Projected Changes in Climate Extremes over the Northeastern United States*

LIANG NING

Key Laboratory of Virtual Geographic Environment Ministry of Education, School of Geography Science, and Jiangsu Key Laboratory for Numerical Simulation of Large Scale Complex System, School of Mathematical Science, Nanjing Normal University, Nanjing, China, and Northeast Climate Science Center, and Department of Geosciences, University of Massachusetts Amherst, Amherst, Massachusetts, and Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing, China

EMILY E. RIDDLE

Department of Geoscience, University of Massachusetts Amherst, Amherst, Massachusetts

RAYMOND S. BRADLEY

Northeast Climate Science Center, and Department of Geosciences, University of Massachusetts Amherst, Amherst, Massachusetts

(Manuscript received 21 February 2014, in final form 3 January 2015)

ABSTRACT

Projections of historical and future changes in climate extremes are examined by applying the bias-correction spatial disaggregation (BCSD) statistical downscaling method to five general circulation models (GCMs) from phase 5 of the Coupled Model Intercomparison Project (CMIP5). For this analysis, 11 extreme temperature and precipitation indices that are relevant across multiple disciplines (e.g., agriculture and conservation) are chosen. Over the historical period, the simulated means, variances, and cumulative distribution functions (CDFs) of each of the 11 indices are first compared with observations, and the performance of the downscaling method is quantitatively evaluated. For the future period, the ensemble average of the five GCM simulations points to more warm extremes, fewer cold extremes, and more precipitation extremes with greater intensities under all three scenarios. The changes are larger under higher emissions scenarios. The inter-GCM uncertainties and changes in probability distributions are also assessed. Changes in the probability distributions indicate an increase in both the number and interannual variability of future climate extreme events. The potential deficiencies of the method in projecting future extremes are also discussed.

1. Introduction

The Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4) indicates that the global mean surface temperature has risen by 0.7°C over the last 100 years, with changes in temperature extremes consistent with warming and substantial increases in heavy precipitation events (Trenberth et al. 2007). Climate extremes have attracted increasing attention because of their large societal impact on multiple sectors, such as agriculture, the economy, and human health (Easterling et al. 2000; Meehl and Tebaldi 2004; Negri et al. 2005). Significant increasing trends in many climate extreme indicators have been reported over many regions using a variety of datasets and methods (Frich et al. 2002; Alexander et al. 2006; Ning and Qian 2009).

General circulation models (GCMs) based on well-established physical principles reproduce observed features of recent climate (Randall et al. 2007), and they are the major tools used to project future climate, including temperature and precipitation extremes (Meehl et al. 2007). However, since some climate extremes, especially precipitation extremes, are mainly controlled by subgrid

---

processes, the coarse resolution of GCMs cannot meet the requirements of end users who need to apply changes of climate extremes to interdisciplinary studies, such as hydrology, conservation, and climate risk assessment (Prudhomme et al. 2002; Diaz-Nieto and Wilby 2005). To bridge this gap, downscaling methods have been developed and widely applied (Wilby and Wigley 1997; Christensen et al. 2007; Maraun et al. 2010). Usually, the downscaling methods are categorized into two kinds: dynamical downscaling, using GCM output to drive regional climate models (RCMs) (e.g., Chen et al. 2003; Plummer et al. 2006), and statistical downscaling, building relationships between observed synoptic circulation and local climate anomalies and then applying this relationship to GCM-simulated synoptic circulation (e.g., Hewitson and Crane 2006; Ning et al. 2012a,b) or building relationships between synoptic-scale outputs of GCMs and local-scale inputs for hydrologic or other land surface process models (e.g., Wood et al. 2002).

Compared to dynamical downscaling, statistical downscaling is generally flexible and less computationally demanding and can effectively remove errors in historical simulated values; however, statistical downscaling is based on the key assumption that the relationship between large-scale weather systems and local climate will remain constant over time (Melillo et al. 2014; Walsh et al. 2014). When applying statistical downscaling methods to simulations of climate extremes, a critical factor is the ability to generate complete distributions of daily variables so they can be used to randomize the downscaled results and thus better represent local variability and extremes (Maraun et al. 2010). The Statistical and Regional Dynamical Downscaling of Extremes for European Regions (STARDEX; http://www.cru.uea.ac.uk/projects/stardex/) project has systematically developed and evaluated 22 different statistical downscaling methods in terms of their ability to reproduce extreme temperature and precipitation indices and has demonstrated that these downscaling methods can be used to construct scenarios of extremes (Goodess et al. 2005). Fowler et al. (2007) reviewed the strengths and weaknesses of downscaling methods for different climate extremes and found that indices related to the frequency of extremes were better reproduced than those related to the magnitude of events, which were generally too moderate compared with observations. Bürger et al. (2012) compared five statistical downscaling methods with respect to their representation of climatic extremes, and all the statistical downscaling methods, especially the bias-correction spatial disaggregation (BCSD) method, were able to reproduce the statistics of observed climate extremes, with relatively better performances on temperature-related indices.

To evaluate the performances of statistical downscaling methods over specific regions, Haylock et al. (2006) compared six statistical and two dynamical downscaling models with regard to their ability to downscale seven seasonal indices of heavy precipitation in northwestern and southeastern England and found that winter indices and indices indicative of rainfall occurrence are generally downscaled better than summer indices and those indicative of rainfall intensity. Schmidli et al. (2007) also evaluated six statistical downscaling models and three RCMs in their ability to downscale daily extreme precipitation statistics over the European Alps and concluded that there is relatively good agreement between the downscaling models for most precipitation statistics, although they also contribute to the uncertainty of regional climate projections, especially for the summer precipitation climate. Over the northeastern United States, when applying both statistical and dynamical downscaling methods to GCM output, Hayhoe et al. (2008) found that the downscaling methods can improve simulations of spatial and temporal variability in temperature and precipitation across the region. Ahmed et al. (2013) also found that, after applying a statistical downscaling method to both multiple GCM and RCM simulations over the northeastern United States, the agreement among the GCM and RCM outputs increased in predicting extreme events in the future climate.

The northeastern United States encompasses enormous diversity in geography, climate, ecological resources, and human land use. It includes 7 of the 21 regions established for the national Landscape Conservation Cooperative (LCC) program and a human population of 131 million (41% of the U.S. population; Palmer et al. 2011). Consequently, the northeastern United States region poses many unique challenges for understanding, adapting to, and mitigating the effects of climate change (Horton et al. 2014; Corell et al. 2014; Ning and Bradley 2014, 2015). In this paper, to better prepare for future climate impacts, historical and future changes of climate extremes over the northeastern United States are systematically investigated through application of the BCSD statistical downscaling method, and it is applied for the first time to five GCMs from phase 5 of the Coupled Model Intercomparison Project (CMIP5). First, this paper systematically evaluates the performance of BCSD on climate extreme estimation over the northeastern United States during the historical period, since BCSD has become widely used to assess climate change impacts. Then, the downscaling method is applied to late twenty-first-century CMIP5 runs under three future emissions [representative concentration pathways (RCPs)] scenarios (Moss et al. 2010; van Vuuren et al. 2011), and probabilistic estimates for

future changes in temperature and precipitation extremes are determined. We focus on uncertainties resulting from the RCP scenarios and different GCM parameterization schemes, since these are likely to dominate other sources of uncertainty by the end of the twenty-first century (Hawkins and Sutton 2009). The downscaled probabilistic changes in climate extremes provide useful insights that will be of interest to a wide range of stakeholders and decision-makers involved in the influence of climate change on hydrology, ecosystems, and agriculture, as well as in adaption and mitigation efforts across the northeastern United States.

2. Data and methodology

a. Data

The study area ranges from 36° to 50°N and from 90° to 68°W, encompassing the United States and Canada (see Fig. 2). The high-resolution ($1/8°$) observed monthly and daily maximum temperature, minimum temperature, and precipitation data used in the statistical downscaling are for the period 1950–99 (Maurer et al. 2002).

The GCM monthly maximum temperature, minimum temperature and precipitation data for the periods 1950–99 and 2050–99 are taken from World Climate Research Programme (WCRP) CMIP5 under the historical emissions scenario and three future RCP scenarios for five different models: 1) the Community Climate System Model, version 4.0 (CCSM4); 2) the Community Earth System Model, version 1 (Community Atmosphere Model, version 5) [CESM1(CAM5)]; 3) the Hadley Centre Global Environment Model, version 2–Earth System (HadGEM2-ES); 4) the Max Planck Institute Earth System Model, medium resolution (MPI-ESM-MR); and 5) the Max Planck Institute Earth System Model, low resolution (MPI-ESM-LR). These five models were selected because of their better performances on large-scale simulations of historical climate extremes evaluated in previous studies (Sillmann et al. 2013). The corresponding daily maximum and minimum temperatures from these five GCMs are also used to investigate the future changes of daily temperature variability. The historical scenario is forced by observed atmospheric composition changes, and the three future RCP scenarios, which include one high emissions scenario (RCP8.5), one midrange mitigation emissions scenario (RCP4.5), and one low emissions scenario (RCP2.6) (Taylor et al. 2012). The monthly maximum and minimum temperature data from another 10 GCMs—CanESM2, CNRM-CM5, CSIRO Mk3.6.0, GFDL CM3, GISS-E2-H, HadGEM2-AO, IPSL-CM5A-MR, MIROC5, MRI-CGCM3, and NorESM1-M (model name expansions are available at http://www.ametsoc.org/PubsAcronymList)—under these four scenarios are also

used to examine the spreads of future maximum and minimum temperature changes. The data and descriptions of the GCMs can be found at the Program for Climate Model Diagnosis and Intercomparison (PCMDI) website (http://pcmdi9.llnl.gov/esgf-web-fe/).

b. Statistical downscaling procedure

In this study, the BCSD method is applied to the GCMs' output. BCSD is one of the most widely used statistical downscaling methods, used to generate high-resolution temperature and precipitation data with significantly reduced climatological biases (Wood et al. 2002, 2004; Maurer 2007).

