Evaluation of Temperature and Precipitation Trends and Long-Term Persistence in CMIP5 Twentieth-Century Climate Simulations

SANJIV KUMAR

Center for Ocean–Land–Atmosphere Studies, Calverton, Maryland

VENKATESH MERWADE

School of Civil Engineering, Purdue University, West Lafayette, Indiana

JAMES L. KINTER III

Center for Ocean–Land–Atmosphere Studies, Calverton, Maryland, and Department of Atmospheric, Oceanic and Earth Sciences, George Mason University, Fairfax, Virginia

DEV NIYOGI

Department of Agronomy, and Department of Earth, Atmospheric and Planetary Sciences, Purdue University, West Lafayette, Indiana

(Manuscript received 8 May 2012, in final form 26 December 2012)

ABSTRACT

The authors have analyzed twentieth-century temperature and precipitation trends and long-term persistence from 19 climate models participating in phase 5 of the Coupled Model Intercomparison Project (CMIP5). This study is focused on continental areas (60°S–60°N) during 1930–2004 to ensure higher reliability in the observations. A nonparametric trend detection method is employed, and long-term persistence is quantified using the Hurst coefficient, taken from the hydrology literature. The authors found that the multimodel ensemble–mean global land–average temperature trend (0.07°C decade^{-1}) captures the corresponding observed trend well (0.08°C decade^{-1}). Globally, precipitation trends are distributed (spatially) at about zero in both the models and in the observations. There are large uncertainties in the simulation of regional-/local-scale temperature and precipitation trends. The models' relative performances are different for temperature and precipitation trends. The models capture the long-term persistence in temperature reasonably well. The areal coverage of observed long-term persistence in precipitation is 60% less (32% of land area) than that of temperature (78%). The models have limited capability to capture the long-term persistence in precipitation. Most climate models underestimate the spatial variability in temperature trends. The multimodel ensemble–average trend generally provides a conservative estimate of local/regional trends. The results of this study are generally not biased by the choice of observation datasets used, including Climatic Research Unit Time Series 3.1; temperature data from Hadley Centre/Climatic Research Unit, version 4; and precipitation data from Global Historical Climatology Network, version 2.

1. Introduction

A major emerging theme of the World Climate Research Programme (WCRP) Open Science Conference, 2011 was "the need for actionable science. . . [that] requires scientists and projects to move beyond understanding and predicting the Earth's climate to providing the resulting knowledge and information that yield practical solutions" (Asrar et al. 2012). For example, precipitation trends provide valuable information for water or agricultural infrastructure planning and design (Arnell 2011). The Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report (AR5) is currently underway, and a report is anticipated in 2013/14. The IPCC AR5 is based on, among other things, climate model simulations

Corresponding author address: Sanjiv Kumar, Center for Ocean–Land–Atmosphere Studies, 4041 Powder Mill Rd., Suite 302, Calverton, MD 20705.
E-mail: sanjiv@cola.iges.org

DOI: 10.1175/JCLI-D-12-00259.1

© 2013 American Meteorological Society

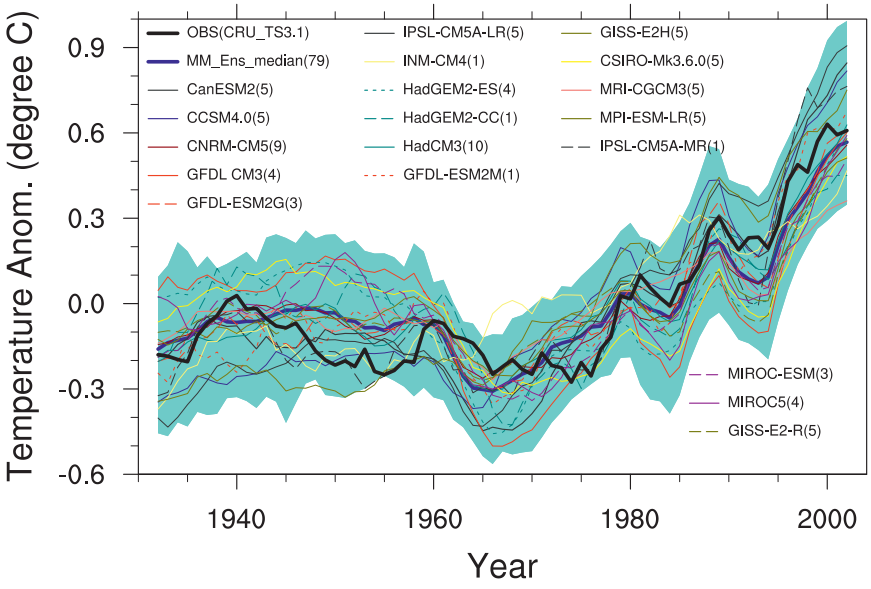

FIG. 1. Global land–average (60°S–60°N) 2-m air temperature anomaly from 19 CMIP5 climate models and from observations (OBS, CRU TS3.1). For models having more than one ensemble member, ensemble averages are shown. Shaded region shows the 95% uncertainty range calculated from 79 ensemble members. Number of ensemble members is shown in parentheses for each model. Five-year moving average is applied to smooth the anomaly curve. Ensemble average is calculated after determining the anomaly for each member.

provided by the world's major climate-modeling centers as part of phase 5 of the Coupled Model Intercomparison Project (CMIP5) (Taylor et al. 2012). This study aims to stretch the CMIP5 climate model evaluation beyond the physical climate features to its practical applications, particularly for the design of water and agricultural infrastructure.

Twentieth-century global warming trends in the "historical" simulations of many CMIP5 models look similar (Fig. 1). It is imperative to establish a robust quantitative measure, in contrast to visual inspection of the temperature time series anomaly plot (e.g., Fig. 1), to differentiate among different CMIP5 climate models. To answer any research or practical question, it is helpful to know which different models and their realizations (ensemble members) should be included in the analyses. Many modeling centers have provided outputs from more than one model, including atmosphere–ocean general circulation models (AOGCMs), earth system models (ESMs), and different versions and resolutions of the same model, following a common set of guidelines (Taylor et al. 2012). The first objective of this study is to propose a robust quantitative measure to evaluate models' relative performance and study inter- and intramodel spread among different CMIP5 climate models using the proposed metric. The proposed metric quantifies climate models' ability to capture twentieth-century observed temperature and precipitation trends.

The global-average temperature trend is an important indicator of global climate change. While considerable attention has been paid to temperature trends, precipitation trends have received less attention in public discussions of climate model results. Precipitation drives the hydrologic cycle, including regional flood and drought conditions. Generally, temperature time series conform to a Gaussian distribution, while precipitation time series do not. Hence, a nonparametric method such as the Mann–Kendall trend test (Kendall 1975; Mann 1945) is better suited to compare temperature and precipitation trends on an equal footing. The second objective of this study is to compare precipitation and temperature trends in CMIP5 climate models.

The hydroclimatic observations show long-term persistence (LTP), which is a major source of uncertainty in a trend analysis (Hurst 1951; Koutsoyiannis and Montanari 2007). The LTP represents the occurrence of similar climate conditions such as wet or dry spells in a cluster time span, for example, multidecadal persistent drought in West Africa (Shanahan et al. 2009). Persistence in a time series leads to an underestimation of variance and, subsequently, overestimation of the statistical significance of the trend (von Storch 1995; Hamed and Rao 1998; Koutsoyiannis 2003). Therefore, investigating climate models' ability to capture the long-term persistence behavior, as found in the observations, is specifically emphasized in this study.

TABLE 1. Temperature trend (°C decade^{-1}) and persistence analysis results (1930–2004) for 60°S–60°N (land only). Here, σ is std dev, sig. area is the area fraction (60°S–60°N, land only) where HC is statistically significant at the 90% confidence interval, M_1 is mean, M_2 is median, MM is multimodel, OBS is observation, r_1 is intramodel, r_2 is intra-institute, r_3 is intermodel, r_4 is model observation, and trend sig. area is the area fraction (60°S–60°N, land only) where the trend is statistically significant at the 95% confidence interval. For models having more than one ensemble member, ensemble averages are shown. For pattern correlation, values represent the mean of the 500 iterations (see text). Statistically significant pattern correlations (95% confidence interval) are shown in boldface.

			Trend		Trend pattern correlation				HC			HC pattern correlation		Trend sig. area	
Model number	Model name	Ensemble members	Global mean	Spatial σ	r_1	r_2	r_3	r_4	M_1	M_2	Sig. area	r_3	r_4	STP	LTP
1	CanESM2	5	0.10	0.05	**0.48**	—	0.07	−0.11	0.64	0.64	0.75	**0.21**	**0.23**	0.71	0.33
2	CCSM4.0	5	0.13	0.05	**0.66**	—	0.13	**0.35**	0.60	0.60	0.61	**0.22**	**0.34**	0.94	0.70
3	CNRM-CM5	9	0.06	0.04	**0.30**	—	0.16	0.00	0.58	0.57	0.50	0.11	0.09	0.58	0.34
4	GFDL CM3	4	0.01	0.05	0.18	−0.04	0.07	0.00	0.64	0.64	0.72	**0.24**	**0.30**	0.15	0.05
5	GFDL-ESM2G	3	0.08	0.06	**0.33**	0.13	0.07	−0.08	0.60	0.59	0.56	**0.23**	**0.24**	0.59	0.32
6	GFDL-ESM2M	1	0.09	0.04	—	0.12	0.14	0.26	0.58	0.57	0.46	0.14	0.08	0.56	0.30
7	HadCM3	10	0.08	0.04	0.20	0.16	0.12	0.04	0.60	0.59	0.58	**0.27**	**0.34**	0.58	0.27
8	HadGEM2-CC	1	0.04	0.06	—	**0.34**	0.20	−0.11	0.65	0.65	0.73	**0.32**	**0.46**	0.41	0.07
9	HadGEM2-ES	4	0.02	0.05	**0.44**	**0.35**	0.13	−0.06	0.68	0.67	0.84	**0.30**	**0.37**	0.26	0.04
10	INM-CM4	1	0.09	0.04	—	—	0.06	−0.14	0.54	0.50	0.26	**0.21**	**0.35**	0.82	0.75
11	IPSL-CM5A-LR	5	0.14	0.05	**0.33**	0.19	0.11	0.21	0.66	0.65	0.79	**0.33**	**0.38**	0.96	0.54
12	IPSL-CM5A-MR	1	0.10	0.04	—	0.19	0.13	0.00	0.65	0.64	0.79	**0.31**	**0.47**	0.83	0.34
13	MPI-ESM-LR	3	0.14	0.05	**0.53**	—	0.12	0.14	0.61	0.60	0.62	**0.26**	**0.29**	0.92	0.61
14	MRI-CGCM3	5	0.05	0.05	**0.56**	—	0.09	0.23	0.59	0.56	0.48	**0.31**	**0.41**	0.40	0.24
15	CSIRO Mk3.6.0	5	0.02	0.04	0.13	—	0.07	0.06	0.64	0.64	0.75	**0.23**	**0.32**	0.21	0.06
16	GISS-E2H	5	0.06	0.06	**0.65**	**0.55**	0.21	−0.03	0.60	0.58	0.57	**0.25**	**0.29**	0.55	0.30
17	GISS-E2-R	5	0.05	0.05	**0.50**	**0.56**	0.18	0.07	0.60	0.59	0.60	**0.25**	**0.32**	0.49	0.25
18	MIROC5	4	0.04	0.05	**0.31**	0.09	0.16	−0.02	0.66	0.67	0.84	**0.20**	**0.20**	0.34	0.07
19	MIROC-ESM	3	0.05	0.08	**0.68**	0.08	0.05	−0.16	0.67	0.66	0.80	**0.23**	**0.25**	0.48	0.14
—	MM mean	79	0.07	0.05	**0.42**	0.23	0.12	0.03	0.62	0.61	0.64	**0.24**	**0.30**	0.57	0.30
—	MM min	—	0.01	0.04	0.13	−0.04	0.05	−0.16	0.54	0.50	0.26	0.11	0.08	0.15	0.04
—	MM max	—	0.14	0.08	**0.68**	**0.56**	0.21	**0.35**	0.68	0.67	0.84	**0.33**	**0.47**	0.96	0.75
—	OBS (CRU TS3.1)	—	0.08	0.08	—	—	—	—	0.69	0.67	0.78	—	—	0.63	0.26

