later); hence, it warrants further investigation. The three observational estimates are generally consistent with each other except for South America, which could be due to poor observational coverage. Averaged across the three observational estimates, the global mean warming rate for the twentieth century is 0.63 °C/century. Significant differences between the CMIP3 and CMIP5 temperature trends are also found when using full global data without observational masking, as well as subsampling of CMIP3 and CMIP5 climate models using available literature (Table 1).

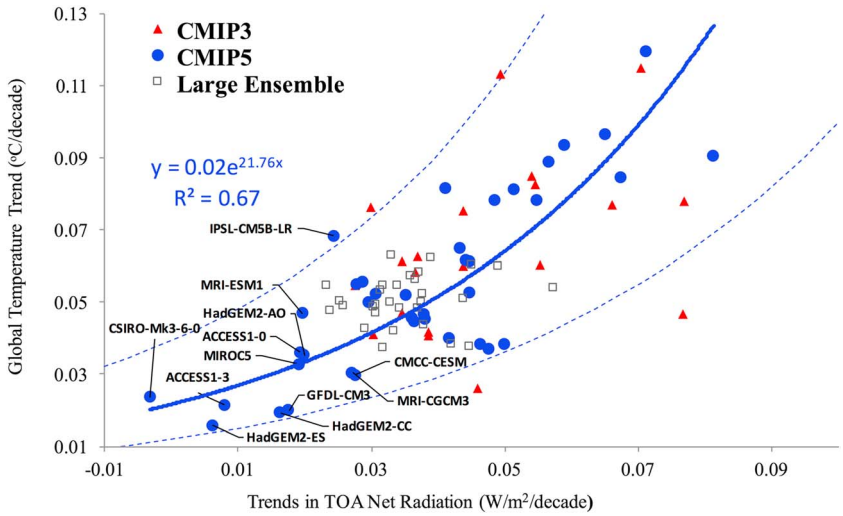

Figure 6. Scatterplot of trends in top of the atmosphere net radiation and global average temperature trends for the twentieth century (1901 to 1998). Data represent the ensemble mean trend in CMIP3 and CMIP5 models and individual member in CESM-LE. The best fit line across 40 CMIP5 models (solid blue line) and its 95% predication interval (blue dotted line) are also shown.

Figure 2 shows temperature trends for the second half of the twentieth century. The warming has accelerated in the second half of the century; the global mean temperature trends are 0.52 ± 0.06, 0.40 ± 0.05, and 0.43 °C per 50 year for CMIP3, CMIP5, and the observations, respectively. The HadCRUT4 global mean temperature trend (0.36) is smaller than the other two observational estimates (0.46 for NOAA MLOST and 0.48 for GISSTEMP), which is mainly due to a smaller temperature trend over the oceans in the former. The HadCRUT4 global mean temperature trend is also not statistically significant. All three observations do not show a statistically significant warming in Europe. A smaller temperature trend in CMIP5 than CMIP3 is also found in the second half of the twentieth century (Figure 2a). These differences also exist without applying observational masking to the CMIP data (Figure S5). While the majority of CMIP3 simulations show statistically significant warming, which is consistent with observations, the majority of CMIP5 simulations does not show significant warming (Figure 2b), highlighting the importance of the difference between CMIP3 and CMIP5. Internal climate variability can also play an important role in addition to external forcing at multidecadal time scales (shown later [also see *DelSole et al.*, 2011]).

Table 1. The Sensitivity of the Twentieth Century Global Temperature Trend (°C/Century) to Observational Masking and Model Selection[a]

	CMIP3	CMIP5	HadCRUT4	NOAA MLOST	GISSTEMP
This study with observational mask (22 CMIP3 and 41CMIP5) (33% global area): Note 1	0.57 ± 0.07	0.47 ± 0.06	0.60	0.66	0.64
Without Observational Masking (100% global area)					
This study global coverage: Note 1 (22 CMIP3 and 41CMIP5)	0.64 ± 0.06	0.54 ± 0.06		Not applicable	
Volcanic CMIP3 model selection using Jones et al. and equivalent CMIP5 (12 CMIP3 and 14 CMIP5; Note 2)	0.62 ± 0.06	0.43 ± 0.07		Not applicable	
Volcanic CMIP3 model selection by using Knutosn et al. and equivalent CMIP5 (7 CMIP3 and 10 CMIP5; Note 3)	0.56 ± 0.07	0.41 ± 0.09		Not applicable	

[a]The error bars denote ±2 times the standard error estimate of the mean. Note 1: Twenty-two CMIP3 models include bccr_bcm2_0, cccma_cgcm3_1_t63, cnrm_cm3, csiro_mk3_0, csiro_mk3_5, gfdl_cm2_0, gfdl_cm2_1, giss_aom, giss_model_e_h, giss_model_e_r, iap_fgoals1_0_g, ingv_echam4, inmcm3_0, ipsl_cm4, miroc3_2_hires, miroc3_2_medres, mpi_echam5, mri_cgcm2_3_2a, ncar_ccsm3_0, ncar_pcm1, ukmo_hadcm3, ukmo_hadgem1. Forty one CMIP5 models are: ACCESS1-0, bcc-csm1-1, CanESM2, CCSM4, CESM1-CAM5, CNRM-CM5, CSIRO-Mk3-6-0, GFDL-CM3, GFDL-ESM2G, GFDL-ESM2M, GISS-E2-H, GISS-E2-R, HadCM3, HadGEM2-CC, HadGEM2-ES, inmcm4, IPSL-CM5A-LR, IPSL-CM5A-MR, MIROC5, MIROC-ESM, MPI-ESM-LR, MRI-CGCM3, NorESM1-M, ACCESS1-3, CESM1-BGC, CESM1-FASTCHEM, CESM1-WACCM, CMCC-CESM, CMCC-CM, CMCC-CMS, CNRM-CM5-2, FGOALS-s2, GFDL-CM2p1, GISS-E2-H-CC, GISS-E2-R-CC, HadGEM2-AO, IPSL-CM5B-LR, MIROC-ESM-CHEM, MPI-ESM-P, MRI-ESM1, and NorESM1-ME. Models whose names are underlined were not included in the recent hiatus analysis shown in Figure 4 because data were not available. Note 2: Twelve CMIP3 models include gfdl_cm2_0, gfdl_cm2_1, giss_model_e_h, giss_model_e_r, inmcm3_0, miroc3_2_hires, miroc3_2_medres, mri_cgcm2_3_2a, ncar_ccsm3_0, ncar_pcm1, ukmo_hadcm3, and ukmo_hadgem1. Fourteen CMIP5 models are: CCSM4, CESM1-CAM5, GFDL-CM3, GFDL-ESM2G, GFDL-ESM2M, GISS-E2-H, GISS-E2-R, HadGEM2-CC, HadGEM2-ES, MIROC5, MIROC-ESM, inmcm4, HadCM3, and MRI-CGCM3. Note 3: Seven CMIP3 models include gfdl_cm2_0, gfdl_cm2_1, giss_model_e_h, giss_model_e_r, miroc3_2_medres, ncar_ccsm3_0, and ukmo_hadgem1. Ten CMIP5 models include CCSM4, GFDL-CM3, GFDL-ESM2G, GFDL-ESM2M, GISS-E2-H, GISS-E2-R, HadGEM2-CC, HadGEM2-ES, MIROC5, and MIROC-ESM.