BCSD contains two major steps: bias correction and spatial disaggregation, and it is applied to each of 12 months (January–December) for each of the three variables (i.e., maximum temperature, minimum temperature, and precipitation), separately. For the bias-correction step, the biases from raw GCM-simulated monthly fields, which are already aggregated to the same coarse 2° resolution, are removed by adjusting the modeled cumulative distribution functions (CDFs) to the observed instrumental CDFs. Then, change factors calculated between the low-resolution bias-corrected GCM fields and observed climatology at 2° are spatially disaggregated to high resolution ($1/8°$) by interpolation. Next, the bias-corrected high-resolution GCM fields are calculated by adding (for maximum or minimum temperature) or multiplying (for precipitation) the high-resolution change factors and the observed climatology at $1/8°$. Finally, for each month, these fields are temporally disaggregated to a daily scale through randomly picking a month from the observational period and shifting the observed daily values by a fixed additive constant so that the monthly mean matches the downscaled monthly value. The major procedures are shown in Fig. 1. For example, for the downscaled maximum temperature of January 1950, one of the 50 observed Januarys (e.g., January 1960) is randomly selected, and then the observed daily maximum temperature values are shifted so that the monthly average maximum temperature equals the downscaled monthly average maximum temperature. The procedure is similar for precipitation, with the only difference being that the observed daily precipitation values are scaled by a multiplicative constant to match the downscaled monthly precipitation. For more details about the BCSD method, readers are referred to previous studies (Wood et al. 2002, 2004; Maurer 2007). The downscaled climate extremes from BCSD over British Columbia for present climate and future projection were evaluated and compared with several other downscaling methods (Bürger et al. 2012, 2013). When assessing the impact of climate changes on streamflow in the Colorado River basin, Harding et al. (2012) pointed out that the BCSD may perform poorly where daily meteorological
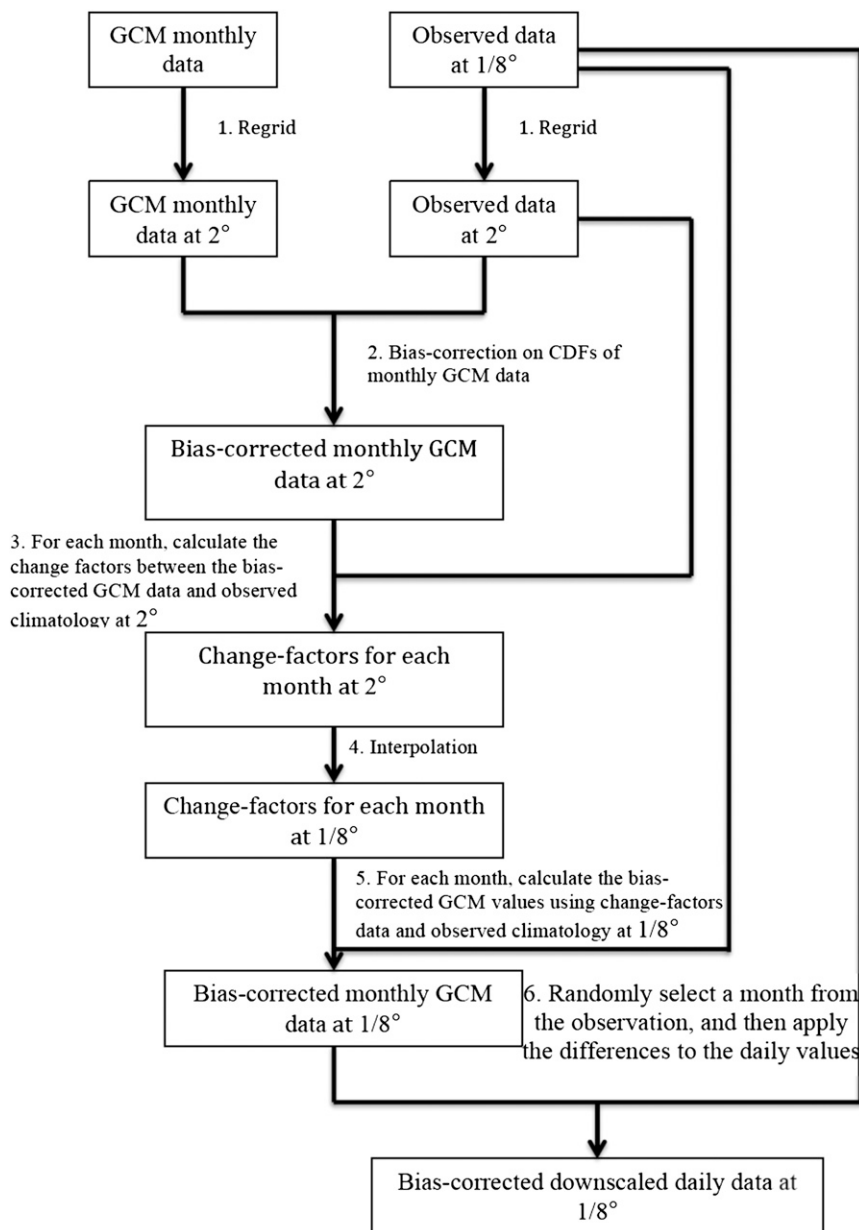
FIG. 1. The major procedures of the BCSD downscaling method.

climatologies are highly skewed or exhibit threshold behavior (e.g., the U.S. Southwest, with intermittency and resulting skewness of precipitation). Ahmed et al. (2013) applied the statistical downscaling and bias correction (SDBC), which is a modification of BCSD method, to six GCMs and four RCMs to predict future climate extremes in the U.S. Northeast. They indicate that bias correction is an effective tool to derive fine-resolution predictions from both GCM and RCM outputs, and can lead to a higher level of agreement among the models in predicting the magnitude and capturing the spatial pattern of the extreme climate indices.

c. Indices of climate extremes

The 11 indices of climate extremes used in this study along with their expansions are listed in Table 1. Most of these indices are taken from the dictionary of the European Climate Assessment and Dataset project (ECA&D), which has been commonly used in previous studies about climate extremes (Frich et al. 2002; Meehl and Tebaldi 2004; Alexander et al. 2006). Precise definitions of the full list of indices are available at the ECA&D website (http:// eca.knmi.nl/indicesextremes/indicesdictionary.php). The definitions of GSL and high values of HWI are adopted

TABLE 1. The definitions and units of the 11 climate extreme indices used in this study (for precise definitions, see http://eca.knmi.nl/indicesextremes/indicesdictionary.php).

	Index	Definition	Unit
Temperature indices	GSL	Growing season length: length between the first period and last period of 5 continuous days with daily average temperature higher than 5°C.	Days
	Tx90p	Warm days: number of days with maximum temperature higher than 90th percentile of daily maximum temperature.	Days
	Tx10p	Cold days: number of days with maximum temperature lower than 10th percentile of daily maximum temperature.	Days
	FD	Frost days: number of days with daily minimum temperature lower than 0°C.	Days
	HWI	Heat wave intensity: maximum average daily minimum temperature of 3 continuous days.	Degrees Celsius
	Tn90p	Warm nights: number of days with minimum temperature higher than 90th percentile of daily minimum temperature.	Days
	Tn10p	Cold nights: number of days with minimum temperature lower than 10th percentile of daily minimum temperature.	Days
Precipitation indices	R95p	Number of days with daily precipitation larger than 95th percentile of daily precipitation amount.	Days
	R95pTOT	Total precipitation amount due to daily precipitation larger than 95th percentile of daily precipitation amount.	Millimeters
	Rx5day	Maximum total precipitation over 5 continuous days.	Millimeters
	R10mm	Number of days with daily precipitation larger than 10 mm.	Days

from Frich et al. (2002) and Meehl and Tebaldi (2004), separately. Since we are concerned about the changes of the absolute extreme precipitation amount more than the percentage, the definition of R95pTOT is slightly different from the original definition in the ECA&D dictionary.

These 11 indices cover a wide range of interdisciplinary interests and concerns about the potential impacts of climate change. For example, the total number of FD and the GSL are important in future agricultural decision-making. High values of HWI can have serious impacts on human health. The number of days with daily precipitation larger than 10 mm is an important factor when driving hydrological models.

3. Results

First, the downscaled climate extremes are systematically compared with observations described in section 2a over the historical period 1950–99 over the northeastern United States. Then future changes of the climate extreme indices are presented, based on the ensemble average across the five GCMs under the three RCP scenarios, and the corresponding uncertainties from the emissions scenarios and different GCMs' responses to the increase in greenhouse gases are discussed.

a. BCSD's performance on reproducing historical climate extremes

In a previous study, Bürger et al. (2012) showed that BCSD performed differently in reproducing different observed climate extremes. In this section, BCSD's

performances in reproducing the 11 climate indices over the northeastern United States are first quantitatively evaluated on three aspects:

1) The downscaled averages are compared with observations through a t test.
2) The downscaled variances are compared with observations through an F test.
3) The downscaled CDFs are compared with observation through a Kolmogorov–Smirnov test (K–S test).

In each test, the percentage of the grid points with significant differences ($p = 0.05$) between downscaled results and observations is calculated for each index. In this case, a smaller percentage indicates better performance.

Figure 2 shows the spatial distributions of the differences between downscaled and observed averages of the 11 extreme indices over the period 1950–99. Among the seven temperature extremes, the downscaling generally overestimates the magnitudes of GSL and HWI and underestimates the magnitudes of FD over most of the region, indicating mainly warm biases on these three indices. In Figs. 2 and 3, color scales are applied so that red indicates warm (dry) biases and blue indicates cold (wet) biases. The biases range from −3 to +6 days for GSL, from −4 to −1 days for FD, and from 0° to +0.5°C for HWI. The corresponding percentages of grid points with significant differences based on a t test (stippled region in Fig. 2) for these three indices are 10.8%, 53.1%, and 29.6%, respectively. For the other four temperature extremes (i.e., Tx90p, Tx10p, Tn90p, and Tn10p) the magnitudes of differences between the downscaled results
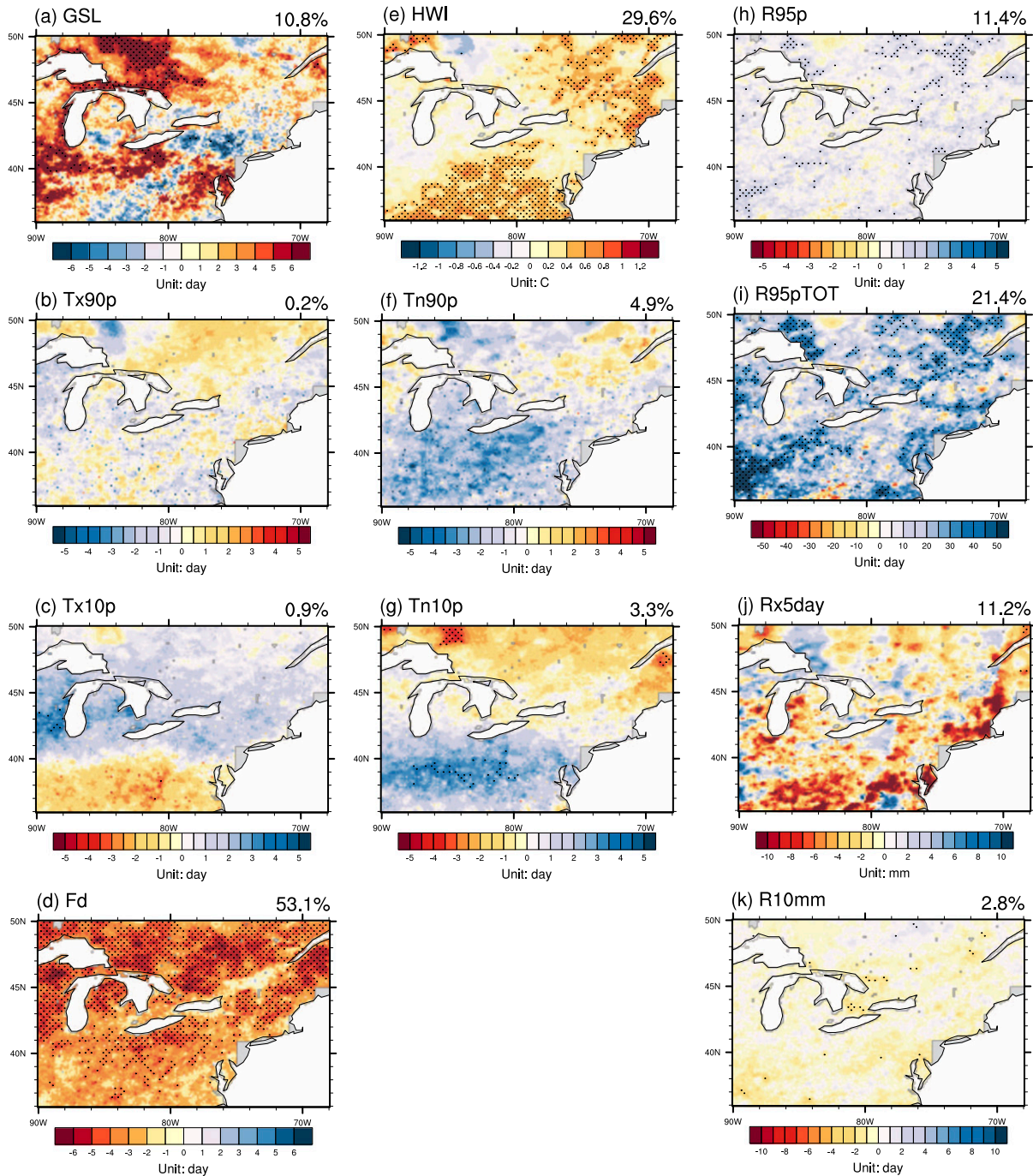
FIG. 2. The spatial distributions of differences between the downscaled and the observed averages of the 11 extreme indices for the period 1950–99. The number in the upper-right corner of each panel shows the percentage of grid points with significant differences. Stippling indicates changes significant at the $p = 0.05$ level based on a t test.

and observations are really small (within ± 3 days), and fewer than 5% of total grid points show significant differences. It can be concluded that the downscaling is good at reproducing the means of GSL, Tx90p, Tx10p, Tn90p, and

Tn10p, while the performances of FD and HWI are not as good as the previous five indices.