The IPCC literature is rich in global-scale evaluations of climate models' performances, and these global evaluations are often used to produce regional-scale climate projections. For example, since the multimodel mean captures the global-average temperature trends well, regional-scale climate projections are also based on the multimodel mean (Randall et al. 2007; Christensen et al. 2007). The regional or local details are critical for the planning and design of infrastructure as well as for policy decisions. It is important to evaluate the reliability of the projections of global climate models in a given region vis-à-vis their global performance. Therefore, the third objective of this study is to analyze the effect of multimodel averaging on regional planning decisions.

2. Data and methodology

Nineteen climate models from twelve different modeling institutes are analyzed. The models are listed in Table 1 for temperature and Table 2 for precipitation (see Table 3 for full model names). The CMIP5 historical simulations, in which all forcings have been applied to the models for the period 1850–2005, including anthropogenic greenhouse gas concentrations and volcanic aerosols, were used in this study. In 15 out of 19 models analyzed in this study, multiple realizations of the historical simulation were produced by applying the same external forcings to runs with different initial conditions, typically taken from different times in the control run for that model (Taylor et al. 2012). All ensemble members of a given model are treated independently, resulting in a total of 79 members. All climate model outputs are regridded to a common resolution (2.5° × 2.5°) using an area-average preserving method, and monthly outputs are aggregated to produce mean annual time series for precipitation and temperature. The Climatic Research Unit Time Series 3.1 (CRU TS3.1) for temperature and precipitation are used as observations in this study (Mitchell and Jones 2005). We have analyzed model simulations and observation in the global continental area between 60°S and 60°N for the period 1930–2004 (75 years) to ensure higher reliability in the observations. We have also investigated the effect of observational data uncertainty by repeating the observational analysis with alternative datasets: the Hadley

TABLE 2. As in Table 1, but for precipitation trend (mm decade^{-1}) and persistence analysis results (1930–2004). Here, Y represents skewness coefficient.

Model number	Model name	Ensemble members	Trend			Trend pattern correlation				HC			Trend sig. area	
			Global mean	Spatial σ	Spatial Y	r_1	r_2	r_3	r_4	M_1	M_2	Sig. area	STP	LTP
1	CanESM2	5	−1.46	10.07	−0.46	**0.41**	—	0.12	0.10	0.51	0.50	0.07	0.11	0.10
2	CCSM4.0	5	1.24	8.92	0.38	**0.31**	—	0.12	0.08	0.51	0.50	0.06	0.07	0.07
3	CNRM-CM5	9	−1.36	8.25	−0.63	**0.28**	—	0.11	−0.02	0.52	0.50	0.08	0.08	0.07
4	GFDL CM3	4	−3.49	11.22	−1.24	**0.57**	0.08	**0.18**	**0.20**	0.51	0.50	0.07	0.16	0.14
5	GFDL-ESM2G	3	−1.17	11.15	−0.05	**0.25**	0.10	0.06	**0.16**	0.52	0.50	0.13	0.11	0.09
6	GFDL-ESM2M	1	−0.02	9.77	0.28	—	**0.19**	0.08	**0.21**	0.52	0.50	0.10	0.06	0.05
7	HadCM3	10	−1.04	9.17	−0.28	**0.20**	**0.14**	0.09	0.05	0.52	0.50	0.11	0.09	0.07
8	HadGEM2-CC	1	−4.22	10.02	−0.98	—	**0.28**	**0.14**	−0.02	0.53	0.50	0.17	0.13	0.10
9	HadGEM2-ES	4	−3.63	10.32	−1.00	**0.34**	**0.29**	**0.16**	0.03	0.53	0.50	0.20	0.15	0.11
10	INM-CM4	1	0.17	8.17	0.03	—	—	0.11	0.07	0.52	0.50	0.08	0.08	0.07
11	IPSL-CM5A-LR	5	0.36	10.59	0.03	**0.40**	**0.34**	0.11	0.07	0.53	0.50	0.17	0.16	0.14
12	IPSL-CM5A-MR	1	−0.26	8.67	1.01	—	**0.34**	0.10	0.05	0.53	0.50	0.18	0.11	0.09
13	MPI-ESM-LR	3	1.07	10.06	0.25	**0.41**	—	0.04	−0.11	0.53	0.50	0.15	0.13	0.10
14	MRI-CGCM3	5	−3.72	19.50	−2.06	**0.73**	—	**0.15**	0.12	0.52	0.50	0.13	0.21	0.19
15	CSIRO Mk3.6.0	5	−3.11	13.26	−1.49	**0.32**	—	**0.13**	0.12	0.52	0.50	0.11	0.10	0.09
16	GISS-E2H	5	−3.83	16.17	−1.09	**0.62**	0.45	**0.14**	0.06	0.52	0.50	0.10	0.19	0.17
17	GISS-E2-R	5	−3.19	14.02	−1.35	**0.48**	0.45	**0.16**	0.11	0.52	0.50	0.09	0.18	0.16
18	MIROC5	4	−3.51	16.60	−1.93	**0.66**	0.21	**0.15**	**0.18**	0.53	0.50	0.15	0.15	0.14
19	MIROC-ESM	3	−2.70	11.51	−0.77	**0.60**	0.21	**0.16**	**0.23**	0.53	0.50	0.19	0.17	0.12
—	MM mean	79	−1.78	11.45	−0.60	**0.44**	**0.26**	0.12	0.09	0.52	0.50	0.12	0.13	0.11
—	MM min	—	−4.22	8.17	−2.06	0.20	0.08	0.04	−0.11	0.51	0.50	0.06	0.06	0.05
—	MM max	—	1.24	19.50	1.01	**0.73**	**0.45**	**0.18**	**0.23**	0.53	0.50	0.20	0.21	0.19
—	OBS (CRU TS3.1)	—	0.78	12.20	−0.40	—	—	—	—	0.57	0.52	0.32	0.25	0.18

Centre/Climatic Research Unit, version 4 (HadCRUT4), temperature analysis (Morice et al. 2012), and the Global Historical Climatology Network, version 2 (GHCN v2), precipitation analysis (Vose et al. 1992).

We have employed a nonparametric trend detection technique. This method has the following advantages: 1) there are no prior assumptions about the shape of the trend, for example, it need not be linear; 2) the

TABLE 3. Model name expansions.

Model name	Expansion
CanESM2	Second Generation Canadian Earth System Model
CCSM4.0	Community Climate System Model, version 4.0
CNRM-CM5	Centre National de Recherches Météorologiques Coupled Global Climate Model, version 5
GFDL CM3	Geophysical Fluid Dynamics Laboratory Climate Model, version 3
GFDL-ESM2G	Geophysical Fluid Dynamics Laboratory Earth System Model with OLD ocean component
GFDL-ESM2M	Geophysical Fluid Dynamics Laboratory Earth System Model with MOM4 ocean component
HadCM3	Hadley Centre Coupled Model, version 3
HadGEM2-CC	Hadley Centre Global Environmental Model, version 2 (Carbon Cycle)
HadGEM2-ES	Hadley Centre Global Environmental Model, version 2 (Earth System)
INM-CM4	Institute of Numerical Mathematics Coupled Model, version 4
IPSL-CM5A-LR	L'Institut Pierre-Simon Laplace Coupled Model, version 5, coupled with NEMO (low resolution)
IPSL-CM5A-MR	L'Institut Pierre-Simon Laplace Coupled Model, version 5, coupled with NEMO (medium resolution)
MPI-ESM-LR	Max Planck Institute Earth System Model, low resolution
MRI-CGCM3	Meteorological Research Institute Coupled Atmosphere–Ocean General Circulation Model, version 3
CSIRO Mk3.6.0	Commonwealth Scientific and Industrial Research Organisation Mark, version 3.6.0
GISS-E2H	Goddard Institute for Space Studies Model E, coupled with the HYCOM ocean model
GISS-E2-R	Goddard Institute for Space Studies Model E, coupled with the Russell ocean model
MIROC5	Model for Interdisciplinary Research on Climate, version 5
MIROC-ESM	Model for Interdisciplinary Research on Climate, Earth System Model

interannual-to-interdecadal variability in temperature time series is better captured relative to a regression-based linear trend line estimate; 3) it is distribution independent, that is, data need not conform to a Gaussian distribution; and 4) this method provides a quantitative measure of each model's relative performance, which is generally difficult to get in a typical anomaly time series plot (see Fig. 1). The magnitude of trends is determined using the Theil–Sen approach (TSA; Theil 1950; Sen 1968). If $x_1, x_2, x_3, \ldots, x_n$ is a time series X_t of length n, then the TSA slope is given by Eq. (1):

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

The trend's statistical significance is determined using the Mann–Kendall test (Mann 1945; Kendall 1975). The Mann–Kendall test is distribution independent, robust against outliers, provides a more powerful analysis method for nonnormally distributed data (Yue et al. 2002; Onoz and Bayazit 2003), and has been extensively applied in hydroclimatic trend studies [Kumar et al. (2009), and references therein]. The Mann–Kendall test requires serially independent data. We have employed the Mann–Kendall test under two considerations: 1) short- and 2) long-term persistence. A significant lag-1 autocorrelation (90% confidence interval) is removed from the data (i.e., X_t) using a trend-free prewhitening (TFPW) technique for calculating the statistical significance of trends under short-term persistence (STP) consideration (Yue et al. 2002).