Figure 2 also shows temperature trends from the CESM-LE climate simulations for the second half of the twentieth century (Figure 2). The CESM-LE mean trend is smaller than both the CMIP3 and CMIP5 ensemble mean trends at global scale, and in four continents. A smaller trend in CESM-LE can be partially related to the aerosol indirect effect that resulted in a reduction in twentieth century warming relative to an earlier version of the same model (CCSM4) that did not include aerosol indirect effect [*Hurrell et al*., 2013]. The CESM-LE temperature trend spread (standard deviation) is smaller by almost 60% at global scale and by 50% at continental scale compared to the CMIP5 simulations; the CMIP5 trend spread is generally comparable to the CMIP3 trend spread. The interquartile ranges of trends simulated by the CESM-LE generally do not capture the observed trend except for Europe. In some cases, e.g., in Africa and Australasia, the observational estimates lie outside the CESM-LE simulation range. Overall, the observations are better captured by the CMIP simulations compared to the CESM-LE simulations; i.e., observations are within the interquartile range of the CMIP simulations. This result emphasizes the importance of using a multimodel ensemble for studying long-term temperature trends. Nevertheless, the CESM-LE provides valuable data to assess the role of internal variability.

Figure 3 shows temperature trends during the recent hiatus period at global and continental scales. Both the CMIP3 and CMIP5 simulations have difficulty in capturing the recent observed hiatus, which lies at the lower end of the CMIP simulations (Figure 3). The MME mean temperature trends are not significantly different between the CMIP3 and CMIP5 simulations except for Australasia and between the CMIP5 and the CESM-LE simulations except for North America. The temperature trend spread from the CMIP simulations is comparable to the trend spread from the CESM-LE simulations, indicating the potential role of internal variability during the recent hiatus (shown later). This result supports the suitability of the CESM-LE simulations for studying decadal climate variability. We also show the statistical significance of trends in Figure 3 for the sake of completeness, although 15 years is not long enough to have a robust estimate of trend. It is interesting to note that all three observations as well as a majority of climate simulations do not show statistically significant trends during this period; one exception is HadCRUT4 for Africa.

Figure 4 shows the twentieth century temperature trends at regional scales for North America and Europe and surrounding oceanic regions. Out of eight regions, four regions do not have a statistically significant warming in the observations or in several of the climate simulations; these regions are the North Atlantic, central and eastern North America, and northern Europe. These regions are known to have considerable influence of decadal to multidecadal climate variability [*Kumar et al*., 2013b; *Meehl et al*., 2015]. Overall, Figure 4 suggests a greater role of internal variability at regional scale even for a century-scale temperature trend. This issue is further discussed in section 3.3 by comparing temperature trend uncertainty in CMIP5 simulations with the CESM-LE simulations at various spatial and temporal scales.

3.2. Potential Sources of Difference in the Twentieth Century Temperature Trend Between CMIP3 and CMIP5

Figure 5a shows the spatial pattern of the difference between CMIP5 and CMIP3 mean temperature trends for the twentieth century. For the CMIP5 simulations, the slower warming rates are spatially extensive covering both continents and oceans, and particularly in tropical and subtropical regions. The difference in warming rate is slightly greater over the continents than over the oceans (Figure 5a). There are localized areas where the difference is notable, e.g., parts of the northern China desert regions, northern India, and southwestern North America and parts of the western-central Australian deserts and eastern Africa. All these regions tend to be dry areas, suggesting that dust aerosols may contribute to the reduced warming [*Huang et al*., 2014]. The spatial pattern of the slower warming rates in CMIP5 resembles the natural and anthropogenic aerosol (or well-mixed greenhouse gas) response [*Xie et al*., 2013]. Cooling (or warming) responses are subdued in the extratropical North Atlantic and the southern oceans due to heat transport by the deep ocean layers and reorganization of ocean currents [*Xie et al*., 2013]. A notable exception of the slower rate is the Eurasian part of the Arctic, which may be related to faster sea ice melting in CMIP5 as noted by *Stroeve et al*. [2012].