For R95p and R95pTOT, the downscaling over-estimates the magnitudes over most of the region, and
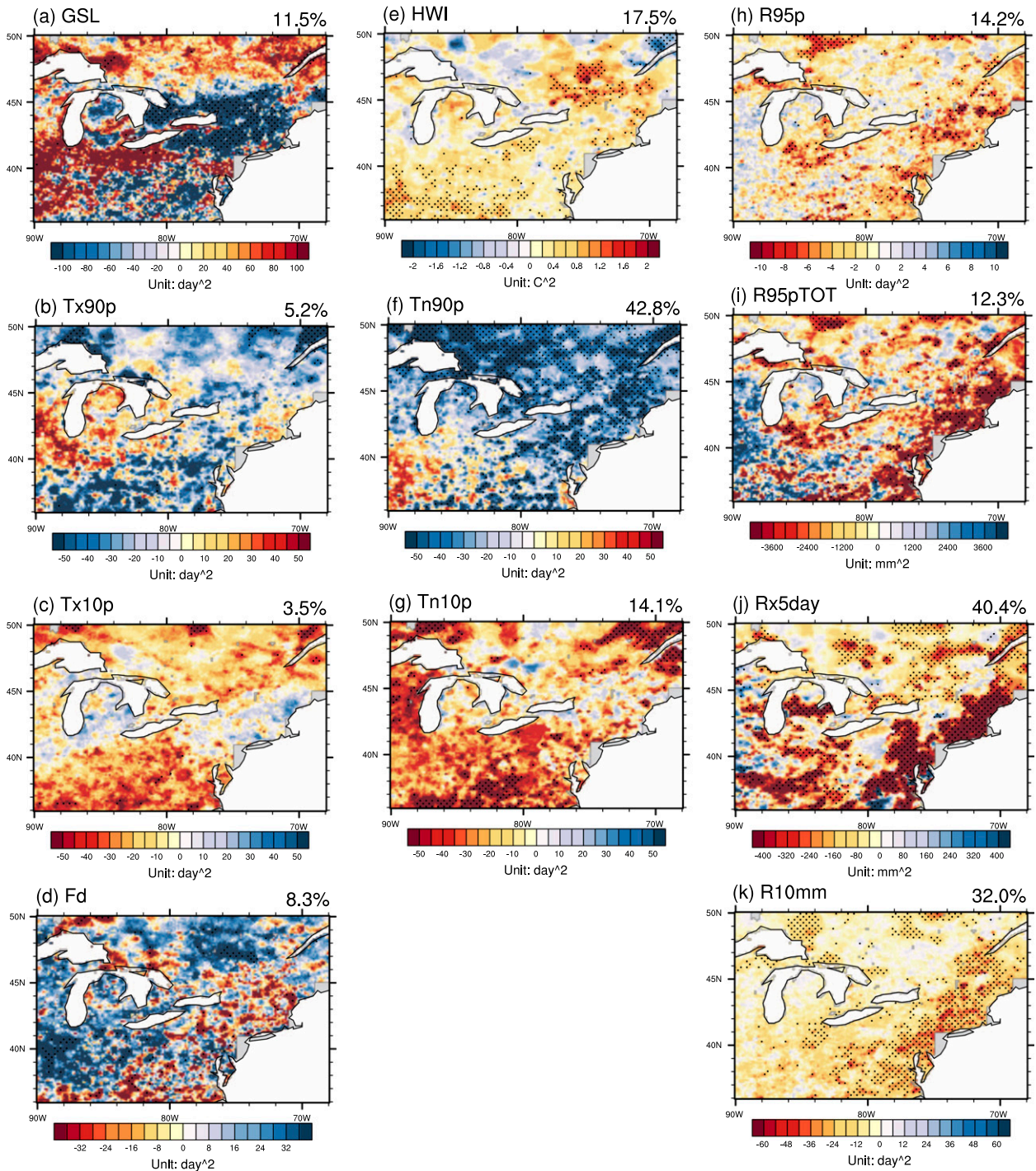
FIG. 3. The spatial distributions of differences between the downscaled and the observed variances of the 11 extreme indices for the period 1950–99. The number in the upper-right corner of each panel shows the percentage of grid points with significant differences. Stippling indicates changes significant at the $p = 0.05$ level based on an F test.

the percentages of grid points with significant differences are 11.4% and 21.4%. The bias distribution of Rx5day is not uniform, with some significant differences over the southern part and coastal region, and the

percentage of grid points with significant differences is 11.2%. For the R10mm, the downscaled results show uniform dry biases (~2 days) over the whole region, and 2.8% of total grid points have significant differences.

Generally, for these four precipitation extremes, BCSD does not perform as well as for temperature extremes.

The variances of the downscaled extreme indices are compared with the observation in Fig. 3. For GSL and FD, BCSD underestimates the variances toward the coastal region but overestimates the variances over the rest of the region, and the percentages of grid points with significant differences based on an F test (stippled region in Fig. 3) are 11.5% and 8.3%. BCSD overestimates the HWI variances over nearly the whole region, with 17.5% of total grid points showing significant differences. For Tx90p, Tx10p, Tn90p, and Tn10p, BCSD usually underestimates the variances over most of the region, and corresponding percentages are 5.2%, 3.5%, 42.8%, and 14.1%.

The variances of R95p are underestimated over the eastern part of the study region but overestimated over the western part of the study region. The corresponding percentage of grid points with significant differences is 14.2%. The spatial patterns of R95pTOT and Rx5day are similar, with significant negative differences along the coastal region and overestimation over a small area in the west, with overall percentages of grid points with significant differences of 12.3% and 40.4%. For R10mm, BCSD slightly underestimates the variances over most of the region, and the percentage of grid points with significant differences is 32.0%. From this figure, it can be concluded that the variances of the seven temperature extremes (except Tn90p and HWI) are usually better simulated by the downscaling method than those of the four precipitation extremes. Among the four precipitation extremes, the variances of the two extremes defined by absolute values (i.e., Rx5day and R10mm) are not reproduced as well as the other two, defined by the relative thresholds (i.e., R95p and R95pTOT).

The third test is the K–S test, which examines whether the CDFs of the downscaled extremes are significantly different from the observed CDFs, and the corresponding p values of the test are given in Fig. 4. Among the seven temperature extremes, GSL, Tx90p, Tx10p, Tn90p, and Tn10p usually have small differences (large p values) between the downscaled and observed CDFs over the whole domain with fewer than 10% of total grid points showing significant differences (stippled region in Fig. 4). For FD and HWI, the differences are larger, and about 20% and 30% of total grid points show significant differences. For R95p, Rx5day, and R10mm, the p values are usually smaller than 0.5, and the percentages for the first two extremes are about 5%, while the percentage for R10mm is larger, at 1.5%. The p values for R95pTOT are larger than those other three precipitation extremes, and 20.1% of total grid points show significant differences.

From these three figures, it can be concluded that BCSD performs differently for the three aspects of the 11 extreme indices, generally reproducing the character of temperature extremes better than precipitation extremes, consistent with conclusions from Bürger et al. 2012. Table 2 summarizes the percentages of grid points with significant differences for three tests over all extreme indices. Among the seven temperature indices, BCSD properly reproduces the means, variances, and CDFs on GSL, Tx90p, Tx10p, and Tn10p, with all percentages smaller than 15%, showing a relatively higher confidence level. For FD and HWI, BCSD does not perform as well as other temperature extremes in reproducing the means, with percentages of 53.1% and 29.6%, respectively. The downscaled variances of Tn90p are also significantly different from observations, over 42.8% of total grid points. This means the confidence in these three temperature indices is lower than for the previous four.

Among the four precipitation extremes, downscaling only shows high confidence for R95p, with all percentages smaller than 15% on three aspects. BCSD does not perform well in reproducing means of R95pTOT (21.4%) and variances of Rx5day (40.4%) and R10mm (32%), indicating lower confidence in these three extremes.

Figure 5 shows regionally averaged time series for GSL, the first of the 11 climate extremes indices. The observed time series is shown, along with time series from each individual downscaled GCM and the ensemble average. From the figure, it can be concluded that the downscaled ensemble regional average (average = 218.79 days and standard deviation = 9.99 days) is close to the observation (average = 217.05 days and standard deviation = 10.35 days) for both average and variance. Similar results for the remaining 10 climate extremes are shown in Figs. S1–S7 of the supplementary material.