The LTP is quantified in terms of the Hurst coefficient (HC; Hurst 1951). The concept of the Hurst coefficient is based on the hypothesis that hydroclimatic time series (i.e., X_t) exhibit scale invariant properties (Koutsoyiannis 2003), that is,

$$[Z_i^{(k)} - k\mu] \triangleq \left(\frac{k}{l}\right)^H [Z_j^{(l)} - l\mu], \tag{2}$$

where $\triangleq$ represents equality in finite dimensional joint distribution, μ is the expected value of X_t, $Z_i^{(k)}$ is the aggregated stochastic process at a scale k (an integer greater than 1) defined as

$$Z_i^k = \sum_{p=(i-1)k+1}^{ik} X_p, \tag{3}$$

where $Z_j^{(l)}$ is the aggregated stochastic process at scale l, and H is the Hurst coefficient. For example, at given scale $l = 5$, the aggregated stochastic process is as follows:

$$Z_1^5 = \sum_{p=1}^{5} x_p, Z_2^5 = \sum_{p=6}^{10} x_p, \ldots, Z_m^5 = \sum_{p=(m-1)\times 5+1}^{m\times 5} x_p.$$

For a stationary and positively correlated time series, H varies in the range (0.5, 1); $H = 0.5$ indicates independence and increasing values of H represent increasing LTP intensities (Koutsoyiannis and Montanari 2007). In this study, the Hurst coefficient is estimated from the detrended data as the maximum likelihood estimator of a fractional Gaussian noise process [Eq. (2)] (Mandelbrot 1965; McLeod and Hipel 1978; Hamed 2008). Non-Gaussian data can also follow the fractional Gaussian noise process (Koutsoyiannis 2000, 2003). A full description of the application of these statistical methods is given in Kumar et al. (2009).

Figure 2 shows an example of two temperature anomaly series in which one has LTP and other does not have LTP. Linear trends using conventional regression methods and their significance are also shown (Fig. 2, Table 4). The magnitude of the trend does not differ significantly between the two methodologies. This provides consistency between this study and other studies that have used a conventional regression method for estimating the magnitude of the trend. However, the significance of the trend differs substantially, particularly after consideration of LTP. For example, a negative temperature trend in the eastern United States, generally referred to as the ''warming hole'' (Pan et al. 2004; Meehl et al. 2012), is statistically significant using a conventional method as well as the Mann–Kendall test using STP only. After consideration of LTP, the variance of the eastern United States temperature increases by a factor of 4 (not shown), which results in a statistically insignificant trend. Since, LTP is not apparent in the southwestern Russia temperature data, the warming trend in the region is statistically significant in both methodologies. A full description of the statistical method is beyond the scope of this study and can be found in the referred literature.

3. Results

a. Temperature and precipitation trends

The geographic distributions of observed and multimodel ensemble–average twentieth-century temperature trends (1930–2004) are shown in Fig. 3. For the multimodel-ensemble average, temperature trends from different ensemble members within the same model are averaged first, and then the temperature trends for different models are averaged. The multimodel ensemble–average results are virtually identical if we average all 79 ensemble members with equal weight (not shown). The

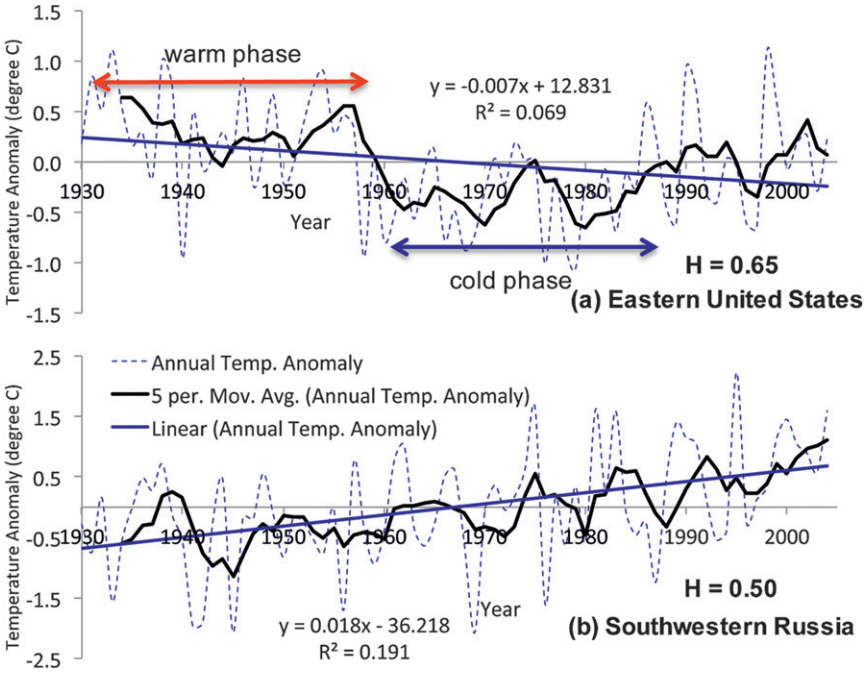

FIG. 2. Temperature anomaly time series for: (a) eastern United States and (b) southwestern Russia. These regions are shown in Fig. 3a. Significant negative temperature trends in the eastern United States become statistically insignificant after consideration of LTP. Temperature anomaly series LTP for southwestern Russia is not apparent.

observation shows fastest warming in north-central Asia and warming holes in the eastern United States, south-central China, west-central South America, and west-central Africa. The multimodel-ensemble average shows a warming trend in all regions, with less spatial variability compared with the observation. The warming holes region in southeastern China shows less warming relative to other regions in the multimodel ensemble–average results. A number of factors, for example, cloud–aerosol interactions, land use change, and natural climate variability could be responsible for less warming in southeastern China (Giorgi et al. 2003; Pielke et al. 2011). Extreme temperature trends (both positive and negative), as found in the observations, are underestimated by the multimodel-ensemble average.

Figure 4 shows the intra- and intermodel spread of the spatial distribution (land only) of the temperature trend from 19 CMIP5 climate models. The spatial distribution of the temperature trend generally conforms to a Gaussian distribution (not shown). The multimodel ensemble–average temperature trend (0.07°C decade^{-1}) is consistent with the observed temperature trend (Fig. 1). The intermodel spread of the ensemble-mean temperature trend from each model ranges from 0.01° to 0.14°C decade^{-1} (Table 1). The top five (i.e., the fastest warming) models are the IPSL-CM5A-LR, MPI-ESM-LR, CCSM4.0, IPSL-CM5A-MR, and CanESM2 models; the bottom five (i.e., the slowest warming) models are GFDL CM3, HadGEM2-ES, HadGEM2-CC, CSIRO Mk3.6.0, and MIROC5. Generally, intra- is lower than intermodel spread in the temperature trends, with a few exceptions,

TABLE 4. Comparison of temperature trends from different methodologies in the eastern United States (30°–41°N, 265°–280°E) and southwestern Russia (45°–60°N, 40°–60°E). For the conventional method (this study), magnitude (TSA trend) is provided in degrees Celsius per decade. Boldface values are statistically significant (i.e., trend, HC, and lag-1 correlation). Mann–Kendall Z is the trend test statistic. For a statistically significant trends at a 95% confidence interval, Mann–Kendall $Z \geq 1.96$ (for increasing trend) or Mann–Kendall $Z \leq -1.96$ (for decreasing trend). See Kumar et al. (2009) for details.

	Eastern United States	Southwestern Russia
Conventional method		
Magnitude	**−0.07**	**0.18**
t value	−2.31	4.14
This study		
Lag-1 autocorrelation	**0.29**	**0.21**
HC	**0.65**	0.50
TSA trend	−0.08	**0.16**
Mann–Kendall Z	−2.43	3.46
Mann–Kendall Z with STP	−2.91	3.23
Mann–Kendall Z with LTP	−1.44	3.23

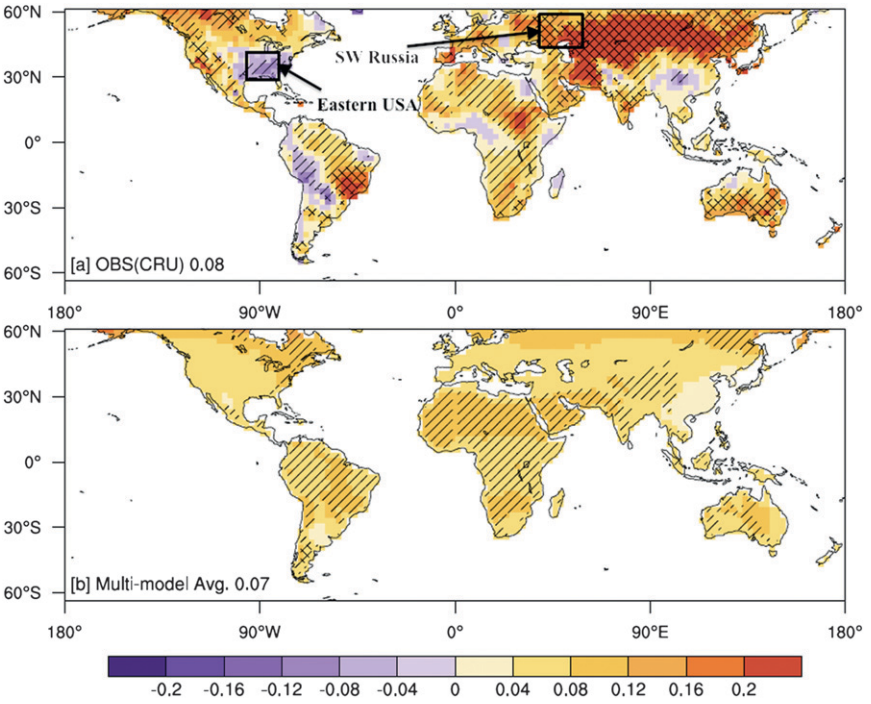

FIG. 3. Geographic distribution of temperature trends (°C decade^{-1}; 1930–2004) from (a) observation and (b) multimodel-ensemble average. Hatching represents statistically significant trends at 95% confidence interval: single hatching shows that the trend is significant with consideration of STP and cross hatching shows that the trend is significant with consideration of both STP and LTP. For the multimodel average, the statistical significance of a trend is determined using the rule of majority, that is, at least 40 or more ensemble members (from 79) show a statistically significant trend.

such as in the GFDL CM3 model. Most CMIP5 models underestimate the spatial variability of the temperature trend (Fig. 4). The overall model spatial standard deviation (0.05°C decade^{-1}) is 40% lower than the observed spatial standard deviation (0.08°C decade^{-1}) in the temperature trend.