Figure 5b shows the first two empirical orthogonal functions (EOFs) of the twentieth century temperature trend uncertainty in CMIP3 (left column), CMIP5 (middle column), and CESM-LE simulations (right column). EOFs were computed by using the correlation matrix for the twentieth century temperature trends in 66 CMIP3 ensemble members, and 138 CMIP5 ensemble members (both weighted using the square root of weights in equation (2)), and for 1920 to 2004 temperature trends in 30 CESM-LE simulations. The first EOF

Table 2. Global Average TOA Rn and Individual Component Trends From 1901 to 1998 (Unit: W/m^2/Century)[a]

	CMIP3 (20 models)	CMIP5 (40 models)
Rn	0.47 ± 0.05	0.37 ± 0.05
Downward shortwave	0.23 ± 0.07	0.15 ± 0.01
Upward (reflected) shortwave	0.29 ± 0.24	0.95 ± 0.23
Upward (outgoing) longwave	-0.53 ± 0.20	-1.20 ± 0.18

[a]The error bars denote ±2 times the standard error estimate of the mean.

explains the majority of the variance in CMIP3 (57%) and CMIP5 (56%) and shows a spatially contiguous pattern with an enhanced equatorial response. The spatial pattern of the temperature trend difference in Figure 5a resembles the first EOF pattern, which represents the well-mixed greenhouse gas or aerosol response [*Liu et al*., 2005; *Xie et al*., 2013]. The spatial pattern of trend variability EOFs in CESM-LE simulations differs considerably from the external forcing responses. The first EOF in the CESM-LE explains 14% of the variance and resembles the El Niño–Southern Oscillation/Pacific Decadal Oscillation pattern [*Trenberth and Fasullo*, 2013; *Newman et al*., 2016]. The second internal variability EOF that explains 12% of the variance resembles a La Niña-like negative phase of Interdecadal Pacific Oscillation [*Meehl et al*., 2011].

To further ascertain the role of radiative forcing we analyzed trends in top-of-atmosphere net radiation (abbreviated as TOA Rn hereafter). Changes in atmospheric composition due to greenhouse gas emissions result in a net increase in absorbed radiation that warms the atmosphere. The twentieth century trends in TOA Rn explain 67% of the variance ($r = 0.82$) in global average temperature trends across the 40 CMIP5 climate models (Figure 6; data for GFDL-CM2p1 not available). The multimodel mean trend in TOA Rn is significantly smaller in CMIP5 (0.37 ± 0.05 W m^{-2}/century) than CMIP3 (0.47 ± 0.05 W m^{-2}/century). An analysis of the individual components of TOA Rn shows that the outgoing/reflected shortwave radiation contributes most to the differences in TOA Rn trends (Table 2). An opposite sign of change in the outgoing longwave radiation (a negative trend) compared to the reflected shortwave radiation (a positive trend) can be due to reorganization of convective clouds as noted by other studies [e.g., *Kato*, 2009; *Brown et al*., 2014].

The TOA Rn trends from 20 CMIP3 models (data for csiro_mk3_0 and ingv_echam4 not available) and their global temperature trends are aligned with the upper 2/third of the CMIP5 range (Figure 6). Twelve CMIP5 models that show smaller TOA net radiation trends than the CMIP3 minimum trends are CSIRO-Mk3-6-0, HadGEM2-ES, ACCESS1-3, HadGEM2-CC, GFDL-CM3, MIROC5, ACCESS1-0, MRI-ESM1, HadGEM2-AO, IPSL-CM5B-LR, MRI-CGCM3, and CMCC-CESM (highlighted in Figure 6). These 12 models are referred to as G12, hereafter. These models also generally show a lower temperature trend with one notable exception of IPSL-CM5B-LR, which lies outside the CMIP5 95% uncertainty range (Figure 6). The G12 models contribute most to the differences between CMIP3 and CMIP5 twentieth century temperature trends. This difference is not statistically significant after removing the G12 models from CMIP5 ensembles (not shown). We note that the G12 includes climate models from all over the world as well as several major climate modeling centers; hence, such a biased CMIP5 model selection may not be justified. *Knutti et al*. [2013] found that some of the G12 models, e.g., ACESS1-0, ACESS1-3, and HadGEM2-CC, may have similar model physics. Nevertheless, we investigate the origin of this discrepancy between the G12 and other climate models.

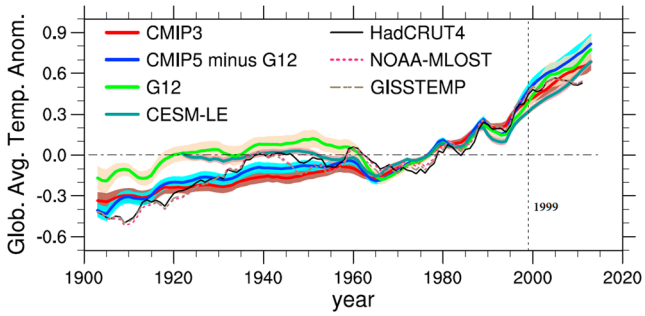

Figure 7. Time series of global average annual temperature anomaly. Five-year running mean is applied to smooth the data. Model data are masked for long-term observation availability as described in text. Observations represent average of three data HadCRUT4, NOAA-MLOST, and GISSTEMP. The shaded regions represent the 95% confidence interval estimate of mean. Observations are shown only up to 2013, whereas model outputs are up to 2015.

Figures 7 and 8 show the global average temperature and TOA Rn anomaly time series from various groups of climate simulations. We computed anomalies with respect to the 1961 to 1990 climatology in individual climate simulations and observations. There are several

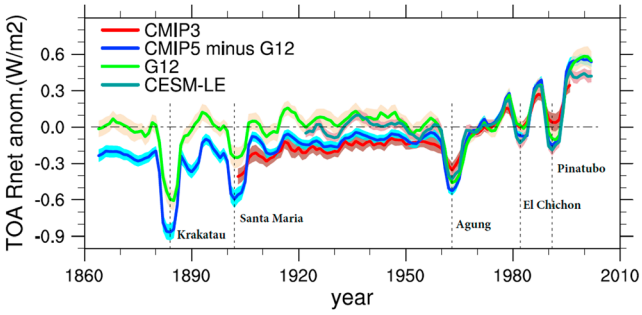

Figure 8. Global average top-of-the atmosphere (TOA) annual net radiation anomaly. Five-year running mean is applied to smooth the data. The shaded regions represent the 95% confidence interval estimate of mean. All major volcanic eruptions are also shown.