b. Future changes of climate extremes

Figure 6 shows the spatial distributions of ensemble-averaged future changes of GSL under the three RCP scenarios (RCP2.6, RCP4.5, and RCP8.5). In this section, the future changes are defined by the difference between the future period average (2050–99) and the historical period average (1950–99). A Student's t test is employed for the significance test. The ensemble-averaged GSL will increase under all three scenarios, because the growing season will start earlier and end later as a result of the background of continuous warming, with larger magnitude increases in the higher emissions scenarios. Under the RCP2.6 scenario, the increases range from 10 days in the southern part of the domain to 30 days in the northern part of the U.S.
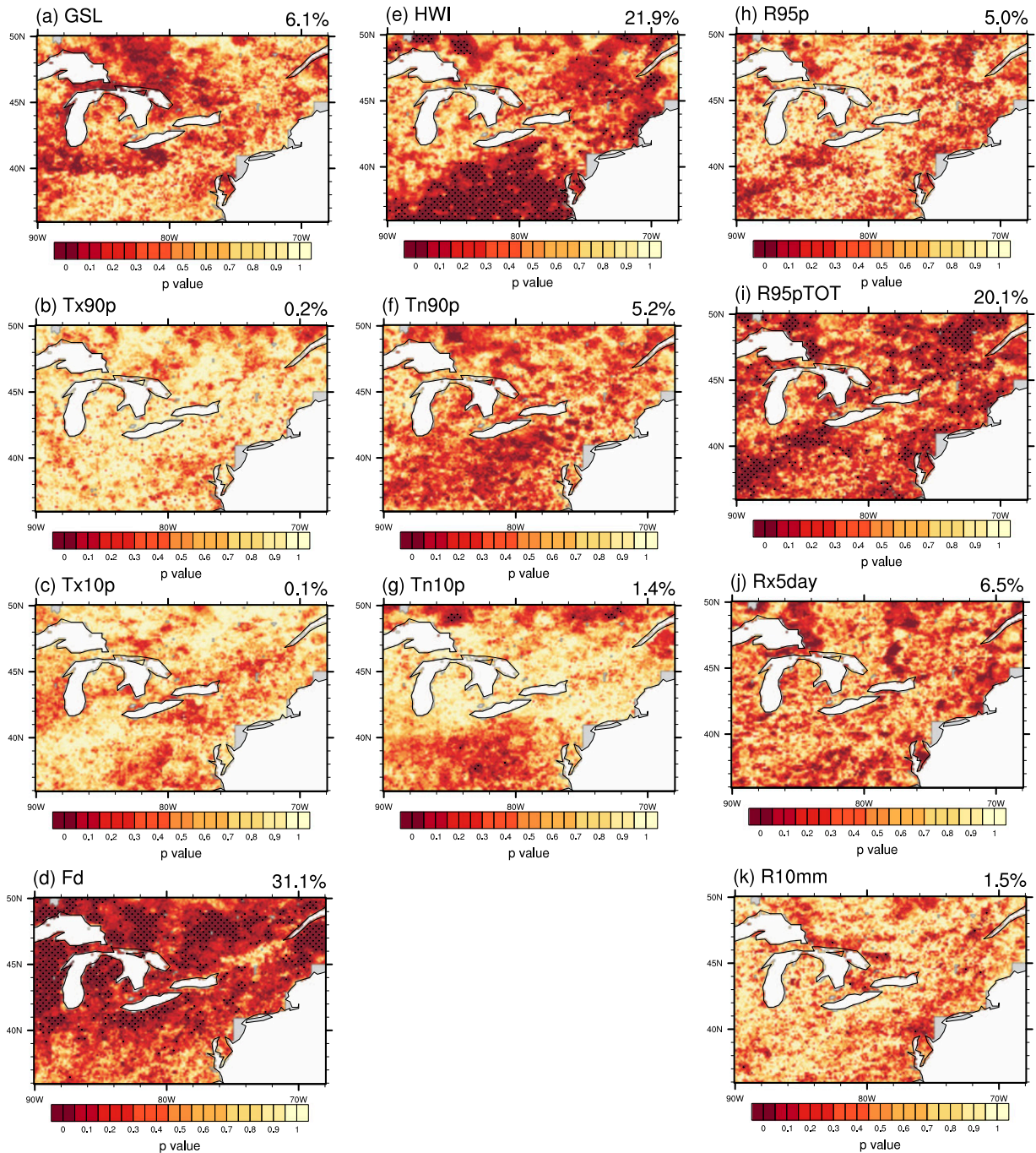
FIG. 4. The spatial distributions of differences between the downscaled and the observed CDFs of the 11 extreme indices for the period 1950–99. The number in the upper-right corner of each panel shows the percentage of grid points with significant differences. Stippling indicates changes significant at the $p = 0.05$ level based on a K–S test.

Northeast (Fig. 6a). These magnitudes increase to 20–40 days, with the largest increases around the Great Lakes under the RCP4.5 scenario (Fig. 6b). Under the RCP8.5 scenario, the increases are even higher, ranging from 30 days at the southern part to more than 50 days over the

northern part and Great Lakes (Fig. 6c). The increase in GSL is larger over the Appalachian Mountains (35°–40°N, 78°–82°W) than in the surrounding region, indicating that the impact of climate change on the agricultural growing season is enhanced over the region with

TABLE 2. Summary of percentages of the grid points with significant differences ($p = 0.05$) between observation and downscaled results over 11 climate extremes for the period 1950–99.

	Indices	t test	F test	K–S test
Temperature indices	GSL (days)	10.8%	11.5%	6.1%
	Tx90p (days)	0.2%	5.2%	0.2%
	Tx10p (days)	0.9%	3.5%	0.1%
	FD (days)	53.1%	8.3%	31.1%
	HWI (°C)	29.6%	17.5%	21.9%
	Tn90p (days)	4.9%	42.8%	5.2%
	Tn10p (days)	3.3%	14.1%	1.4%
Precipitation indices	R95p (days)	11.4%	14.2%	5.0%
	R95pTOT(mm)	21.4%	12.3%	20.1%
	Rx5day (mm)	11.2%	40.4%	6.5%
	R10mm (days)	2.8%	32.0%	1.5%

higher elevations (Diaz et al. 2014). All of these changes are significant at the 5% level.

Ensemble-averaged future changes in warm days (Tx90p) and cold days (Tx10p) are shown in Fig. 7. The threshold values are described in Table 1. Because of the warmer future climate, there are significant increases in the number of warm days and decreases in the number of cold days for all three emissions scenarios. Warm days increase by 10–30 days under the RCP2.6 scenario, with larger increases over the southern part of the domain, especially over the Ohio River valley (Fig. 7a). This spatial pattern is more obvious in the RCP4.5 scenario (Fig. 7c), with increases of 20 days over the northern part and 50–60 days over the southern part and Ohio River valley. Under the RCP8.5 scenario, there are over 70-day increases over the southern part of the domain, compared with about 50-day increases over the northern part (Fig. 7e). The spatial patterns of future changes in cold days (Figs. 7b,d,f) are similar to those of GSL, with larger magnitudes around the Great Lakes and smaller magnitudes over the southern part of the region. The decreases range from 6 to 18 days for the RCP2.6 scenario, from 9 to 24 days for the RCP4.5 scenario, and from more than 20 to 30 days for the RCP8.5 scenario. The smaller magnitudes of decreases in cold days compared with increases in warm days indicate that the lower tail of the daily maximum temperature distribution is not as sensitive to greenhouse gas increases as the upper tail, which exhibits a larger and more spatially varied response.

Two additional indices describing cold extremes (FD and Tn10p in Figs. S8 and S9b,d,f of the supplementary material, respectively) show ensemble-averaged future decreases with similar spatial patterns to those of cold days (Figs. 7b,d,f). Likewise, warm nights (Figs. S9a,c,e) show future increases with similar spatial patterns to warm days (Figs. 7a–c). Increases in the HWI (Fig. 8), however, show a spatial pattern that is unique from the
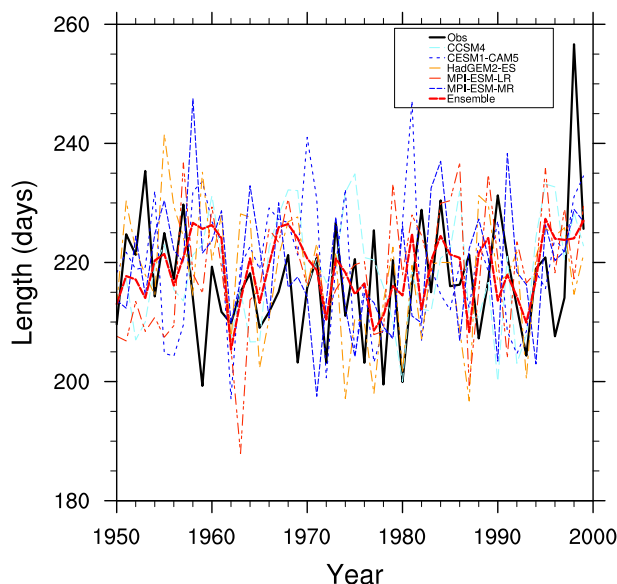

FIG. 5. The observed and downscaled time series of GSL (days) for period 1950–99.

other indices previously discussed, with more uniform increases over the whole region (1°C for RCP2.6, 3°C for RCP4.5, and 5°C for RCP8.5) and slightly higher increases of about 0.5°C over the Appalachian Mountains.

Because of increases in water vapor brought about by future warming (Collins et al. 2010), all four indices relevant to daily precipitation [i.e., R95p (Figs. 9a,c,e), R95pTOT (Figs. 9b,d,f), Rx5day (Fig. S10 in the supplemental material), and R10mm (Fig. S11 in the supplemental material)] show increases throughout the region because of higher specific humidity (Ning et al. 2012b), with magnitudes increasing with higher emissions scenarios. Within the four indices, only changes of the maximum total precipitation amount of 5 continuous days are not significant at the 5% level across the entire domain (Fig. S10). The physical mechanism underlying this relationship is that larger convective available potential energy (CAPE) will be induced by future larger specific humidity (Riemann-Campe et al. 2009), so stronger convection will occur as a result of the larger atmospheric instability, resulting in more extreme precipitation. These increases tend to be larger along the coast and smaller inland, consistent with previous studies (Ahmed et al. 2013). For example, for R95p and R95pTOT, the increases range from 1 to 3 days and from 20 to 30 mm for the RCP2.6 scenario (Figs. 9a,b), from 1 to 3 days and from 30 to 40 mm for RCP4.5 scenario (Figs. 9c,d), and from 2 to 5 days and 40 mm for the RCP8.5 scenario (Figs. 9e,f). All of these increases are significant at the 5% level. One potential reason for this spatial pattern is that water vapor increases are larger
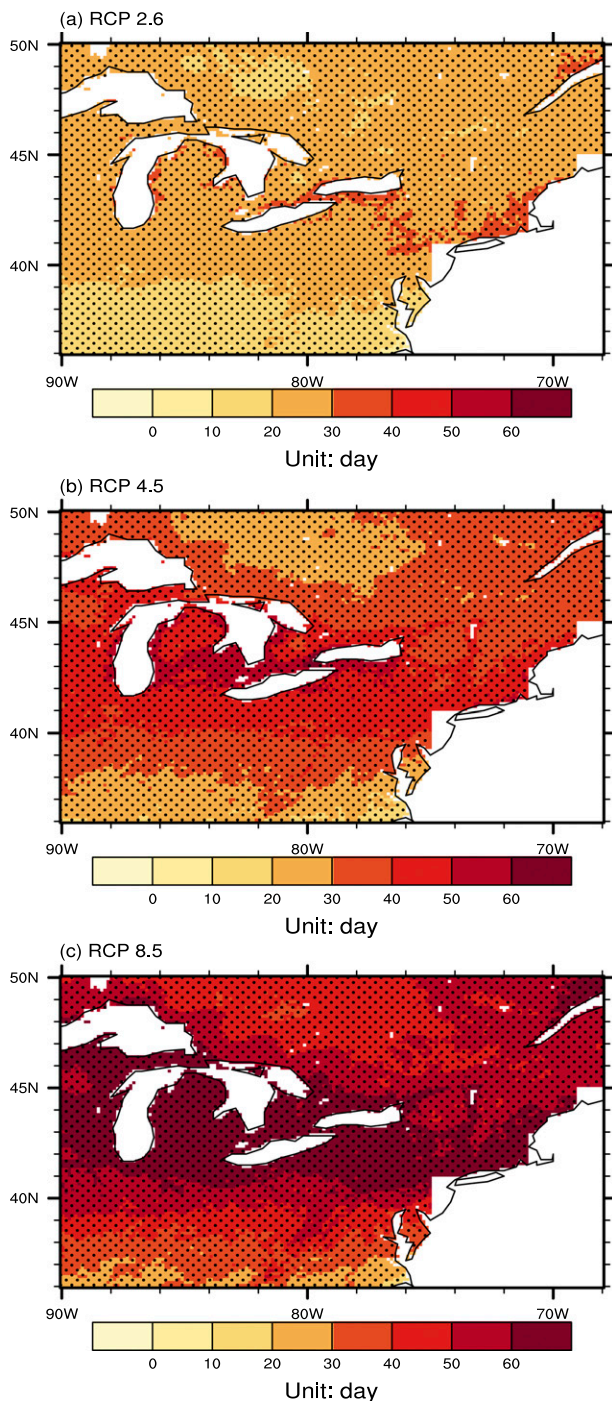
FIG. 6. The spatial distributions of future changes of GSL under the (a) RCP2.6, (b) RCP4.5, and (c) RCP8.5 emissions scenarios (days) for the period 2050–99 relative to period 1950–99. Stippling indicates changes significant at the $p = 0.05$ level.

be examined further in a future analysis. Another characteristic of the spatial distribution is that, for the two indices relevant to the precipitation intensity (i.e., R95pTOT and Rx5day), there are also significant increases over the Ohio River valley in addition to along the coastal region (Figs. 9b,d,f and Figs. S10a–c), especially for the higher emissions scenarios. We suggest that this indicates future warming over the Gulf of Mexico will lead to more water vapor being transported to the Ohio River valley to enhance precipitation intensity in that region, but further analysis of this hypothesis is needed.