Figures 5a–c show intramodel, intra-institute, and intermodel pattern correlation of the spatial distribution of the (land only) temperature trend. For a given model, two different ensemble members are randomly selected for intramodel pattern correlation: a second model from the same institute as the given model whose ensemble member is randomly selected to compute the intra-institute pattern correlation and a second model from the remaining 18 models whose ensemble member is randomly selected to compute the intermodel pattern correlation. This process is repeated 500 times for randomly selected ensemble members of the given model and the second model. The mean and the 95% uncertainty range of 500 iterations are shown in Fig. 5. The 500 iterations are used to provide a reasonably large number to sample all possible combinations of a different model and/or ensemble member. The significance of pattern correlation is calculated using the equivalent sample size method (Zwiers and von Storch 1995) implemented in the NCAR Command Language (NCAR 2012). The pattern correlation for the temperature trend is highest when comparing different ensemble members from the same model, lower when comparing different models from the same institute, and lowest for different models from different institutes (Table 1). The intramodel pattern correlation is significant for most CMIP5 models (Fig. 5a). The intra-institute pattern correlations are generally not significant except for the GISS and the two Hadley Centre models (Fig. 5b). The intermodel pattern correlations are generally positive but not significant among CMIP5 models (Fig. 5c). This issue is particularly important for making multimodel average–based climate change projections at regional scale (e.g., Fig. 3b). The pattern correlations between simulated and observed temperature trend are not statistically significant for most CMIP5 models except for CCSM4.0 (Fig. 5d).

Observed and multimodel ensemble–average precipitation trends are shown in Fig. 6. Globally, the precipitation trends are spatially distributed about zero in both the multimodel-ensemble average (mean of −1.78 and standard deviation of 11.45 mm decade^{-1}) and the observation (mean of 0.78 and standard deviation of

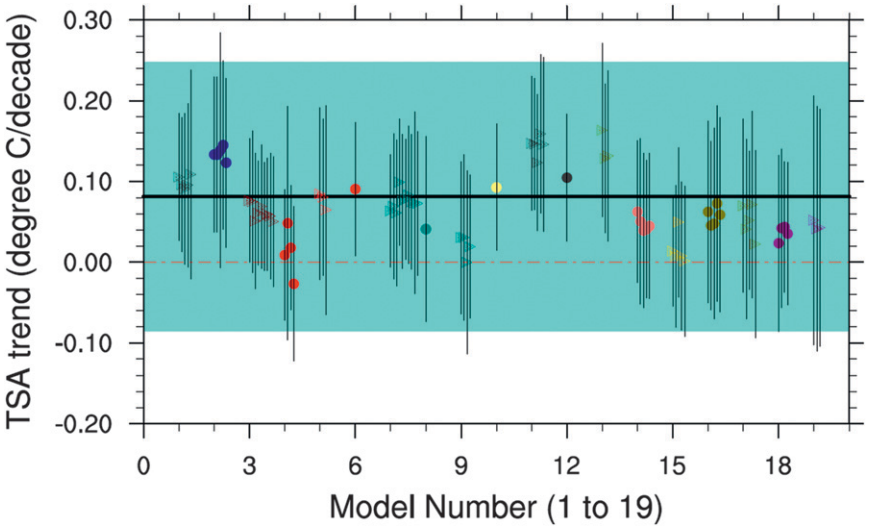

FIG. 4. Mean (colored markers) and 95% range (error bar) of the spatial distribution of temperature trends for all models (60°S–60°N, land only). Marker color is the same as in Fig. 1, and individual ensemble members for a given model are displaced along the *x* axis according to model number (Table 1). Mean observed temperature trend is shown by the thick black line and color shading represents 95% range of spatial distribution of an observed temperature trend (std dev = 2).

12.20 mm decade^{-1}). For many models, the spatial distribution of the precipitation trend is negatively skewed (Table 2). The median and the 95% range of the spatial distribution of the precipitation trends are shown in Fig. 7. Interestingly, the CMIP5 climate models capture the spatial variability in precipitation trends (6% underestimation) better compared to the spatial variability in the temperature trends (40% underestimation). The intramodel, intra-institute, and intermodel spread in the pattern correlation of a precipitation trend are shown in Figs. 8a–c, which is similar to the pattern correlation results for the temperature trends, that is, highest for intramodel, followed by intra-institute, and then intermodel. One important difference between the precipitation and temperature trend results is that the relative performance/ranking of the models for the spatial correlation of the precipitation trend is different from the relative performance of the models for the temperature trends (Figs. 5, 8). For example, the MIROC and GFDL models have significant pattern correlation (~0.20) with the observed precipitation trend (Fig. 8d).

By comparing Tables 1 and 2 and related figures (Figs. 4, 5, 7, and 8), it is remarkable to see that the model skill for the geographical distribution of the precipitation trends is comparable to that for the temperature trends. It is generally believed that climate models simulate temperature better than precipitation (Randall et al. 2007); we do not find this assertion true for mean annual temperature and precipitation trends in CMIP5 twentieth-century climate simulations. For example, the average intramodel spatial correlations for temperature and precipitation trends are 0.42 and 0.44, respectively. One model (CCSM4.0) has a statistically significant pattern correlation with observations of the temperature trends, and five models (all GFDL and MIROC) show statistically significant pattern correlations with observations of precipitation trends. For all climate models, the global land–average precipitation trends are comparable to the observations. The spatial variability (standard deviation) is better captured in the precipitation trend than the temperature trend.

b. *Long-term persistence in temperature and precipitation (Hurst coefficient)*

The observed and multimodel ensemble–average Hurst coefficient for temperature, which is a measure of LTP, is shown in Fig. 9. The global-average (land only) Hurst coefficient for temperature in observations is 0.69 and in the multimodel-ensemble average is 0.62. Intermodel spread in global-average Hurst coefficient for temperature ranges from 0.54 to 0.68 (Table 1). The land area fraction with a statistically significant Hurst coefficient (greater than 0.58; 10% significance level) is 0.78 in observations and 0.64 in the multimodel-ensemble average. Both models and observations have statistically significant Hurst coefficients for temperature in Africa, the southern half of Asia, and northern South America (Fig. 9). The observed temperature persistence in southeastern North

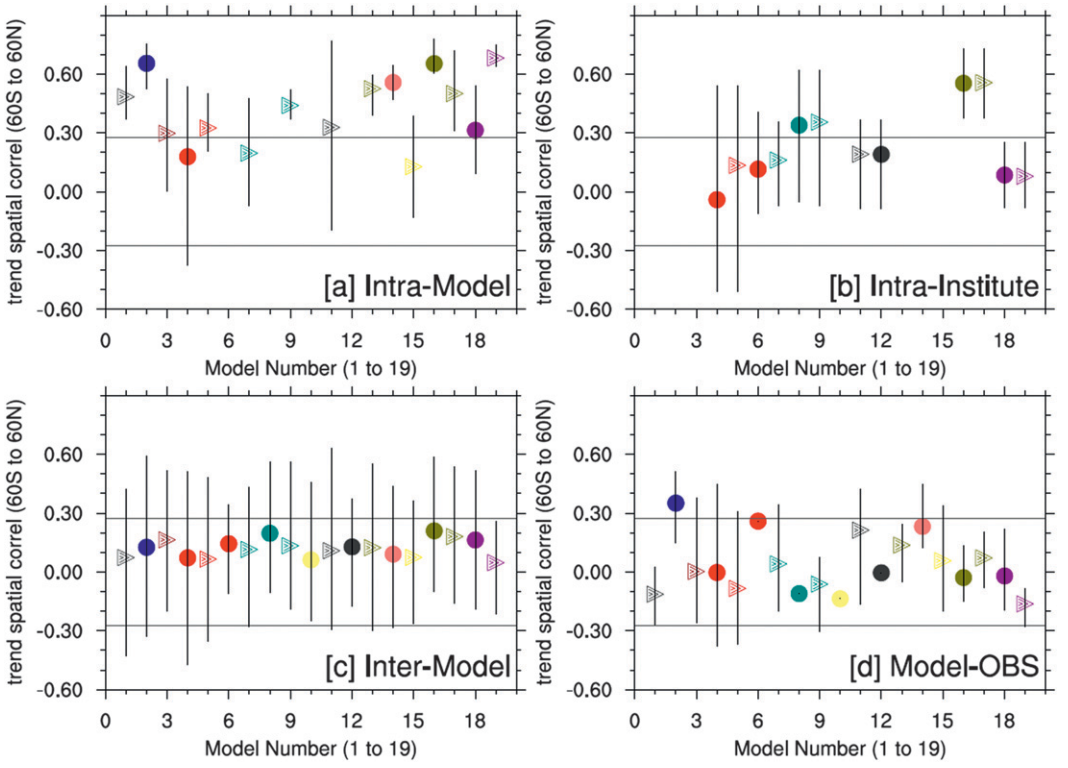

FIG. 5. (a) Intramodel, (b) intra-institute, (c) intermodel, and (d) model–OBS spread for pattern correlation of temperature trend (mm decade^{-1}; 1930–2004). Colored markers show mean value and error bars show 95% range of 500 iterations (see text for details). Thin gray horizontal lines represent statistical significance at the 95% confidence interval of pattern correlation. The statistical significance level is calculated based on the average number of independent samples in the temperature trend data (~50).

America and western Europe is generally not found in CMIP5 simulations (Fig. 9). Most CMIP5 climate models (17 of 19) have a statistically significant pattern correlation with observations for the Hurst coefficient of temperature (Table 1). The intermodel pattern correlation for the Hurst coefficient of temperature is also statistically significant for the same 17 models. Hence, it can be asserted that a significant intermodel pattern correlation can be a good indicator of realistic geographic distribution of projected climate trends and persistence.