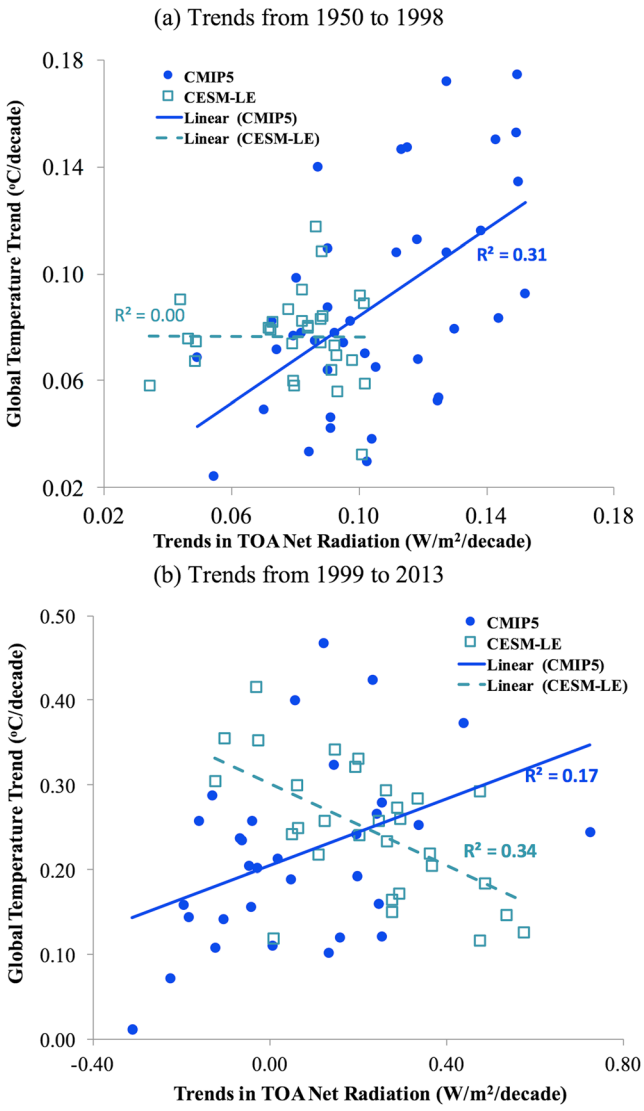

Figure 9. Relationship between top of the atmosphere net radiation and global average temperature trends for the shorter periods in CMIP5 and CESM-LE climate simulations.

notable features: (a) G12 models have a warm bias compared to the remaining CMIP5 models, CMIP3, and observations in the early half of the twentieth century (Figure 7). (b) The warm bias in the G12 climate models can be traced to a net positive (~0.2 W m^{-2}) TOA Rn anomaly at the start of the historical simulations compared to the remaining CMIP5 models (Figure 8), which could be due to a different background aerosol content. That displacement is reduced to 0.1 W m^{-2} after the Santa Maria volcanic eruption (1902) and reduced to 0.0 W m^{-2} after the Agung volcanic eruption (1963), suggesting that the background aerosol concentration difference among the models is time-dependent. (c) Another important difference is that the CMIP5 models have significantly greater reflectance leading to a greater negative TOA Rn anomaly than the CMIP3 models during all major volcanic activity (Figure 8). Significant differences between the G12 temperature trends and the remaining CMIP5 temperature trends also exist at regional scales (Figures S6–S9).

Some of the G12 models have generational representative in the CMIP3, e.g., csiro_mk3_5, gfdl_cm2_1, ipsl_cm4, miro-c3_2_hires, mri_cgcm2_3_2a, and ukmo_hadgem1. The twentieth century temperature and TOA Rn trends in the corresponding subset of the G12 climate models are 55 ± 16% and 65 ± 19% smaller, respectively, than their counterpart in the CMIP3 (Table S4). This result indicates substantial changes in the G12 models from CMIP3 to CMIP5 generations.

Time series derived from the CESM-LE simulations are also shown in Figures 7 and 8. The CESM-LE simulations are closer to the G12 models