From the spatial distributions of the ensemble-averaged future changes of the 11 climate extreme indices, it can be concluded that under future global warming, there will be more warm extremes and fewer cold extremes, and there will be more precipitation extremes, with larger intensities. For the temperature extremes, the magnitudes of the changes usually increase with the higher emissions scenarios. The magnitudes of the precipitation extreme changes also usually increase with the emissions scenario as a result of water vapor increases brought about by the warming. Moreover, the changes of temperature extremes are usually more spatially uniform than the precipitation extremes. This indicates that precipitation extremes are more highly dependent on subgrid processes, such as local convection. Comparing with the similar results based on raw GCM simulations (Sillmann et al. 2013), although these changes are also spatially homogeneous, they are based on influences of the large-scale warming on probability distribution functions (PDFs) over the high-resolution grid points, so some small-scale effects, such as topography, are taken into consideration in the downscaling procedure.

To characterize uncertainties associated with GCM differences, the multimodel ensemble spread is shown in Fig. 10 for regionally averaged changes in the 11 extremes indices and three emissions scenarios. For each of the 11 indices, the sign of the changes (either increasing or decreasing) is the same across all five GCMs and for all three emissions scenarios. For the temperature indices, with the exception of the HWI, the inter-GCM spread does not generally depend too strongly on the emissions scenario. The multimodel ensemble spread of the HWI is about 1.5°C under the RCP2.6 scenario and increases to about 3°C under the RCP8.5 scenario. The GSL and FD have model spreads with magnitudes of approximately 20 days. Warm days and warm nights have model spreads of approximately 25 days, larger than the ranges of cold days and cold nights, which are about only 7 days. Across the seven indices relevant to temperature, HadGEM2-ES always has the largest increasing/decreasing magnitudes, while MPI-ESM-MR

near the coast than they are inland because of warming of sea surface temperatures (SSTs). However, investigation of the corresponding mechanism is beyond the scope of this paper, so this potential mechanism will
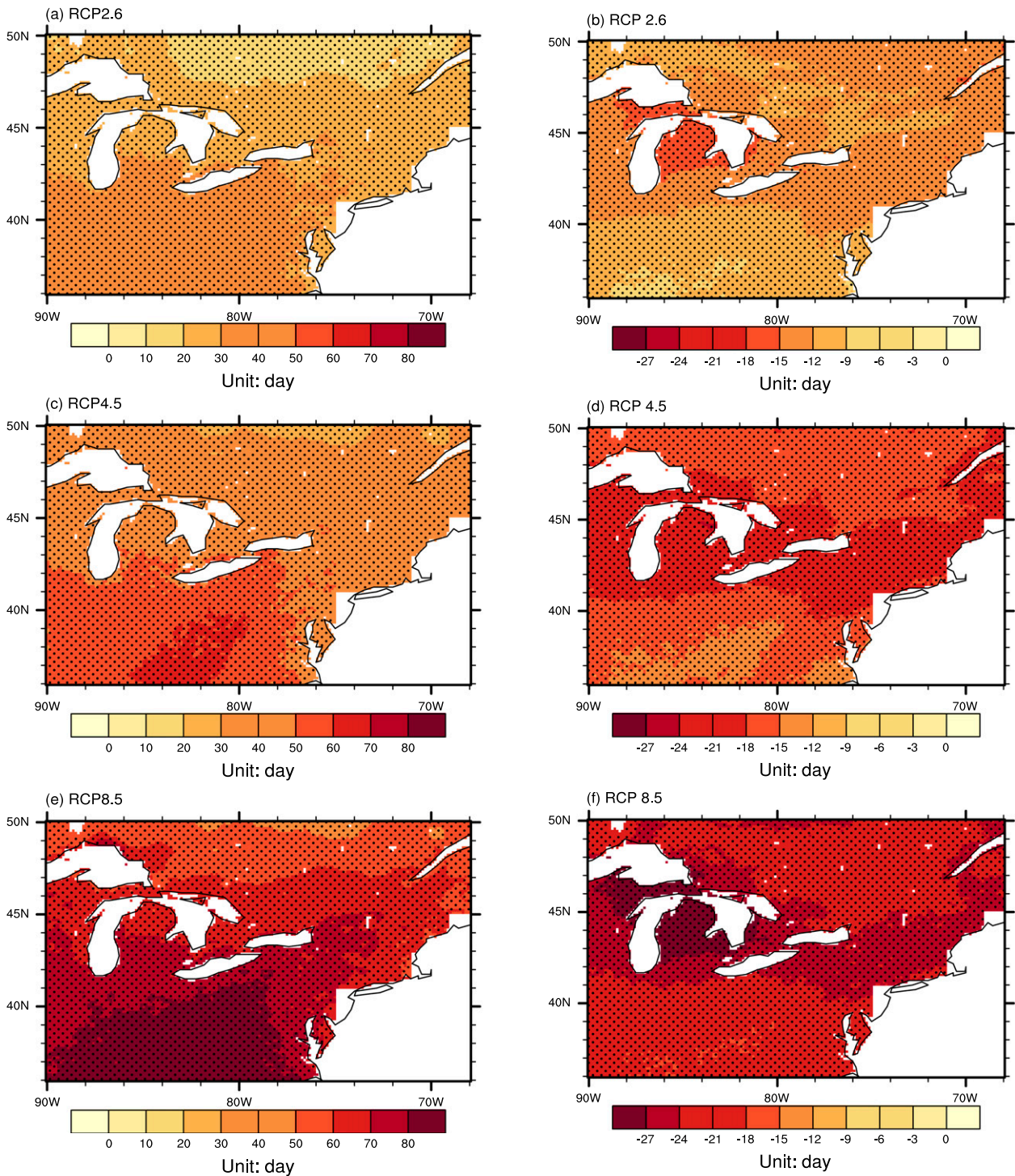
FIG. 7. The spatial distributions of future changes of (a),(c),(e) Tx90p and (b),(d),(f) Tx10p under the (a),(b) RCP2.6, (c),(d) RCP4.5, and (e),(f) RCP8.5 emissions scenarios (days) for the period 2050–99 relative to period 1950–99. Stippling indicates changes significant at the $p = 0.05$ level.
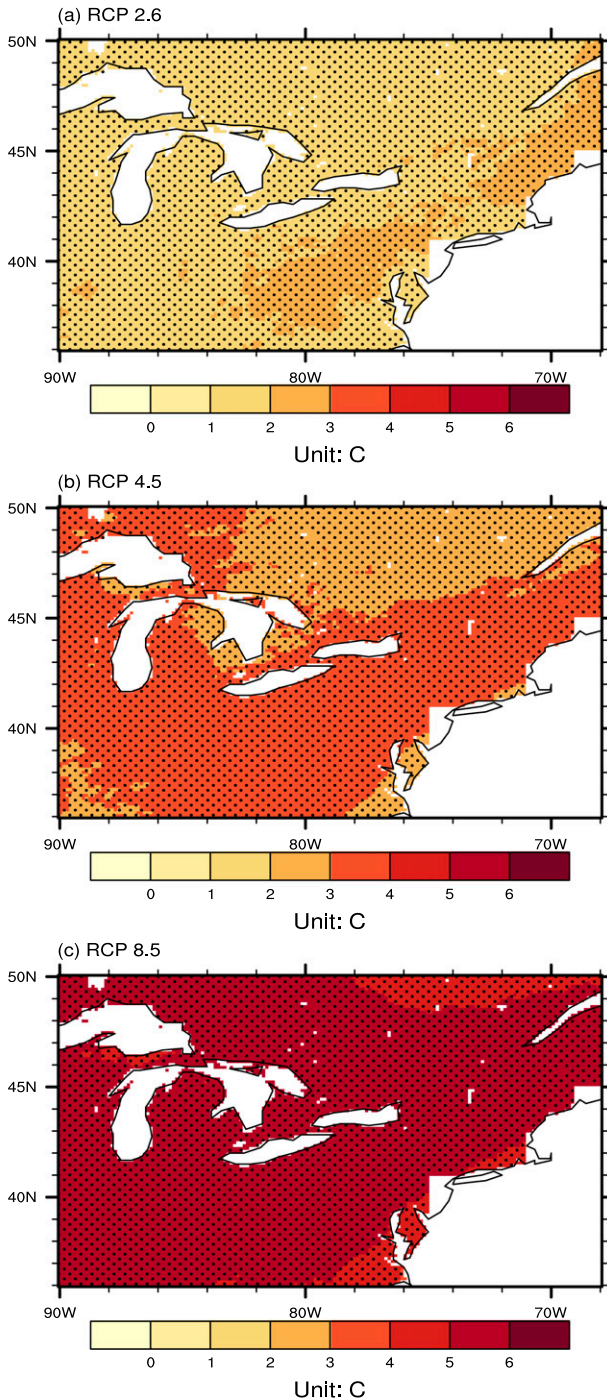
FIG. 8. The spatial distributions of future changes of HWI under the (a) RCP2.6, (b) RCP4.5, and (c) RCP8.5 emissions scenarios (°C) for the period 2050–99 relative to period 1950–99. Stippling indicates changes significant at the $p = 0.05$ level.

and MPI-ESM-LR always have the smallest increasing/ decreasing magnitudes, except for heat wave intensity and warm nights in RCP4.5 and RCP8.5, for which CCSM4 has the smallest increases.

For the four indices relevant to precipitation, multi-model ensemble spread depends on the emissions scenario, with the largest spread for the RCP8.5 scenario and the smallest spread for the RCP4.5 scenario. R95p and R95pTOT show the largest model spreads of about 2.5 days and 65 mm under the RCP8.5 scenario, followed by 2 days and 55 mm under the RCP2.6 scenario, and the smallest spreads of 1.5 days and 35 mm under the RCP4.5 scenario. For Rx5day and R10mm, the model spreads under RCP2.6 and RCP8.5 are similar, with magnitudes of 6 mm and 4 days, and the ranges under RCP4.5 are smaller, with magnitudes of 4.5 mm and 3 days. This indicates that these inter-GCM uncertainties do not change with the emission scenarios linearly. Among the five GCMs, CCSM4 always has the smallest increases of all four precipitation indices for three scenarios, and CESM1 (CAM5) and MPI-ESM-MR have the largest increases for the RCP2.6 and RCP8.5 scenarios (Fig. 10b), suggesting that the different responses may be due to different cloud microphysics parameterization schemes.

Although the magnitudes of the future changes increase significantly with the emissions scenarios, the inter-GCM uncertainties for the temperature-relevant indices do not change very much. In contrast, the inter-GCM uncertainties for the precipitation-relevant indices are more dependent on the emissions scenario, indicating that the GCMs' responses to the emissions scenarios on temperature extremes are more consistent than for precipitation extremes.