The LTP consideration in the trend analysis does not affect the magnitude of trends, but it reduces the land area fraction with statistically significant trends. For example, statistically significant temperature trends in Africa (STP case) do not remain statistically significant after consideration of LTP in both the multimodel-ensemble average and the observation (Fig. 3). For the multimodel-ensemble average, the land area fraction with statistically significant temperature trends reduces from 0.57 in the STP case to 0.30 in the LTP case (Table 1). The corresponding land area fractions in the observations are 0.63 in the STP case and 0.26 in the LTP case. The absence of observed LTP in north-central Asia (Fig. 9) could have partially contributed to statistically significant temperature trends in the region (Fig. 3). Generally, a higher trend magnitude is more likely to be statistically significant compared to lower trend magnitude (see Figs. 3 and 6 and Kumar et al. 2009). North-central Asia is warming almost twice as fast as the global-average warming trends (Fig. 3). Overall, CMIP5 climate models capture the LTP in temperature reasonably well.

The land area fraction with statistically significant LTP in precipitation is 60% less (0.32, Fig. 10a) than that of temperature (0.78, Fig. 9a). The multimodel-ensemble average does not show LTP in precipitation (Fig. 10b, Table 2). In the observations, the land area fraction with a statistically significant precipitation trend reduces from 0.25 in the STP case to 0.18 in the LTP case. The corresponding land area fractions in the multimodel average are 0.13 in the STP case and 0.11 in the LTP case. Overall, the multimodel ensemble–average results from CMIP5 climate models provide a conservative estimate of regional or local precipitation trends (Fig. 6). For example, CMIP5 climate models (the multimodel-ensemble average) underestimate the land area fraction with statistically significant precipitation

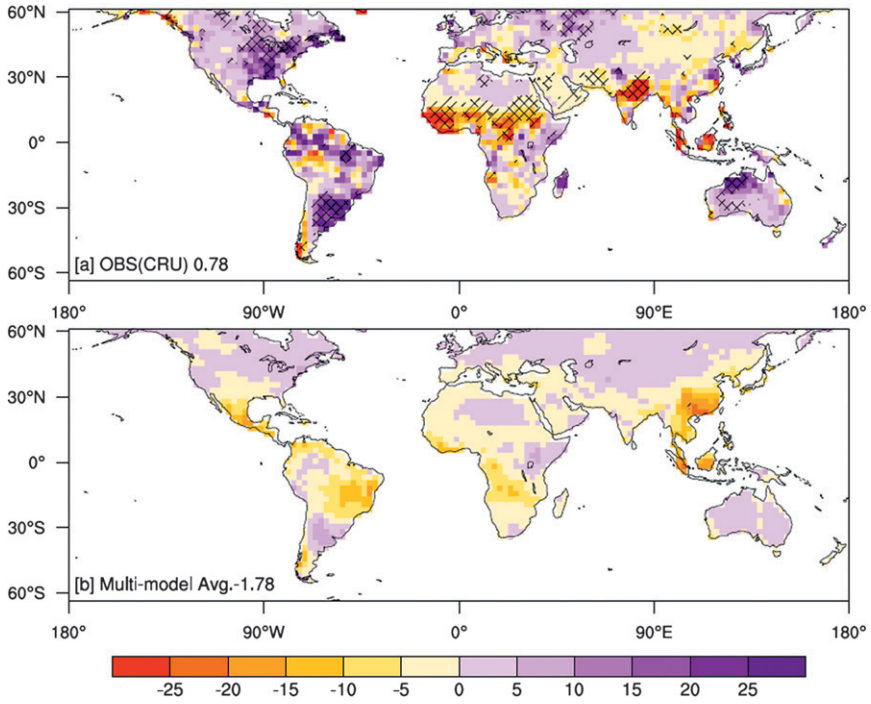

FIG. 6. As in Fig. 3, but for precipitation (mm decade^{-1}).

trends by almost 50% in both the STP and the LTP cases (Table 2).

c. *Effect of multimodel averaging on regional planning*

In the previous sections, it was found that the multimodel-ensemble average is better or at least as good as the best model for estimating the global-average mean temperature and precipitation trends (Tables 1, 2). However, the global-mean values may not be useful in addressing regional issues. For example, it is of interest to know whether India is getting wetter or drier or whether the western United States is warming faster than the eastern United States. To answer such questions, we identified the best simulation, a posteriori, using the pattern correlation with the observations as

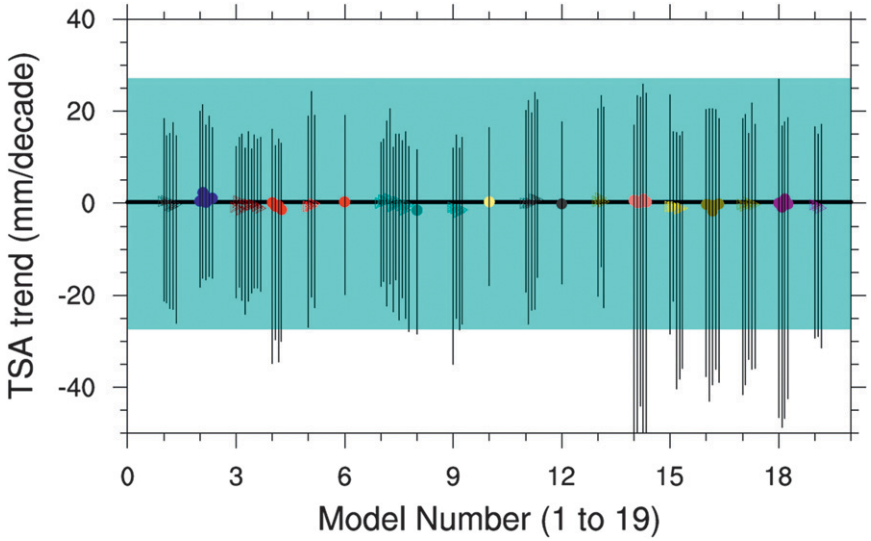

FIG. 7. As in Fig. 4, but for precipitation. Instead of mean, median value of precipitated trends is shown for both models and observations.

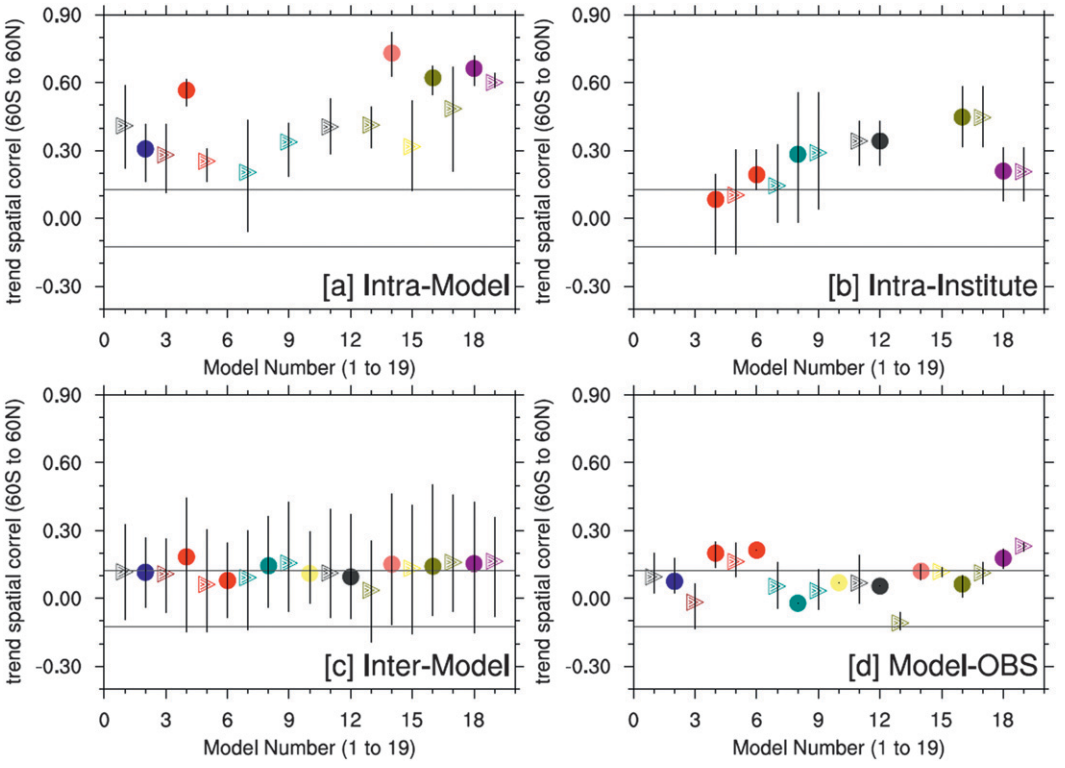

FIG. 8. As in Fig. 5, but for precipitation. Number of independent samples is higher (~250) in precipitation than temperature.

the performance metric. Figure 11 shows the spatial distributions of the temperature and precipitation trends (1930–2004) from the best model and respective best ensemble member in each case. Since the global-mean temperature trend in the best model (CCSM4.0) is significantly higher than the multimodel-ensemble average, its global mean is removed and the multimodel-ensemble global mean is added for presentation in Fig. 11a. In the best model, the Sahel is getting drier, the southern half of South America is getting wetter, central Asia is warming more rapidly, and there are ''warming holes'' in the eastern United States and southeastern China. In the multimodel-ensemble average, these signals are either muted (e.g., the Sahel drying magnitude is reduced) or not present (e.g., the east–west temperature trend gradient in the United States) (Figs. 3b, 6b). This simple exercise shows that there is potentially more useful information available in individual CMIP5 climate simulations (e.g., Fig. 11a) than in the multimodel-ensemble average (Fig. 3b).