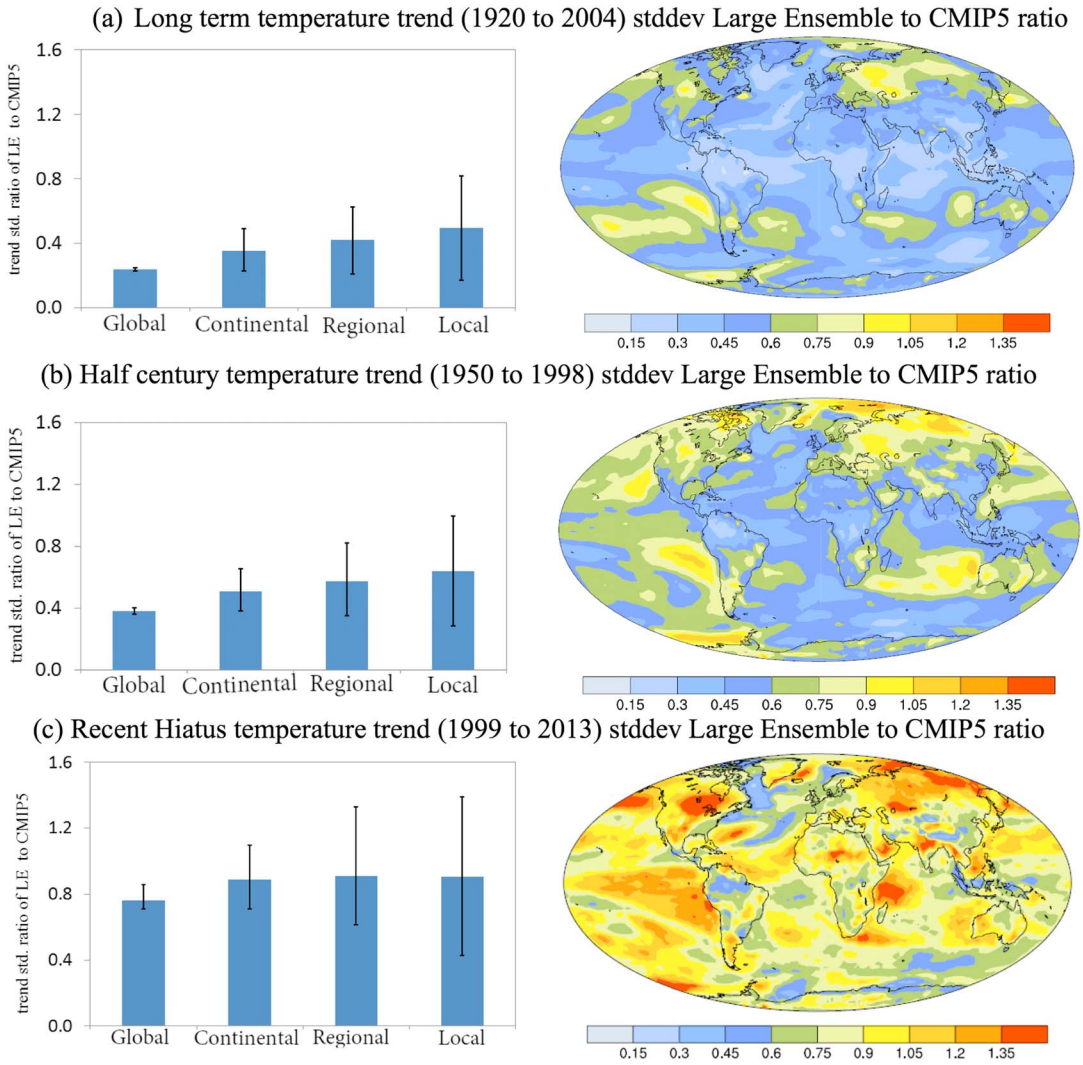

Figure 10. Role of internal variability estimated using CESM-LE at various spatial and temporal scales. It shows the ratio of temperature trend spread (standard deviation) from CESM-LE to CMIP5 simulations. Local trends are calculated for each 2.5° × 2.5° grid box over land only. The error bar denotes the 95% range from global average, land only, and ocean only (at global scales), six continents, 22 regions, 2.5° × 2.5° grid boxes over land.

than the remaining CMIP5 and CMIP3 models during the first half of the twentieth century (Figures 7 and 8). A warm temperature anomaly (~0.2 °C) in the observation, relative to the CMIP simulations, during the 1940's is better captured in the CESM-LE data than in the remaining CMIP5 and CMIP3 simulations (Figure 7).

The TOA Rn analysis is also helpful in quantifying the role of external forcing on the global average temperature trends at multidecadal to decadal time scales. Trends in TOA Rn explain only 31% of the variance ($r = 0.55$) in the 49 year temperature trend (second half of twentieth century) and only 17% of the variance ($r = 0.41$) in the 15 year temperature trend (recent hiatus) in the CMIP5 models (Figure 9). The temperature trend variability in CESM-LE does not have any correlations with TOA Rn trends for the second half of the twentieth century and has a negative correlation ($r = -0.58$) during the recent hiatus period (Figure 9), indicating that internal variability can affect the TOA radiative balance at decadal scales [e.g., *Brown et al*., 2014]. These results suggest an increasing role of internal variability in the coupled climate system at decadal time scales compared to century time scales (Figures 6 and 9).

3.3. Role of Internal Variability at Various Spatial and Temporal Scales

Figure 10 shows the role of internal variability estimated using CESM-LE for the three temporal scales: the twentieth century, second half of the twentieth century, and the recent hiatus period and at various spatial scales: global, continental, regional, and local. We compared the temperature trend uncertainty in the

CESM-LE with the trend uncertainty in the CMIP5 simulations. The uncertainty in the trends is measured as 1 standard deviation of trends in the respective simulations. This measure of uncertainty is the same as in previous studies, [e.g., *Knutti et al*., 2010; *Kay et al*., 2015]. Furthermore, the large sample sizes, e.g., $n = 30$ for CESM-LE and $n = 138$ for CMIP5, justify the use of standard deviation as a measure of uncertainty [*Knutti et al*., 2010]. For the long-term trend we used the 1920–2004 trend period and compared the CMIP5 trends uncertainty in the corresponding period. Here our assumption is that the CMIP5 trend uncertainty is dominated by the uncertainty due to model structure, parameterization differences, and feedback processes, whereas the CESM-LE uncertainty is only due to internal variability. This assumption is supported by the two analyses presented previously: (1) the EOF analysis in Figure 5b and (2) the TOA Rn analysis in Figure 6, where we found that TOA Rn trends are not significantly correlated with temperature trends in the CESM-LE simulations for 1920–2004 (squares in Figure 6) as opposed to CMIP5 where it is significantly correlated ($R^2 = 0.67$). We also found that the role of internal variability (intramodel) in the uncertainty of CMIP5 trends is minor relative to the inter-model variability. We confirmed this by repeating the analysis presented in Figure 10 by using only the first ensemble member from each CMIP5 model (total 41) and then comparing the trend uncertainty with that of CESM-LE (Figure S10, described below).

The role of internal variability estimated using CESM-LE increases as the time scale decreases from century to decadal (Figure 10). The ratio of the global average temperature trend spread of the CESM-LE to that of the CMIP5 ensemble increases from 0.24 for the 85 year trend period (1920–2004), to 0.38 for the 49 year trend period (1950–1998), and to 0.76 for the 15 year trend period (1999–2013). The role of internal variability also increases from global to local scales; e.g., for the long-term trends (1920–2004), the ratio of trend spread of the CESM-LE to the CMIP5 ensemble increases from 0.24 at global scale, to 0.36 at continental scale, to 0.42 at regional scale, and to 0.50 at local scale. There are regions with high internal variability, e.g., north-central North America and southwestern Russia and adjoining regions (Figure 10a, right column). During the recent hiatus, higher internal variability is also found in the central eastern Pacific, which has been associated with the recent hiatus [*Kosaka and Xie*, 2013; *Trenberth and Fasullo*, 2013]. Figure S10 shows the same result as in Figure 10 but using only the first ensemble member from each CMIP5 model (41 total) instead of using all CMIP5 ensemble members (total 138). The latter combines the effects of internal variability (intramodel) and model structural/parameterizations differences (intermodel). The result in Figure S10 is very similar to Figure 10; e.g., the ratio of the global average temperature trend spread of the CESM-LE to that of the CMIP5 ensemble increases from 0.24 for the 85 year trend period (1920–2004), to 0.40 for the 49 year trend period (1950–1998), and to 0.74 for the 15 year trend period (1999–2013). This confirms that intermodel variability is the major contributor to the uncertainty in CMIP temperature trends.