In general, uncertainties represented by the multi-model ensemble spread and uncertainties associated with different possible future RCP scenarios both appear to contribute a substantial portion to the overall uncertainty in future projections of the 11 climate extremes indices. For the temperature indices, emissions scenario–related uncertainties appear to dominate over the model-related uncertainties. In contrast, for the precipitation indices, model-related uncertainties appear to be as important as, if not more so, the uncertainties resulting from the different emissions scenarios. Other uncertainties not characterized in this plot include: 1) uncertainties resulting from lack of invariance in model biases for both model means and interannual variances (e.g., Maurer et al. 2013; C. Spence and C. Brown 2014, personal communication), 2) uncertainties resulting from changing relationships between monthly climate anomalies and the daily weather patterns, 3) uncertainties resulting from natural multidecadal variability not accurately represented in the model spread (e.g., Deser et al. 2012), and 4) uncertainties from sampling errors at the far tails of the temperature and precipitation distributions because of the limited historical record.
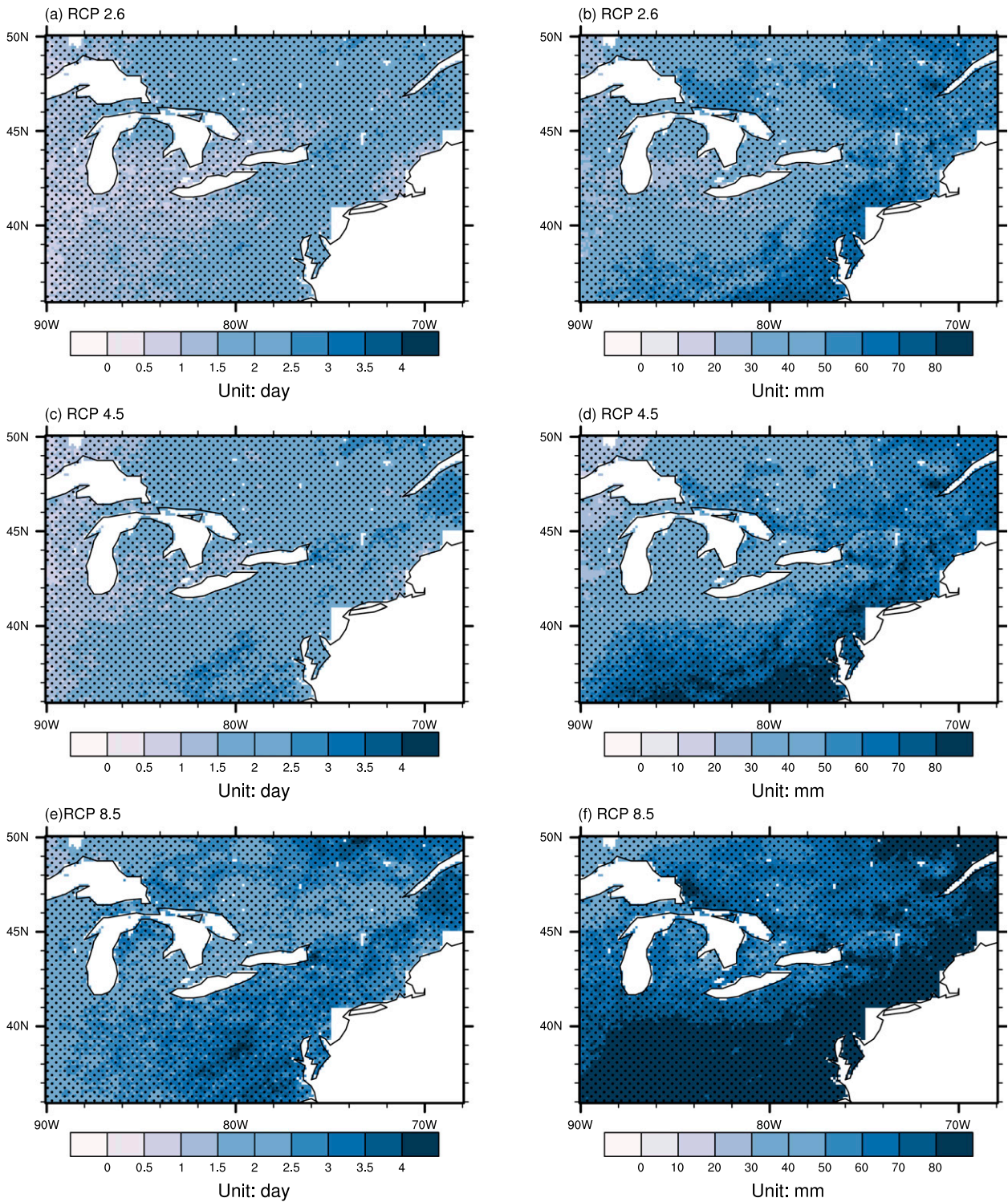
FIG. 9. The spatial distributions of future changes of (a),(c),(e) R95p (days) and (b),(d),(f) R95pTOT (mm) of extreme precipitation under the (a),(b) RCP2.6, (c),(d) RCP4.5, and (e),(f) RCP8.5 emissions scenarios for the period 2050–99 relative to period 1950–99. Stippling indicates changes significant at the $p = 0.05$ level.
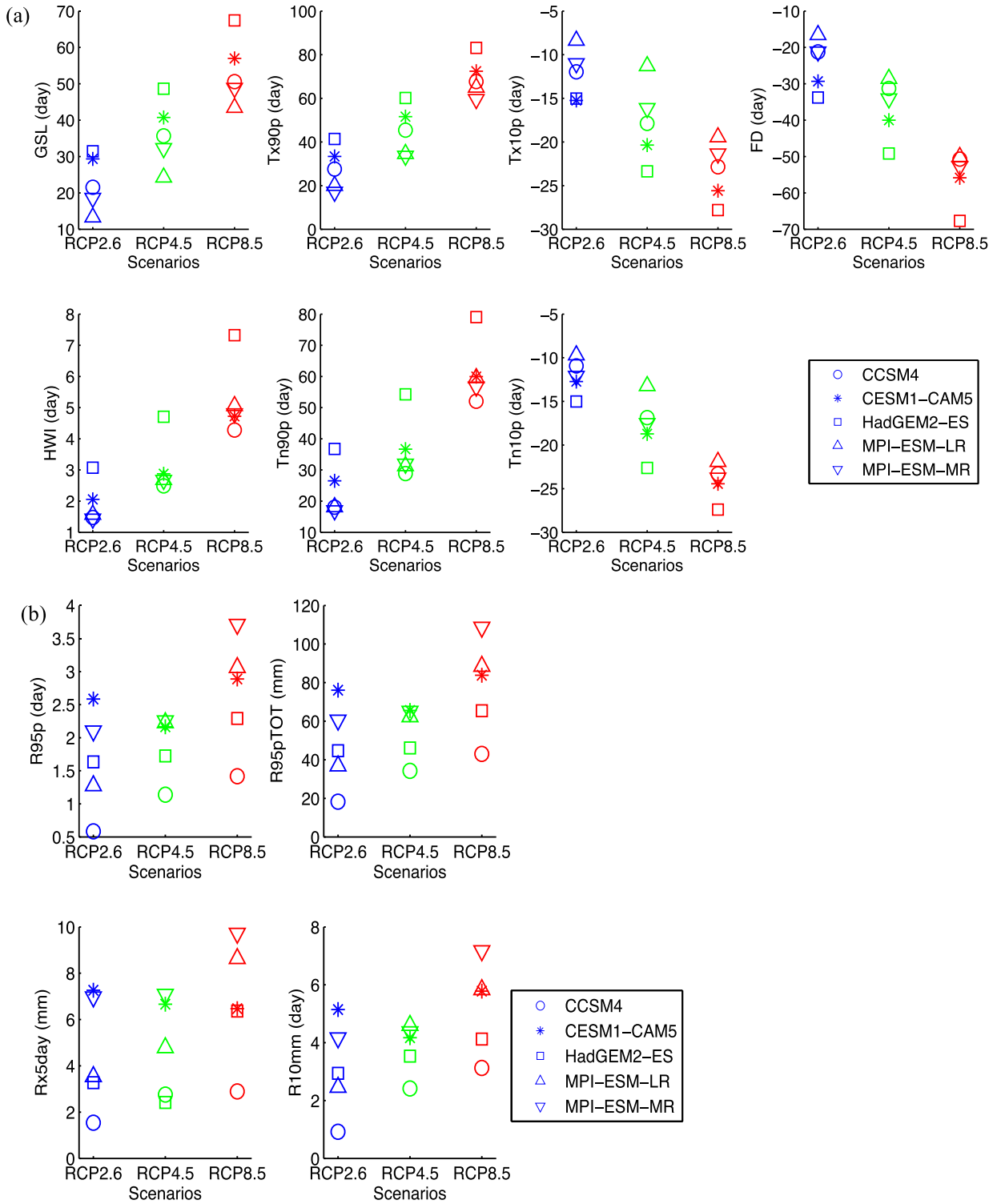
FIG. 10. The multimodel GCM ensemble spread for changes of the (a) temperature and (b) precipitation extreme indices under the RCP2.6 (blue), RCP4.5 (green), and RCP8.5 (red) scenarios for the period 2050–99 relative to period 1950–99.
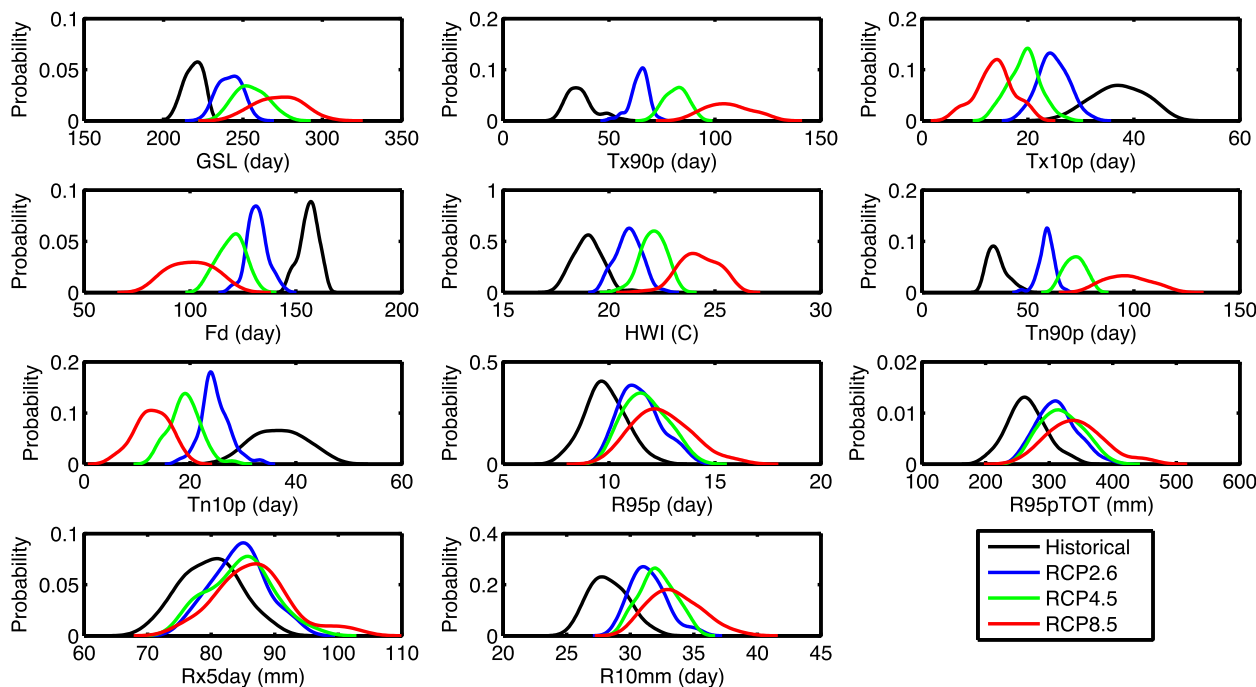
FIG. 11. Probability distributions representing interannual variability of the 11 extremes indices under the historical (black), RCP2.6 (blue), RCP4.5 (green), and RCP8.5 (red) scenarios.