To determine how far away the observed spatial distribution of the temperature trend is from the multimodel-ensemble averages, we identified the 10 ''slowest warming'' to ''fastest warming'' world scenarios based on percentile rankings (10th, 20th, . . . , 100th percentile) for each gridbox using a total of 79 ensemble members from all 19 CMIP5 climate models. In this analysis, each gridbox (total 1902) is treated independently and sampled from the population ($n = 79$), then a world map is created and correlated with the observations. Figure 12a shows the pattern correlation of these 10 slowest to fastest warming scenarios with the observations. The global-mean temperature trend (60°S–60°N, land only) for each scenario is also shown. The 50th percentile scenario global-mean temperature trend corresponds to the multimodel-ensemble average (approximately). There is a systematic increase in the pattern correlation going from slowest to fastest warming scenarios, and the pattern correlation is highest for the warmest 10th percentile of the model simulations. A similar exercise for precipitation trends is shown in Fig. 12b. There is a gradual increase in the pattern correlation of the precipitation trends going from ''dry'' to ''wet'' scenarios, which is highest for the 60th–70th percentile scenarios, and then there is sharp decline in pattern correlation after the 70th percentile scenario. Overall a ''drier'' world scenario seems to be better correlated (all correlations are significant for the 50th or less percentile scenarios) than a ''wetter'' world scenario. On a cautionary note, this result depends upon the

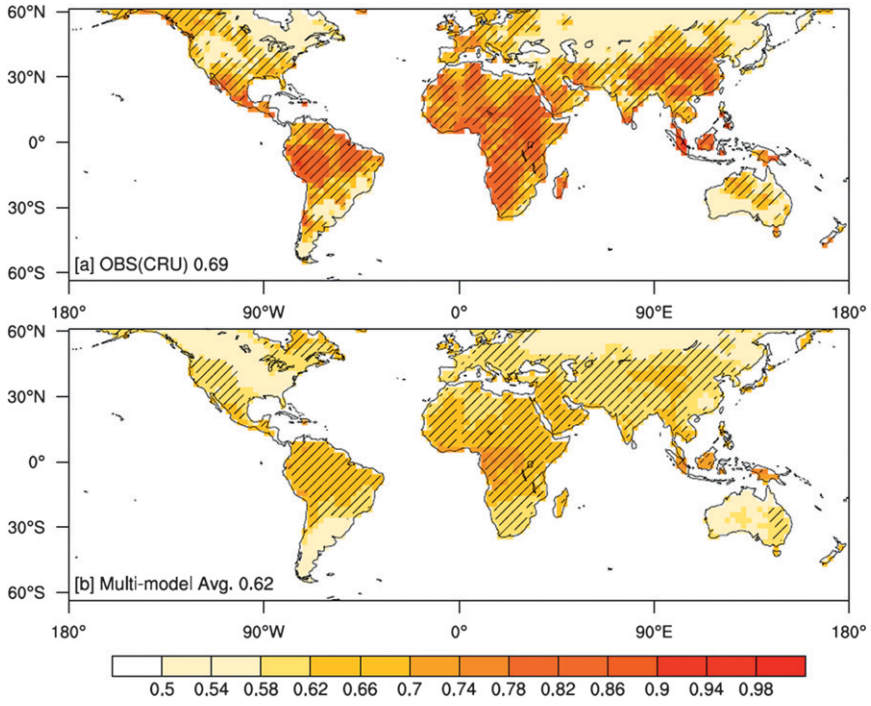

FIG. 9. The HC for temperature from (a) observations and (b) multimodel-ensemble averages. Hatching represents statistically significant HC at 90% confidence interval. For the multimodel ensemble, average statistical significance is calculated using the majority rule as explained in Fig. 2.

question that has been asked, which in this case is related to the spatial distribution of temperature and precipitation trends.

d. Effect of observation uncertainty

CMIP5 twentieth-century climate simulation results (1930–2004) have been compared with observations from the CRU TS3.1 dataset, which has following advantages: 1) global land coverage, 2) available at high resolution (0.5° × 0.5°), and 3) missing data are filled using neighboring stations and nonoverlapping records are also incorporated (Mitchell and Jones 2005). There are several sources of uncertainties in the observation dataset, including structural uncertainty that arises owing to methodology employed in preparing the dataset (Thorne et al. 2005; Pielke et al. 2007), the quality control method, the density of observation stations, and the interpolation and gridding schemes, which may affect the results of this study. To overcome this issue, we have repeated our observational analysis with alternative/better quality-controlled datasets: the HadCRUT4 temperature data (Morice et al. 2012) and GHCN v2 data (Vose et al. 1992). For example, the HadCRUT4 data do not employ any spatial infilling, have better models to describe uncertainties in measurements (e.g., uncertainties from changes in measurement techniques, equipment exposures, and urbanization effect), and have additional newly digitized observations (Morice et al. 2012). HadCRUT4 has a 100-member-ensemble analysis, whose median is employed in this study.

The HadCRUT4 and GHCN v2 data are, however, available at a coarser resolution: 5.0° × 5.0°; hence, the trends and persistence results in the models and in CRU TS3.1 have been degraded to 5.0° × 5.0° resolution for comparison purposes. Figure 13 shows temperature and precipitation trends (1930–2004) from HadCRUT4 and GHCN v2 data, respectively (see Figs. 3a and 6a for comparison). Only 6 months or more of nonmissing data for each year from 1930 to 2004 have been employed to calculate annual-mean temperature and precipitation trends. A substantial land fraction (60°S–60°N)—Africa, South America, and central Asia—has missing data; globally, 0.53 land fraction for temperature and 0.56 land fraction for precipitation have missing data. Table 5 shows a comparison of these data with CRU TS3.1 at collocated regions, that is, regions where alternative observational datasets are available (see Fig. 13). The HadCRUT4 temperature trend and GHCN v2 precipitation trend are highly correlated (spatial) with corresponding CRU TS3.1 data. Higher trend magnitude (0.02°C decade^{-1}) in HadCRUT4 data and (21%) GHCN v2 precipitation data can be expected because

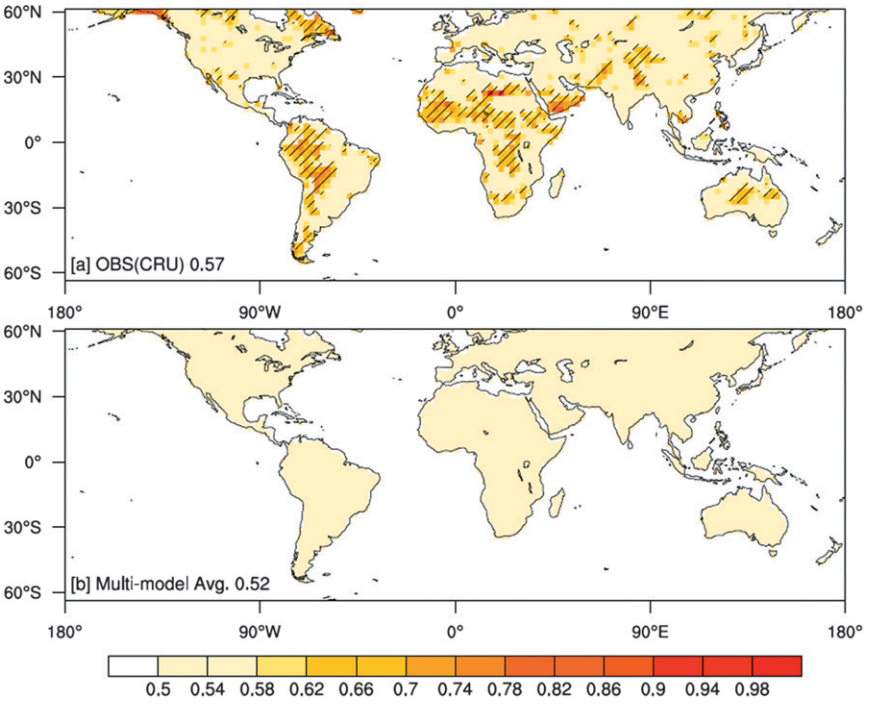

FIG. 10. As in Fig. 9, but for precipitation.

these data do not employ any spatial infilling or interpolation scheme that is likely to smooth out the trend. A statistically significant warming trend in southwestern Russia has been found in both CRU TS3.1 and HadCRUT4 datasets (Figs. 3a and 13a). Similarly, a statistically significant increasing precipitation trend in the eastern United States is also found in both CRU TS3.1 and GHCN v2 datasets.

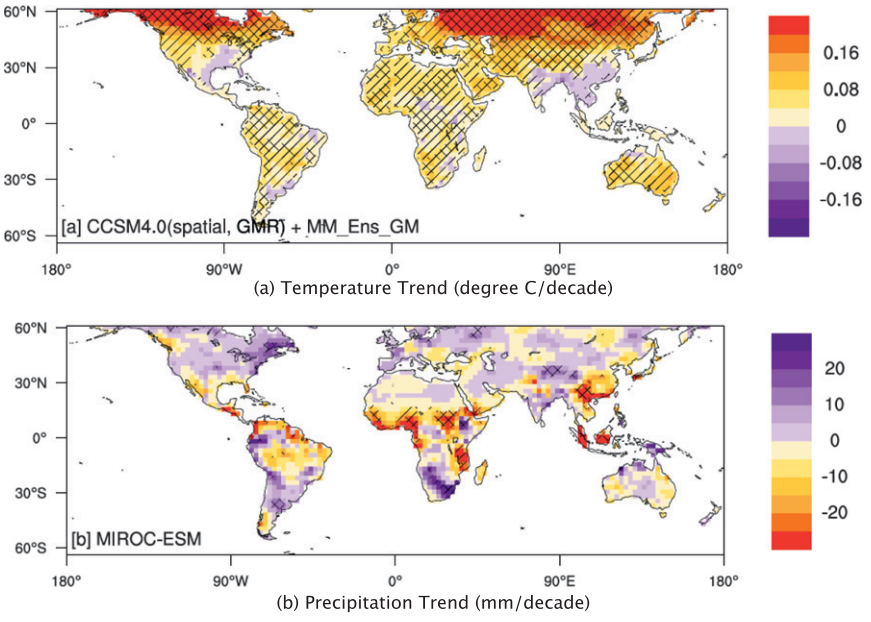

FIG. 11. Spatial distribution of (a) temperature and (b) precipitation trends from models showing highest correlation with the observations. In (a) the global-mean temperature trends from CCSM4.0 are removed (GMR) and replaced with multimodel ensemble–average global-mean trend (MM_Ens_GM). See text for details (section 3c). Hatching is the same as described in Fig. 3.