4. Discussion and Conclusions

A number of studies have investigated the twentieth century temperature trends in the CMIP3 and the CMIP5 ensembles for studying climate sensitivity, feedback parameter, attribution to anthropogenic influence [e.g., *Forster et al*., 2013; *Knutson et al*., 2013a], and the use of different trend calculation methodologies [e.g., *Jones et al*., 2013]. The main contribution of this study lies in evaluating the twentieth century temperature trends in a statistically rigorous way, e.g., considering long-term-persistence in trend significance calculation and making inferences about the role of natural variability at various spatial and temporal scales. We further examine the role of internal variability by analyzing the TOA Rn trends and the CESM-LE climate simulations.

We identified 12 CMIP5 models that fall outside the range of CMIP3 TOA Rn trends (lower side, G12). If these 12 models are excluded, then the difference between the twentieth century temperature trends in the CMIP3 and CMIP5 ensembles are not significant (not shown). Hence, we conclude that these G12 models contributed most to the differences between CMIP3 and CMIP5. We further traced these differences to a net positive TOA Rn anomaly (~0.2 W m^{-2}) in the G12 models compared to other models at the start of the historical runs, which could be due to several reasons, e.g., due to internal climate variability [e.g., *Brown et al*., 2014], a trend or drift in the preindustrial simulations [e.g.,*Knutson et al*., 2013a], or time dependent background aerosol forcing deference (Figure 8). It is also worth noting that *Forster et al*. [2013] found a higher nonGHG forcing (almost two times) in a subset of the G12 models compared to the remaining CMIP5 models.

The role of internal variability is investigated by using the CESM-LE simulations at various spatial and temporal scales. We found that at global scale, the contribution of internal variability to the uncertainty in

temperature trends increases from 24% for the twentieth century to 76% for the recent hiatus period and at regional scales from 43% to almost 100% during the corresponding time period. This analysis complements the statistical analysis results presented in section 3.1. For example, using a statistical method (trend significance), we found a greater role of internal variability at regional scale even for century time scales (Figure 4); this is also supported by the CESM-LE analysis in Figure 10. Our results add to and refine the existing estimates of the role of internal variability in decadal climate predications at global scales [e.g., *Meehl et al*., 2014a, 2014b] and regional climate predications on even longer time scales [*Hawkins and Sutton*, 2009; *Deser et al*., 2012a, 2012b]. For example, *Hawkins and Sutton* [2009] estimated the role of internal variability at local/regional scales for century time scale at less than 10% (Figure 6 in the reference); whereas we found this estimate to be 40 to 50% (Figure 10a). This result has implications for detection and attribution studies at regional scales where the role of internal variability is less recognized [e.g., *Najafi et al*., 2015]. It is likely that differences in methodology, e.g., *Hawkins and Sutton* [2009] estimated internal variability as a residual term of a polynomial fit in the respective climate simulation versus the CESM-LE estimate of internal variability used in the study, and climate projections versus historical climate simulations used here, may be contributing to the revised estimate. We also show that during the recent hiatus period (1999 to 2013) all three observational data sets and the majority of climate simulations do not show statistically significant trends (Figure 3). The main advantage of the CESM-LE is that we can identify process-level details of the role of internal variability, e.g., contribution from oceanic variability such as the Pacific Decadal Oscillation [*Newman et al*., 2016].

A number of studies have emphasized the role of internal variability for the recent hiatus [e.g., *Meehl et al*., 2011, 2014a; *Trenberth and Fasullo*, 2013]. We presented additional evidence by comparing the spatial pattern of radiative forcing response with the internal variability only response (Figure 5) and clarified the role of external forcing versus internal variability. While internal variability plays a major role at decadal time scales, e.g., the recent hiatus, the external forcing dominates longer time scale responses (cf. Figures 5 and 10). Further research is needed, for example, to understand the uncertainty in internal variability estimates considering its importance at decadal and regional scales.

Acknowledgments

The first author was supported by the Canadian Sea Ice and Snow Evolution Network (CanSISE) for a part of this work. The second author was supported by the National Science Foundation (1338427), NOAA (NA15OAR4310160), and the National Aeronautics and Space Administration (NNX14AM19G). The authors acknowledge the support of NOAA's Climate Program Office Modeling, Analysis, Predictions, and Projections (MAPP) Program as part of the CMIP5 Task Force. We also acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups (listed in Figure 1) for their model output. We also thank Matt Newman, Nathan Gillett, Fancis Zwiers, and four anonymous reviewers whose comments led to significant improvement in the manuscript. CMIP3 and CMIP5 data were downloaded from the following website: https://pcmdi9.llnl.gov/projects/esgf-llnl/. The CESM large ensemble data were downloaded from https://www.earthsystemgrid.org/dataset/ucar.cgd.ccsm4.CESM_CAM5_BGC_LE.html

References

Bindoff, N. L., et al. (2013), Detection and attribution of climate change: From global to regional, in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by T. F. Stocker et al., pp. 867–952, Cambridge Univ. Press, Cambridge, and New York.

Brekke, L. D., J. E. Kiang, J. R. Olsen, R. S. Pulwarty, D. A. Raff, D. P. Turnipseed, R. S. Webb, and K. D. White (2009), Climate change and water resources management—A federal perspective, *U.S. Geol. Surv. Circ. 1331*, 65.