Figure 11 shows probability distributions representing interannual variability of the indices over the 50-yr historical and 50-yr future simulations. The indices are averaged over the region and over the five GCMs. Since the downscaling method employs the observed daily probability distributions at the temporal disaggregation step, for the future downscaling, the GCM-simulated future changes of temperatures and precipitation will be transferred to generate totally new probability distributions of daily maximum temperature, minimum temperature, and precipitation, which are different from the historical period. Therefore, when examining the future changes of probability distributions on an interannual scale, the downscaling method can add considerable values to the projections of climate extremes by combining the GCM-predicted changes and observed probability distributions on a daily scale.

When comparing the probability distributions from the four scenarios (Fig. 11), obvious shifts can be seen in the distributions of the seven extreme indices relevant to the temperature. Five of seven temperature-relevant indices (i.e., GSL, Tx90p, FD, HWI, and Tn90p) have larger widths compared with the historical period, indicating not only shifts in the index magnitudes, but also larger interannual variability in the future. The remaining two indices (i.e., Tx10p and Tn90p) have smaller widths compared with the historical period,

showing reduced interannual variability in the future. In both cases, the interannual variability is dependent on the emissions scenario, with higher emissions scenarios causing larger variance shifts compared with the historical simulations.

Of particular note are the distributions of warm days and warm nights under the highest emissions scenario (RCP 8.5). Because of increases in the mean index values, together with a quite remarkable broadening of the interannual PDF, we see that extreme warm temperature days and nights are expected to occur 2.5–3 times as often during the most severe years in the future, as they do during severe years in the historical record. Even mild years in the future, under this emissions scenario, are expected to be more severe than the most severe years experienced in the past. The heat wave index shows a similar characteristic, with the highest emissions scenario leading to a future distribution that does not overlap with conditions that have occurred in the past.

For the four indices relevant to daily precipitation, the distribution shifts and shape changes of the PDFs are smaller than and not as significant as those of temperature, especially for the maximum precipitation amount of 5 consecutive days, consistent with previous results. The magnitudes of the distribution shifts increase with the emissions scenarios, but the differences between RCP2.6 (blue) and RCP4.5 (green) are small.

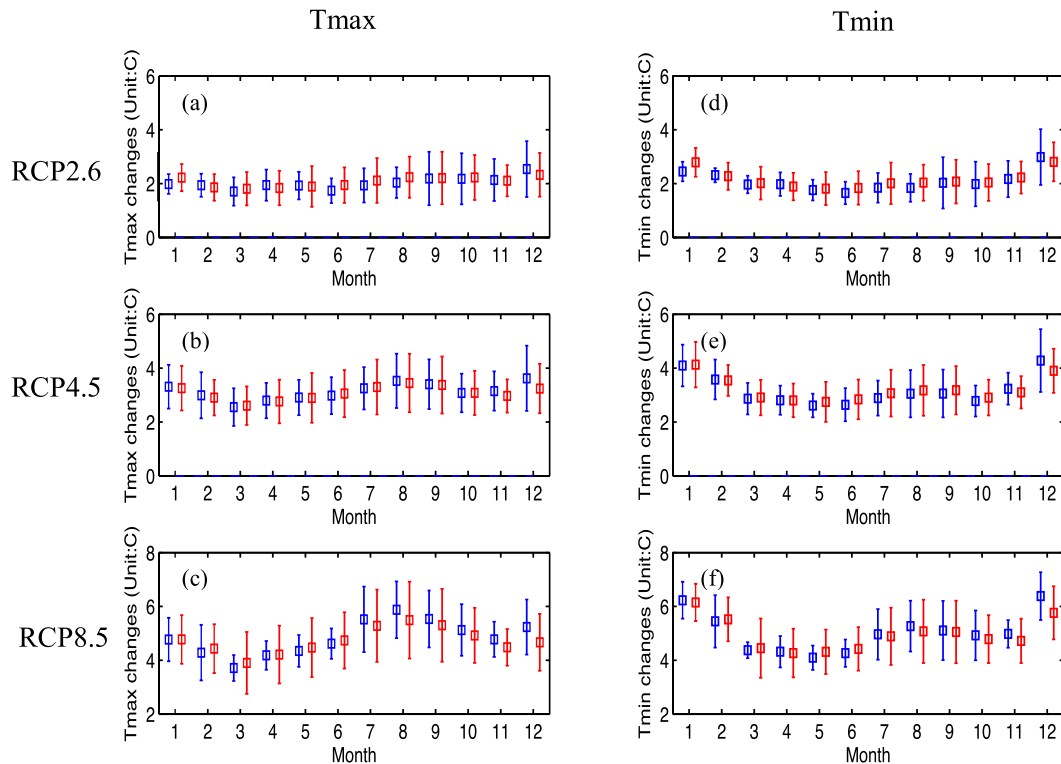Tmax   Tmin

FIG. 12. Comparison of the averages (boxes) and standard deviations (whiskers) of the mean future (a)–(c) maximum temperature and (d)–(f) minimum temperature changes for the 12 months based on the raw simulations of 5 GCMs (blue) and 15 GCMs (red) for the period 2050–99 relative to period 1950–99.

4. Discussion

Selecting global climate models for regional climate change studies through downscaling is a challenging question (Pierce et al. 2009). In this study, the future projections of climate extremes are based on the downscaled results based on five GCMs, which are only part of the CMIP5 data archive. To investigate how representative the warming from the five GCMs is, compared to the whole CMIP5 dataset, the ensemble-averaged future changes of maximum and minimum temperatures from the five GCMs are compared with a larger dataset containing 15 GCMs (Fig. 12). The comparison shows that changes from the five GCMs can really capture the changes over most months from the larger dataset, with small differences under all three emission scenarios. The spreads of the maximum and minimum temperature changes from the five GCMs are also comparable to results from the 15 GCMs, with similar magnitudes of standard deviations over different months. To quantitatively evaluate the differences between the five-GCM subsets and 15-GCM ensembles, a bootstrap with replacement (Efron 1982) is applied to the five-GCM subset for each month to generate 100 five-GCM subsets

through random sampling 100 times. Then, a t test is applied to the ensemble of the 100 five-GCM subsets and 15-GCM ensemble to investigate whether they are statistically different at $p = 0.05$ level. This procedure is repeated 10 times for each month, and frequency of significant differences is calculated. The results show that, for most months, the 100-member ensembles are not significantly different from the 15-GCM ensembles. The only four differences exist in maximum temperature for December under the RCP8.5 scenario ($^1/_{10}$), minimum temperature for January under the RCP2.6 scenario ($^7/_{10}$), December under the RCP4.5 scenario ($^1/_{10}$), and December under the RCP8.5 scenario ($^9/_{10}$), as shown in Fig. 12. Because the future changes of the climate extremes are highly related to the mean changes of maximum and minimum temperature based on the downscaling procedure, it is clear that the results from the five GCMs can capture most of the inter-GCM uncertainties in the CMIP5 archive.

One potential limitation of BCSD is that it applies historical daily maximum temperature, minimum temperature, and precipitation values but shifting/factoring the monthly means to the downscaled monthly means in the future downscaling. That means that BCSD assumes

that the variability of daily maximum and minimum temperatures and number of rainy days do not change in the future, since the historical daily variability is used in step 6 (Fig. 1). How this assumption influences the results is assessed through a comparison of changes of variability of daily maximum and minimum temperatures and number of rainy days from the raw GCM projection.

For maximum temperature (Figs. 13a–c), the future changes of variability defined as standard deviations of daily maximum temperature anomalies after removing the annual cycle show decreases over the northern part and increases over the southern part of the region, with magnitudes within ±0.3°C. The corresponding percentages are ±6% (Figs. S12a–c in the supplemental material). For minimum temperature (Figs. 13d–f), the future changes of variability of daily minimum temperature anomalies show negative values over most of the region, and the magnitudes decrease from about −0.25°C under the RCP2.6 scenario to about −0.6°C under the RCP8.5 scenario. The corresponding percentages range from −4% to −10% (Figs. S12d–f).

The future changes of numbers of rainy days show obvious differences over four seasons (Fig. 14), with mainly increases in winter, decreases in autumn, southern increases–northern decreases in summer, and decreases–increases with scenarios in spring. Usually, the magnitudes of the changes increase with higher RCP scenarios, and all the changes are within ±1.5 days, corresponding to ±10% (Fig. S13 in the supplemental material).

From these results, we can see that GCMs simulate slightly different changes of daily maximum and minimum temperature variability and precipitation frequency, which are inconsistent with the BCSD assumptions of fixed temperature PDF shapes and precipitation frequency. Although the changes are small, this inconsistency may still introduce some biases in the downscaled future projections of climate extremes. Therefore, readers are reminded here about this potential limitation of the downscaling method.

5. Conclusions

In the paper, historical and future changes of temperature and precipitation extremes over the northeastern United States are examined by applying the BCSD statistical downscaling method to five GCMs from the new CMIP5 archive projects under historical and three future emissions scenarios. The performances of the downscaling method in reproducing historical means, variances and CDFs of the observed 11 climate extremes are first quantitatively evaluated. Over the

historical period 1950–99, downscaling shows best performances for GSL, Tx90p, Tx10p, Tn10p, and R95p, followed by lower confidence level for FD, HWI, Tn90p, R95pTOT, Rx5day, and R10mm.

By the end of this century, the downscaling projects more warm temperature extremes, fewer cold temperature extremes, and more and intensified precipitation extremes under all three future emissions scenarios. When applying a Student's t test to these changes, the changes in all temperature extremes and most precipitation extremes, (except the maximum total precipitation amount of 5 continuous days), are significant at the 5% level across the entire domain. The spatial distribution of changes in temperature extremes is usually more uniform than for precipitation extremes, since the precipitation extremes are also influenced by other local factors. These changes (e.g., increases of growing season length, decreases of number of frost days, and increases of number of days with daily precipitation larger than 10 mm) are consistent with previous studies (e.g., Ahmed et al. 2013), indicating a shift toward higher daily temperature and larger daily precipitation in the future PDFs (Hayhoe et al. 2008).

The magnitudes of the changes increase with the emissions scenario and indicate that the temperature extremes are highly sensitive to future greenhouse gas emissions. Precipitation extremes are also sensitive to corresponding water vapor increases brought about by the warming. Although all five GCMs are consistent in showing future increases or decreases of the 11 indices, inter-GCM uncertainties still exist. The inter-GCM uncertainties of the temperature-relevant indices are not sensitive to the emissions scenarios, but the inter-GCM uncertainties of the precipitation-relevant indices are largest under the RCP8.5 scenario. The future probability distributions of temperature-relevant indices have obvious shifts and shape changes compared with the historical period, and the magnitudes increase with the emissions scenario. The shifts and shape changes in the probability distributions of precipitation indices are smaller than those of temperature.