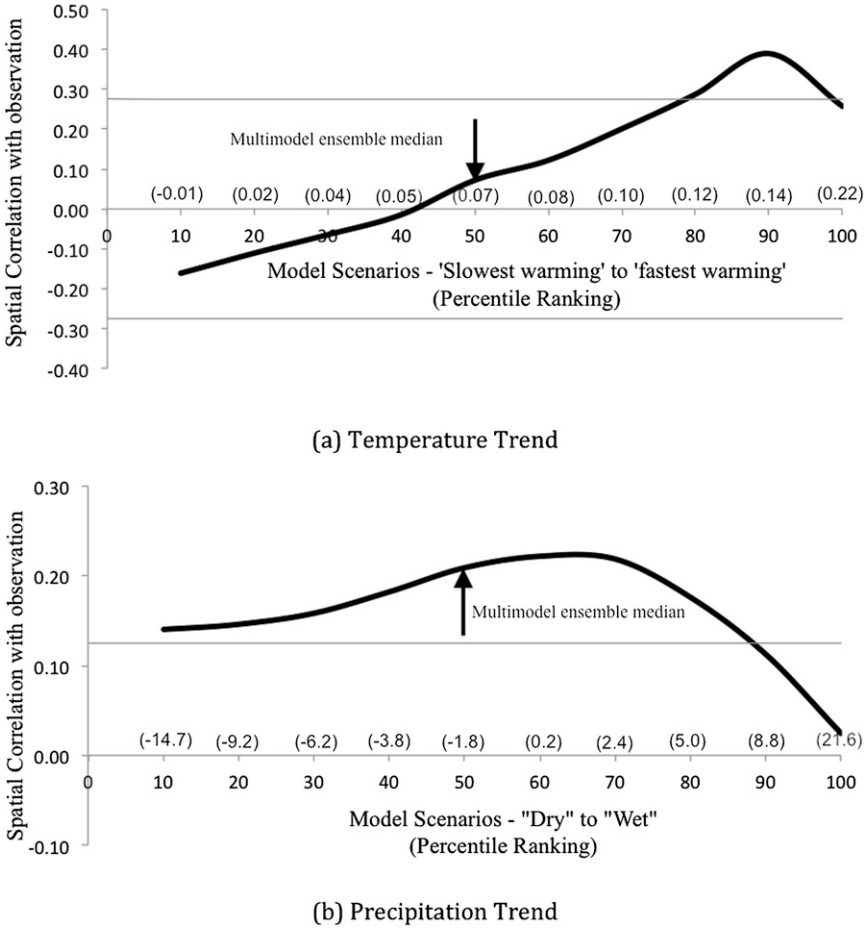

FIG. 12. (a) Comparison of slowest to fastest warming model scenarios with the observations in terms of pattern correlation. The thin gray horizontal line shows statistical significance. (b) As in (a), but for precipitation. Numbers in parentheses on x axis show global-mean trend for the corresponding percentile scenario (temperature: °C decade^{-1}; precipitation: mm decade^{-1}).

Figure 14 shows comparison of CMIP5 model evaluations using HadCRUT4 temperature and GHCN v2 precipitation data and those of CRU TS3.1 data. The relative rankings of models for temperature trend and their spatial correlation with observations are virtually identical across both datasets (Figs. 14a,b). Global-average land precipitation trend magnitude (60°S–60°N) is also highly correlated with the magnitude of the precipitation trend from the limited area where GHCN v2 data are available (R^2 = 0.67; Fig. 14c). The pattern correlation between CMIP5 models and CRU TS 3.1 precipitation trends for the global land area (60°S–60°N) is poorly correlated with the pattern correlation between CMIP5 models and GHCN v2 precipitation trends in the limited area (R^2 = 0.15, not shown). However, in the collocated regions, both pattern correlations are significant (R^2 = 0.49, Fig. 14d). This indicates that a model with the highest precipitation trend pattern correlation for all land areas may not have the highest pattern correlation in a limited subdomain. Overall, the findings of this study are generic enough and independent of the choice of observation datasets.

4. Discussion and conclusions

We have documented promising results for trends and long-term persistence in ''historical'' simulations by CMIP5 twentieth-century climate simulations over continental areas (60°S–60°N). We found that CMIP5 climate models simulate the global-average temperature and precipitation trends (1930–2004) well, that is, the multimodel ensemble–average trends are consistent with the observations. Results for long-term persistence in temperature are also promising. These results can be seen as an improvement over earlier climate models, which generally had difficulties in simulating long-term

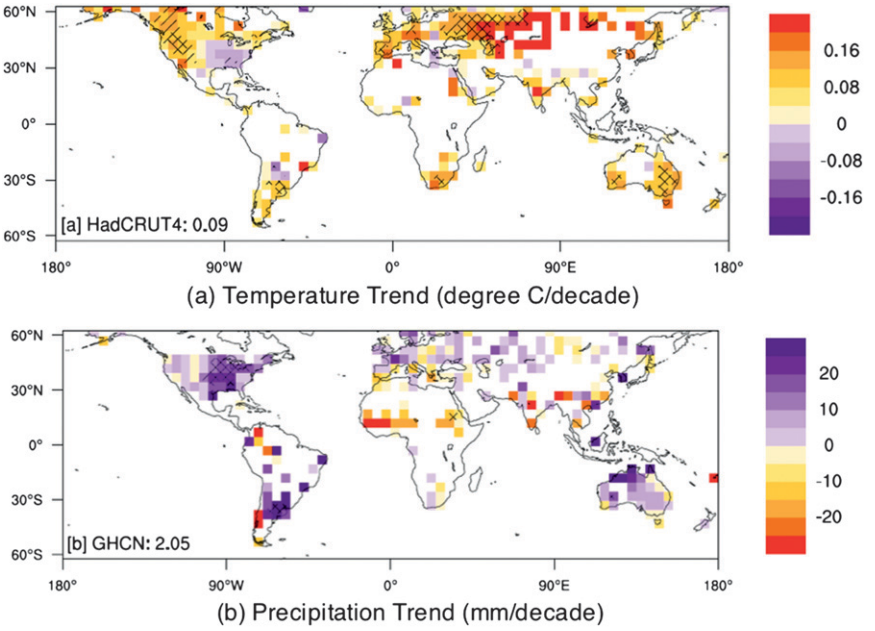

FIG. 13. Temperature and precipitation trends (1930–2004) in alternative observation datasets: (a) HadCRUT4 for temperature and (b) GHCN v2 for precipitation. Only 6 months or greater of nonmissing data for each year from 1930 to 2004 have been employed to calculate annual trend. Hatching is as described in Fig. 3.

persistence (e.g., HadCM2 and HadCM3; Giorgi 2002). The CMIP5 climate models also show some improvements in capturing features like the warming holes for which CMIP3 climate models were challenged (Pan et al. 2004; Kunkel et al. 2006; Zhou and Yu 2006). The CMIP5 climate models, however, continue to have difficulties in simulating observed geographic distributions of temperature and precipitation trends. The simulation of long-term persistence in precipitation is also poor.

Few studies have quantified twentieth-century temperature and precipitation trends using CMIP3 climate simulations at global and regional scales. For example, IPCC AR4 includes a twentieth-century temperature anomaly plot at global and continental scales without explicit trend quantification from the individual model (Hegerl et al. 2007). Hence, it is difficult to compare CMIP5 results from this study with CMIP3 results from the literature. Zhou and Yu (2006) have analyzed twentieth-century temperature trends from 19 CMIP3 climate models over China and the globe. The global-mean temperature trend in CMIP5 twentieth-century climate simulations (0.07°C decade^{-1}; 60°S–60°N, land only; 1930–2004) is the same as CMIP3 climate simulations for a somewhat different region and period (global mean including ocean, 1880–1999; Zhou and Yu 2006). Recently, Sakaguchi et al. (2012a,b) have evaluated the simulation skill for temperature trends from selected CMIP3 and CMIP5 climate models. They found limited skill in the simulation of temperature trends at regional scales in these climate models.

The relative performance of the CMIP5 models for precipitation and temperature is not the same. The model that provides the best simulation of the temperature trend does not necessarily provide the best simulation of the precipitation trend. For example, CCSM4.0 has the highest pattern correlation for temperature trends, whereas GFDL and MIROC models show the highest pattern correlation for precipitation trends. Also, identifying a best model depends on the purpose of the analysis or the question being asked. For example, CCSM4.0 captures the spatial pattern of temperature

TABLE 5. Comparison of CRU TS3.1 with alternative observation datasets in collocated regions for 1930–2004: HadCRUT4 for temperature and GHCN v2 for precipitation. Global land–mean trend (for temperature: °C decade^{-1}; for precipitation: mm decade^{-1}), trend spatial variability (one std dev; for temperature: °C decade^{-1}; for precipitation: mm decade^{-1}) and trend spatial correlation are provided.

For temperature	HadCRUT4		CRU TS3.1
Global land–mean trend	0.09		0.07
Trend spatial variability	0.08		0.08
Trend spatial correlation		0.87	
For precipitation	GHCN v2		CRU TS3.1
Global land–mean trend	2.05		1.69
Trend spatial variability	17.13		10.51
Trend spatial correlation		0.72	

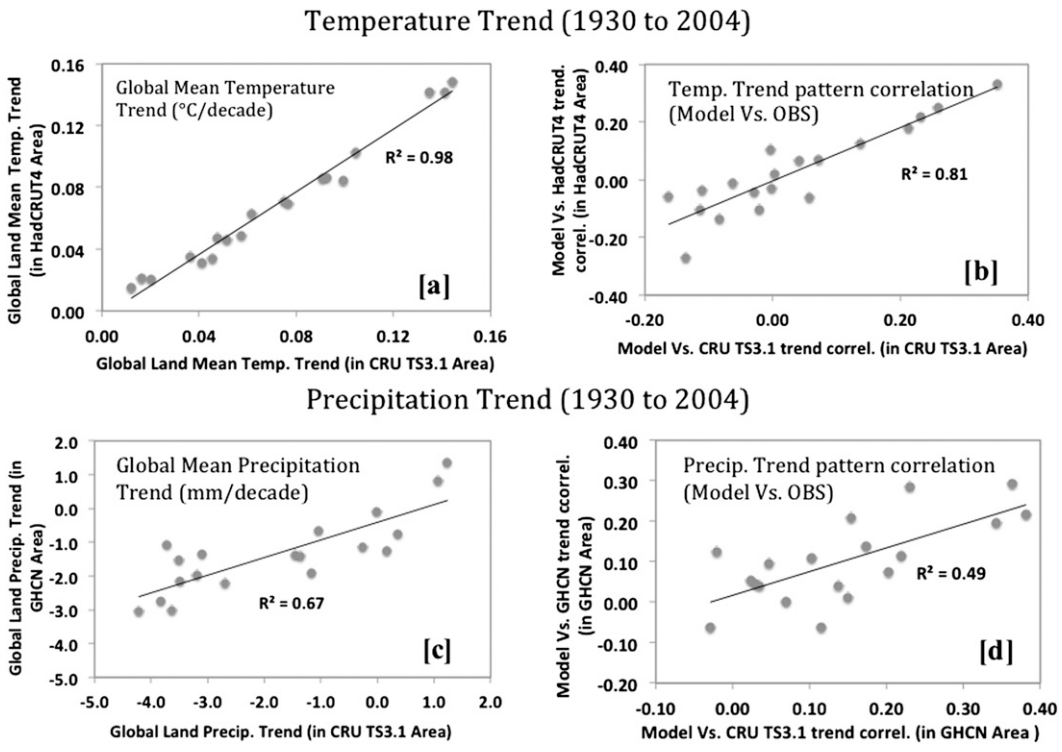

FIG. 14. (a)–(d) Evaluation of CMIP5 twentieth-century climate simulation with CRU TS3.1 observation data vis-à-vis alternative observational datasets shown in Fig. 13 (HadCRUT4 for temperature and GHCN v2 for precipitation). Each gray circle represents ensemble-average results from individual models for a total of 19 models. In (a) and (c), the *x* axis shows model-simulated trends in CRU TS3.1 data coverage area and the *y* axis shows model-simulated trends in alternative observational datasets coverage area. In (b) and (d), the *x* axis shows pattern correlation between model-simulated and CRU TS3.1 trends and the *y* axis shows pattern correlation between model-simulated and alternative observational datasets trends. Please note alternative observation data has almost 50% less spatial coverage than CRU TS3.1 data.