Brown, P. T., W. Li, L. Li, and Y. Ming (2014), Top-of-atmosphere radiative contribution to unforced decadal global temperature variability in climate models, *Geophys. Res. Lett.*, *41*, 5175–5183, doi:10.1002/2014GL060625.

Dai, A. G., J. C. Fyfe, S. P. Xie, and X. G. Dai (2015), Decadal modulation of global surface temperature by internal climate variability, *Nat. Clim. Change*, *5*(6), 555–559.

DelSole, T., M. K. Tippett, and J. Shukla (2011), A significant component of unforced multidecadal variability in the recent acceleration of global warming, *J. Clim.*, *24*(3), 909–926.

Deser, C., R. Knutti, S. Solomon, and A. S. Phillips (2012a), Communication of the role of natural variability in future North American climate, *Nat. Clim. Change*, *2*, 775–779, doi:10.1038/nclimate1562.

Deser, C., A. S. Phillips, V. Bourdette, and H. Teng (2012b), Uncertainty in climate change projections: The role of internal variability, *Clim. Dynam.*, *38*, 527–546, doi:10.1007/s00382-010-0977-x.

Flato, G., et al. (2013), Evaluation of climate Models, in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by T. F. Stocker et al., pp. 741-866, Cambridge Univ. Press, Cambridge.

Forster, P. M., T. Andrews, P. Good, J. M. Gregory, L. S. Jackson, and M. Zelinka (2013), Evaluating adjusted forcing and model spread for historical and future scenarios in the CMIP5 generation of climate models, *J. Geophys. Res. Atmos.*, *118*, 1139–1150, doi:10.1002/jgrd.50174.

Fyfe, J. C., N. P. Gillett, and F. W. Zwiers (2013), Overestimated global warming over the past 20 years, *Nat. Clim. Change*, *3*, 767–769.

Giorgi, F. (2002), Variability and trends of sub-continental scale surface climate in the twentieth century. Part I: Observations, *Clim. Dyn.*, *18*, 675–691.

Hansen, J., R. Ruedy, M. Sato, and K. Lo (2010), Global surface temperature change, *Rev. Geophys.*, *48*, RG4004, doi:10.1029/2010RG000345.

Hawkins, E., and R. Sutton (2009), The potential to narrow uncertainty in regional climate predictions, *Bull. Am. Meteorol. Soc.*, *90*, 1095–1107.

Hay, L. E., S. L. Markstrom, and C. Ward-Garrison (2011), Watershed-scale response to climate change through the twenty-first century for selected basins across the United States, *Earth Interact.*, *15*, 1–37, doi:10.1175/2010EI370.1.

Huang, J. P., T. H. Wang, W. C. Wang, Z. Q. Li, and H. R. Yan (2014), Climate effects of dust aerosols over East Asian arid and semiarid regions, *J. Geophys. Res. Atmos*, *119*, 11,398–11,416, doi:10.1002/2014JD021796.

Hurrell, J. W., et al. (2013), The community Earth system model a framework for collaborative research, *Bull. Am. Meteorol. Soc.*, *94*(9), 1339–1360.

Intergovernmental Panel on Climate Change (IPCC) (2007), Summary for policymakers, in *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by S. Solomon et al., pp. 1–18, Cambridge Univ. Press, Cambridge.

Intergovernmental Panel on Climate Change (IPCC) (2013), Summary for policymakers, in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by T. F. Stocker et al., pp. 1–27, Cambridge Univ. Press, Cambridge.

Jones, G. S., P. A. Stott, and N. Christidis (2013), Attribution of observed historical near-surface temperature variations to anthropogenic and natural causes using CMIP5 simulations, *J. Geophys. Res. Atmos.*, *118*, 4001–4024, doi:10.1002/jgrd.50239.

Karl, T. R., A. Arguez, B. Y. Huang, J. H. Lawrimore, J. R. McMahon, M. J. Menne, T. C. Peterson, R. S. Vose, and H. M. Zhang (2015), Possible artifacts of data biases in the recent global surface warming hiatus, *Science*, *348*(6242), 1469–1472.

Kato, S. (2009), Interannual variability of the global radiation budget, *J. Clim.*, *22*(18), 4893–4907, doi:10.1175/2009JCLI2795.1.

Kay, J. E., et al. (2015), The Community Earth System Model (CESM) CESM-LE Project: A community resource for studying climate change in the presence of internal climate variability, *Bull. Am. Meteorol. Soc.*, doi:10.1175/BAMS-D-13-00255.1, in press.

Knutson, T. R., F. R. Zeng, and A. T. Wittenberg (2013a), Multimodel assessment of regional surface temperature Trends: CMIP3 and CMIP5 twentieth-century simulations, *J. Clim.*, *26*(22), 8709–8743.

Knutti, R., and J. Sadlacek (2013), Robustness and uncertainties in the new CMIP5 climate model projections, *Nat. Clim. Change*, *3*, 363–373, doi:10.1038/nclimate1716.

Knutti, R., R. Furrer, C. Tebaldi, J. Cermak, and G. A. Meehl (2010), Challenges in combining projections from multiple climate models, *J. Clim.*, *23*, 2739–2758.

Knutti, R., D. Masson, and A. Gettelman (2013), Climate model genealogy: Generation CMIP5 and how we got there, *Geophys. Res. Lett.*, *40*, 1194–1199, doi:10.1002/grl.50256.

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

Koutsoyiannis, D., and A. Montanari (2007), Statistical analysis of hydroclimatic time series: Uncertainty and insights, *Water Resour. Res.*, *43*, W05429, doi:10.1029/2006WR005592.

Kumar, S., V. Merwade, J. Kam, and K. Thurner (2009), Streamflow trends in Indiana: Effects of long term persistence, precipitation and subsurface drains, *J. Hydrol.*, *374*(1-2), 171–183.

Kumar, S., V. Merwade, J. L. Kinter III, and D. Niyogi (2013a), Evaluation of temperature and precipitation trends and long-term persistence in CMIP5 twentieth-century climate simulations, *J. Clim.*, *26*, 4168–4185.

Kumar, S., J. Kinter, P. A. Dirmeyer, Z. Pan, and J. Adams (2013b), Multidecadal climate variability and the "warming hole" in North America: Results from CMIP5 twentieth- and twenty-first-century climate simulations*, *J. Clim.*, *26*(11), 3511–3527.

Kumar, S., R. P. Allan, F. Zwiers, D. M. Lawrence, and P. A. Dirmeyer (2015), Revisiting trends in wetness and dryness in the presence of internal climate variability and water limitations over land, *Geophys. Res. Lett.*, *42*, 10867–10875, doi:10.1002/2015GL066858.

Liu, Z., S. Varus, F. He, N. Wen, and Y. Zhong (2005), Rethinking tropical ocean response to global warming: The enhanced equatorial warming, *J. Clim.*, *18*, 4684–4700.

Meehl, G. A., J. A. Arblaster, J. T. Fasullo, A. Hu, and K. Trenberth (2011), Model-based evidence of deep-ocean heat uptake during surface-temperature hiatus period, *Nat. Clim. Change*, *1*, 360–364, doi:10.1038/nclimate1229.

Meehl, G. A., H. Teng, and J. M. Arblaster (2014a), Climate model simulations of the observed early-2000s hiatus of global warming, *Nat. Clim. Change*, *4*, 898–902, doi:10.1038/nclimate2357.

Meehl, G. A., et al. (2014b), Decadal climate prediction – An update from the trenches, *Bull. Am. Meteorol. Soc.*, *95*, 243–267.

Meehl, G. A., J. M. Arblaster, and C. T. Y. Chung (2015), Disappearance of the southeast U.S. "warming hole" with the late 1990s transition of the Interdecadal Pacific Oscillation, *Geophys. Res. .Lett.*, *42*, 5564–5570, doi:10.1002/2015GL064586.

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

Najafi, M. R., F. W. Zwiers, and N. P. Gillett (2015), Attribution of Arctic temperature change to greenhouse gas and aerosol inuences, *Nat. Clim. Change*, doi:10.1038/NCLIMATE2524, in press.

Newman, M., et al. (2016), The Pacific Decadal Oscillation, revisited, *J. Clim.*, doi:10.1175/JCLI-D-15-0508.1, in press.

Onoz, B., and M. Bayazit (2003), The power of statistical tests for trend detection, *Turk. J. Eng. Environ. Sci.*, *27*(4), 247–251.

Sale, M. J., S.-C. Kao, M. Ashfaq, D. P. Kaiser, R. Martinez, C. Webb, and Y. Wei (2012), Assessment of the effects of climate change on federal hydropower, technical manual 2011/251, Oak Ridge Nat. Lab., Oak Ridge, Tenn.

Santer, B. D., et al. (2007), Identification of human-induced changes in atmospheric moisture content, *Proc. Natl. Acad. Sci. U.S.A.*, *104*(39), 15, 248–15, 253.

Schmidt, G. A., D. T. Schindell, and K. Tsigaridis (2014), Reconciling warming trends, *Nat. Geosci.*, *7*, 158–160.

Sen, P. K. (1968), Estimates of the regression coefficients based on Kendall's tau, *J. Am. Stat. Assoc.*, *63*, 1379–1389.

Sheffield, J., et al. (2013), North American climate in CMIP5 experiments, Part II: Evaluation of historical simulations of intraseasonal to decadal variability, *J. Clim.*, *26*, 9247–9290.

Smith, T. M., R. W. Reynolds, T. C. Peterson, and J. Lawrimore (2008), Improvements to NOAA's historical merged land-ocean surface temperature analysis (1880–2006), *J. Clim.*, *21*, 2283–2296.

Sperber, K. R., H. Annamalai, I.-S. Kang, A. Kitoh, A. Moise, A. Turner, B. Wang, and T. Zhou (2013), The Asian summer monsoon: an inter-comparison of CMIP5 vs. CMIP3 simulations of the late 20th century, *Clim. Dyn.*, *41*(9–10), 2711–2744, doi:10.1007/s00382-012-1607-6.

Stroeve, J. C., V. Kattsov, A. Barrett, M. Serreze, T. Pavlova, M. Holland, and W. N. Meier (2012), Trends in Arctic sea ice extent from CMIP5, CMIP3 and observations, *Geophys. Res. Lett.*, *39*, L16502, doi:10.1029/2012GL052676

Swart, N. C., J. C. Fyfe, E. Hawkins, J. E. Kay, and A. Jahn (2015), COMMENTARY: Influence of internal variability on Arctic sea-ice trends, *Nat. Clim. Change*, *5*(2), 86–89.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl (2012), An overview of CMIP5 and the experiment design, *Bull. Am. Meteorol. Soc.*, *93*, 485–498.

Thiel, H. (1950), A rank-invariant method of linear and polynomial analysis, Part 3, *Ned. Akad. Wetensch Proc.*, *53*, 1397–1412.

Trenberth, K. E., and J. T. Fasullo (2013), An apparent hiatus in global warming? *Earth's Future*, *1*, 19–32, doi:10.1002/2013EF000165.

Walsh, J., et al. (2014), *Ch. 2: Our Changing Climate. Climate Change Impacts in the United States: The Third National Climate Assessment, U.S. Global Change Res. Program*, edited by J. M. Melillo, T. Richmond, and G. W. Yohe, doi:10.7930/JOKW5CXT.

Wan, H., X. Zhang, F. W. Zwiers, and S.-K. Min (2014), Attributing Northern high-latitude precipitation change over the period 1966–2005 to human influence, *Clim. Dyn.*, doi:10.1007/s00382-014-2423-y.

Xie, S.-P., B. Lu, and B. Xiang (2013), Similar spatial patterns of climate responses to aerosol and greenhouse gas changes, *Nat. Geosci.*, *6*, 828–832, doi:10.1038/ngeo1931.

Yue, S., P. Pilon, B. Phinney, and G. Cavadias (2002), The influence of autocorrelation on the ability to detect trend in hydrological series, *Hydrol. Processes*, *16*(9), 1807–1829.