trends better than all other models; however, it underestimates long-term persistence and overestimates the global-average temperature trends (50% or more; Table 1). Therefore, the selection of a particular model for a specific analysis depends on the purpose of the study. With respect to intramodel, intra-institute, and intermodel spread, we found that the models from different institutes have the highest spread, followed by different models from the same institute. Different ensemble members of the same model have least spread.

From a regional natural resource planning perspective, multimodel-ensemble averages provide conservative value for planning or design. For example, the India and West Africa regions are drying much faster (-20 mm decade^{-1}) in the observations than simulations by the multimodel-ensemble average (-5 mm decade^{-1}). Similarly, north-central Asia is warming twice as fast as the global-average warming, which is not found in the multimodel-ensemble average. In Fig. 12, we attempted to compare models' climate projections in a given region vis-à-vis the models' global performance. We found that the multimodel-ensemble average does not represent the best geographic distribution of temperature and precipitation trends. If we can identify the best model for a given purpose, a simple exercise like Fig. 11a can potentially provide more useful information. These examples show that current climate models can bracket the observed trends; further investigation is needed in the area of model selection for regional-scale natural resource planning.

The choice of observation dataset is found to have a minimal effect on the results of this study, particularly for the temperature trends. There are several possible reasons: 1) model performances have been ranked with respect to their global-mean value (e.g., global-mean trend) and pattern correlation for global distribution; 2) different observational datasets use the same historical station observations as their basic input, for example, CRU TS3.1 uses GHCN v2 precipitation data (Mitchell and Jones 2005); and 3) additional newly digitized data brought in from the HadCRUT4 dataset lie outside the analysis of this study (60°S–60°N, land only; 1930–2004), for example, additional sea surface temperature data for the 1850–70 period or improved Arctic coverage mainly

in Russia and Canada. We found that for the precipitation trend, the pattern correlation in the limited area (GHCN area) is not the same as the pattern correlation for global land coverage (CRU TS3.1 area). Thus, results of this study should be used with caution for a limited area.

Overall, we have documented promising results from CMIP5 twentieth-century climate simulations. We expect these results should be useful for both climate scientists as well as natural resource planners. For example, the result that CMIP5 simulates long-term persistence in temperature but not in precipitation is useful for climate scientists, whereas the result that precipitation trends from the multimodel-ensemble average are conservative is useful for water resources planners. A nonparametric trend detection method employed in this study provides a quantitative measure for model intercomparison as well as for variable intercomparison, for example, temperature and precipitation. These results can have important implications for the forthcoming IPCC AR5 in two areas: 1) model selection based on historical performance and 2) regional climate projections.

Acknowledgments. This project received support from the following grants: National Science Foundation (NSF) Rapid Response Grant (RAPID) 1128164, NSF 0947837, NSF 0830068, NSF Career AGS 0847472, NOAA NA09OAR4310058, and NASA NNX09AN50G. We 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 Table 1) 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 leads the development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. We also acknowledge high-performance computing support provided by NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation. We are thankful to three anonymous reviewers for their helpful comments.

REFERENCES

Arnell, N. W., 2011: Incorporating climate change into water resources planning in England and Wales. *J. Amer. Water Resour. Assoc.,* **47,** 541–549.

Asrar, G., A. Busalacchi, and J. Hurrel, 2012: Developing plans and priorities for climate science in service to society. *Eos, Trans. Amer. Geophys. Union,* **93,** 128–128.

Christensen, J. H., and Coauthors, 2007: Regional climate projections. *Climate Change 2007: The Physical Science Basis,* S. Solomon et al., Eds., Cambridge University Press, 847–940.

Giorgi, F., 2002: Variability and trends of the sub-continental scale surface climate in the twentieth century. Part II: AOGCM simulations. *Climate Dyn.,* **18,** 675–691.

——, X. Q. Bi, and Y. Qian, 2003: Indirect vs. direct effects of anthropogenic sulfate on the climate of East Asia as simulated with a regional coupled climate-chemistry/aerosol model. *Climatic Change,* **58,** 345–376.

Hamed, K. H., 2008: Trend detection in hydrologic data: The Mann–Kendall trend test under the scaling hypothesis. *J. Hydrol.,* **349,** 350–363.

——, and A. R. Rao, 1998: A modified Mann–Kendall trend test for autocorrelated data. *J. Hydrol.,* **204,** 182–196.

Hegerl, G. C., and Coauthors, 2007: Understanding and attributing climate change. *Climate Change 2007: The Physical Science Basis,* S. Solomon et al., Eds., Cambridge University Press, 663–745.

Hurst, H. E., 1951: Long term storage capacities of reservoirs. *Trans. Amer. Soc. Civ. Eng.,* **116,** 776–808.

Kendall, M. G., 1975: *Rank Correlation Methods.* 4th ed. Griffin, 202 pp.

Koutsoyiannis, D., 2000: A generalized mathematical framework for stochastic simulation and forecast of hydrologic time series. *Water Resour. Res.,* **36,** 1519–1533.

——, 2003: Climatic change, the Hurst phenomenon, and hydrological statistics. *Hydrol. Sci. J.,* **48,** 3–24.

——, 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,** 171–183.

Kunkel, K. E., X.-Z. Liang, J. Zhu, and Y. Lin, 2006: Can CGCM simulate the twentieth-century "warming hole" in the central United States? *J. Climate,* **19,** 4137–4153.

Mandelbrot, B. B., 1965: Une classe de processus stochastiques homothetiques a soi: Application a la loi climatologique de H.E. Hurst. *C. R. Acad. Sci. Paris,* **260,** 3274–3276.

Mann, H. B., 1945: Non-parametric tests against trend. *Econometrica,* **13,** 245–259.

McLeod, A. I., and K. W. Hipel, 1978: Preservation of the rescaled adjusted range: 1. A reassessment of the Hurst Phenomenon. *Water Resour. Res.,* **14,** 491–508.

Meehl, G. A., J. M. Arblaster, and G. Branstator, 2012: Mechanisms contributing to the warming hole and the consequent U.S. east–west differential of heat extremes. *J. Climate,* **25,** 6394–6408.

Mitchell, T. D., and P. D. Jones, 2005: An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *Int. J. Climatol.,* **25,** 693–712, doi:10.1002/joc.1181.

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.

NCAR, 2012: The NCAR Command Language (version 6.0.0). UCAR/NCAR/CISL/VETS. [Available online at http://www.ncl.ucar.edu/.]

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

Pan, Z., R. W. Arritt, E. S. Takle, W. J. Gutowski Jr., C. J. Anderson, and M. Segal, 2004: Altered hydrologic feedback in a warming climate introduces a "warming hole." *Geophys. Res. Lett.,* **31,** L17109, doi:10.1029/2004GL020528.

Pielke, R. A., Sr., and Coauthors, 2007: Unresolved issues with the assessment of multidecadal global land surface temperature trends. *J. Geophys. Res.,* **112,** D24S08, doi:10.1029/2006JD008229.

——, and Coauthors, 2011: Land use/land cover changes and climate: Modeling analysis and observational evidence. *Wiley Interdiscip. Rev.: Climate Change,* **2,** 828–850, doi:10.1002/wcc.144.

Randall, D. A., and Coauthors, 2007: Climate models and their evaluation. *Climate Change 2007: The Physical Science Basis,* S. Solomon et al., Eds., Cambridge University Press, 589–662.

Sakaguchi, K., X. Zeng, and M. A. Brunke, 2012a: Temporal- and spatial-scale dependence of three CMIP3 climate models in simulating surface temperature trends in the twentieth century. *J. Climate,* **25,** 2456–2470.

——, ——, and ——, 2012b: The hindcast skill of the CMIP ensembles for the surface air temperature trend. *J. Geophys. Res.,* **117,** D16113, doi:10.1029/2012JD017765.

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

Shanahan, T. M., and Coauthors, 2009: Atlantic forcing of persistent drought in West Africa. *Science,* **324,** 377–380.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.,* **93,** 485–498.

Theil, H., 1950: A rank-invariant method of linear and polynomial regression analysis. III. *Proc. Ned. Akad. Wet.,* **53,** 1397–1412.

Thorne, P. W., D. E. Parker, J. R. Christy, and C. A. Mears, 2005: Uncertainties in climate trends. *Bull. Amer. Meteor. Soc.,* **86,** 1437–1442.

von Storch, H., 1995: Misuses of statistical analysis in climate research. *Analysis of Climate Variability: Applications of Statistical Techniques,* H. von Storch and A. Navarra, Eds., Springer, 11–26.

Vose, R. S., R. Helm, R. L. Schmoyer, T. R. Karl, P. M. Steurer, J. K. Eischeid, and T. C. Peterson, 1992: The Global Historical Climatology Network: Long-term monthly temperature, precipitation, sea level pressure, and station pressure data. Rep. ORNL/CDIAC-53, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, TN, 325 pp. [Available online at http://cdiac.ornl.gov/ftp/ndp041/ndp041.pdf.]

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,** 1807–1829.

Zhou, T., and R. Yu, 2006: Twentieth-century surface air temperature over China and the globe simulated by coupled climate models. *J. Climate,* **19,** 5843–5858.

Zwiers, F. W., and H. von Storch, 1995: Taking serial correlation into account in test of mean. *J. Climate,* **8,** 336–351.