These changes of probability distributions indicate not only more climate extreme events in the future, but also different interannual variability, requiring decision-makers to prepare adaptions and mitigations in anticipation of future changes of climate extreme events over this region. The uncertainties, originally from emission scenarios and GCM parameterization schemes, can be reduced by the downscaling method but cannot be totally removed (Ning et al. 2012b). Moreover, the downscaling method's potential deficiency of keeping the variability of daily maximum and minimum temperatures and precipitation frequency unchanged in the
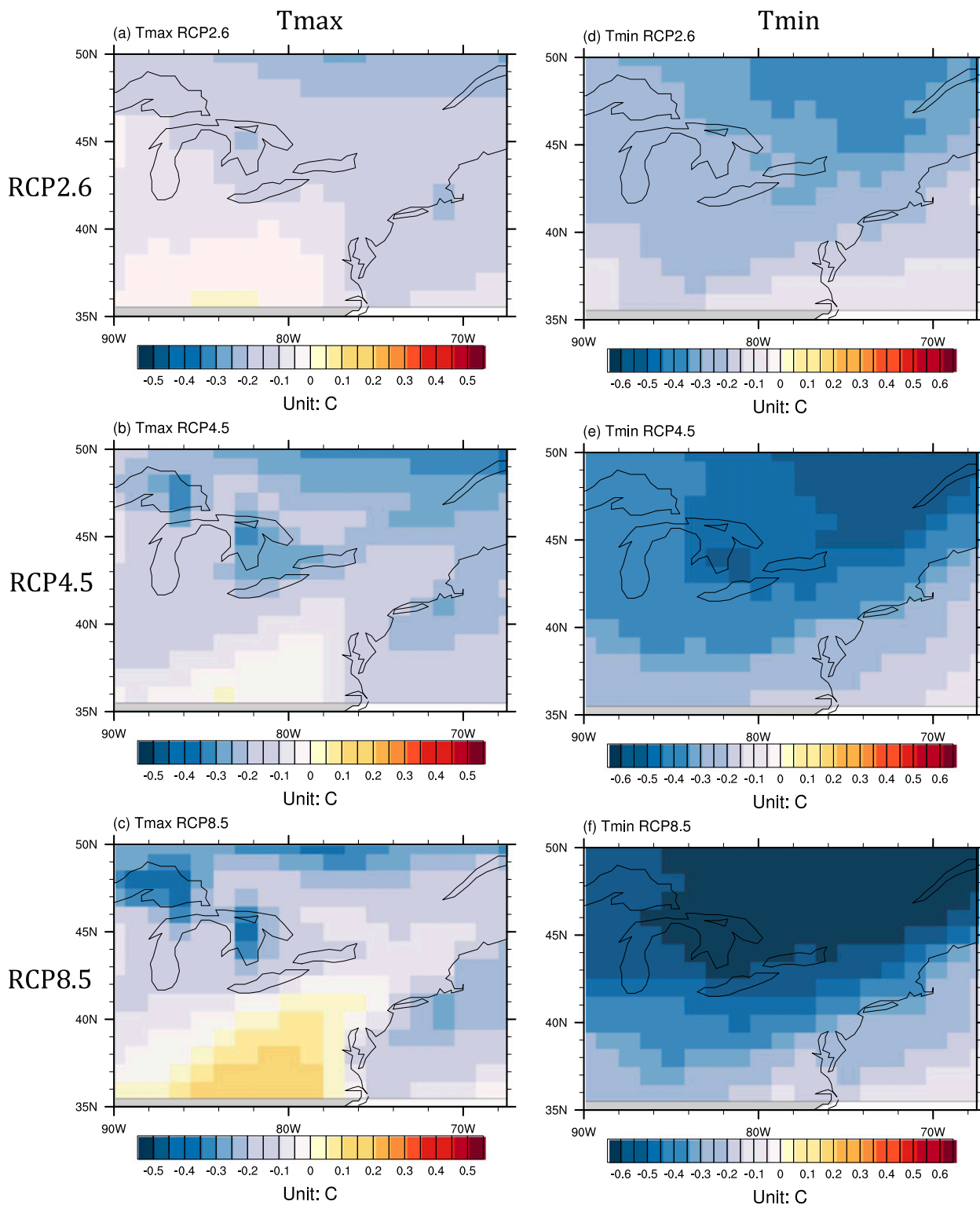
FIG. 13. Spatial distributions of future changes of variability of (a)–(c) daily maximum temperature and (d)–(f) minimum temperature anomalies (°C) from the raw simulations of the five GCMs for the period 2050–99 relative to period 1950–99.
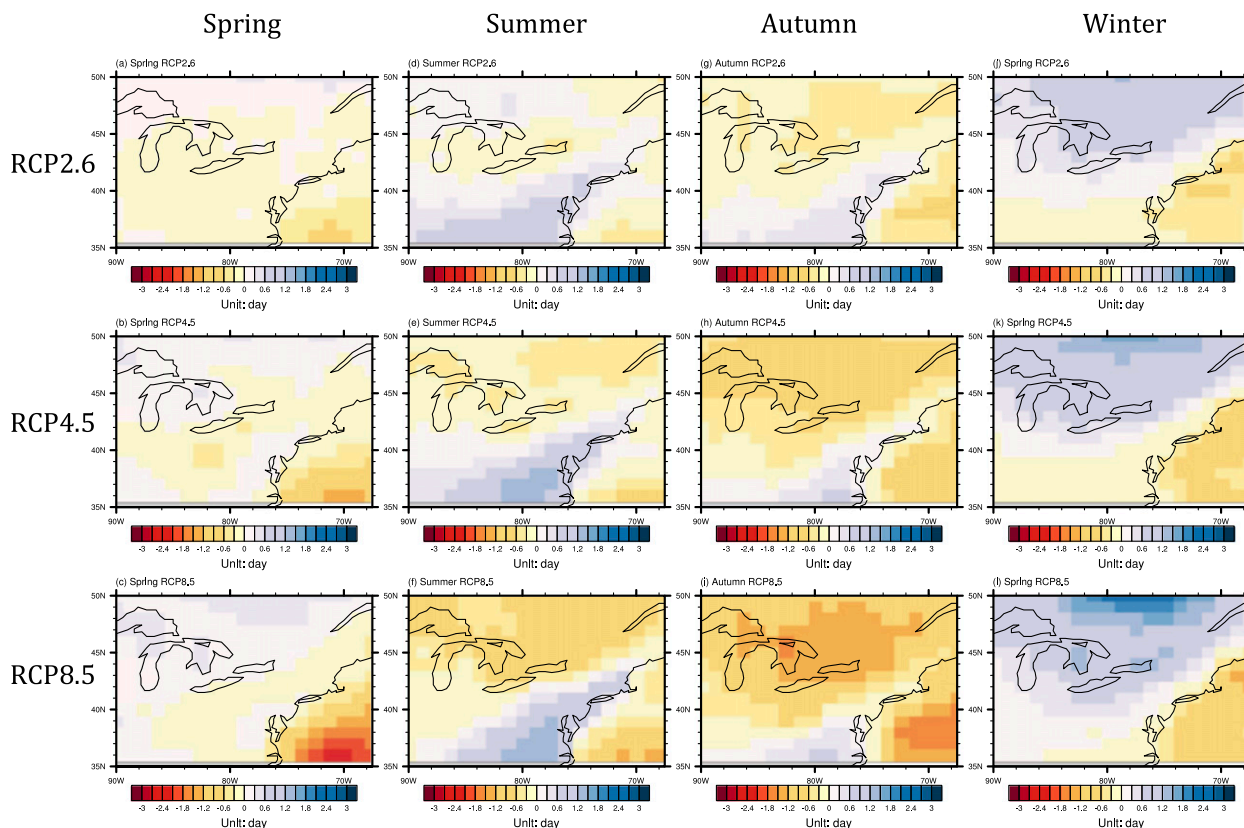
FIG. 14. Spatial distributions of future changes of numbers of rainy days over the four seasons (days) from the raw simulations of the five GCMs for the period 2050–99 relative to period 1950–99.

future may also introduce some new uncertainties, as BCSD may induce some different daily patterns not taken from the GCMs, although the changes of variability of daily maximum and minimum temperatures and numbers of rainy days are not large. Therefore, we would like to remind the readers to pay attention to the uncertainties resulting from the random errors induced by the downscaling method itself when they apply these probabilistic projections of climate extremes to interdisciplinary studies to investigate the influences of climate changes on regional water resources, ecosystems, forests, and agriculture.

REFERENCES

Ahmed, K. F., G. Wang, J. Silander, A. M. Wilson, J. M. Allen, R. Horton, and R. Anyah, 2013: Statistical downscaling and bias correction of climate model outputs for climate change impact assessment in the U.S. northeast. *Global Planet. Change,* **100,** 320–332, doi:10.1016/j.gloplacha.2012.11.003.

Alexander, L. V., and Coauthors, 2006: Global observed changes in daily climate extremes of temperature and precipitation. *J. Geophys. Res.,* **111,** D05109, doi:10.1029/2005JD006290.

Bürger, G., T. Q. Murdock, A. T. Werner, S. R. Sobie, and A. J. Cannon, 2012: Downscaling extremes—An intercomparison of multiple statistical methods for present climate. *J. Climate,* **25,** 4366–4388, doi:10.1175/JCLI-D-11-00408.1.

——, S. R. Sobie, A. J. Cannon, A. T. Werner, and T. Q. Murdock, 2013: Downscaling extremes: An intercomparison of multiple methods for future climate. *J. Climate,* **26,** 3429–3449, doi:10.1175/JCLI-D-12-00249.1.

Chen, M. D., D. Pollard, and E. J. Barron, 2003: Comparison of future climate change over North America simulated by two

regional models. *J. Geophys. Res.,* **108,** 4348, doi:10.1029/2002JD002738.

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. [Available online at https://www.ipcc-wg1.unibe.ch/publications/wg1-ar4/ar4-wg1-chapter11.pdf.]

Collins, M., and Coauthors, 2010: The impact of global warming on the tropical Pacific Ocean and El Niño. *Nat. Geosci.,* **3,** 391–397, doi:10.1038/ngeo868.

Corell, R. W., D. Liverman, K. Dow, K. L. Ebi, K. Kunkel, L. O. Mearns, and J. Melillo, 2014: Research needs for climate and global change assessments. *Climate Change Impacts in the United States: The Third National Climate Assessment,* J. M. Melillo, Terese (T. C.) Richmond, and G. W. Yohe, Eds., U.S. Global Change Research Program, 707–718, doi:10.7930/J03R0QR3.

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

Diaz, H. F., R. S. Bradley, and L. Ning, 2014: Climatic changes in mountain regions of the American Cordillera and the tropics: Historical changes and future outlook. *Arct. Antarct. Alp. Res.,* **46,** 735–743.

Diaz-Nieto, J., and R. L. Wilby, 2005: A comparison of statistical downscaling and climate change factor methods: Impacts on low flows in the River Thames, United Kingdom. *Climatic Change,* **69,** 245–268, doi:10.1007/s10584-005-1157-6.

Easterling, D. R., G. A. Meehl, C. Parmesan, S. A. Changnon, T. R. Karl, and L. O. Mearns, 2000: Climate extremes: Observations, modeling, and impacts. *Science,* **289,** 2068–2074, doi:10.1126/science.289.5487.2068.

Efron, B., 1982: The jackknife, the bootstrap, and other resampling plans. *CBMS-NSF Regional Conference Series in Applied Mathematics,* No. 38, Society for Industrial and Applied Mathematics, 92 pp.

Fowler, H. J., S. Blenkinsop, and C. Tebaldi, 2007: Linking climate change modeling to impacts studies: Recent advances in downscaling techniques for hydrological modelling. *Int. J. Climatol.,* **27,** 1547–1578, doi:10.1002/joc.1556.

Frich, P., L. V. Alexander, P. Della-Marta, B. Gleason, M. Haylock, A. M. G. Klein Tank, and T. Peterson, 2002: Observed coherent changes in climatic extremes during the second half of the twentieth century. *Climate Res.,* **19,** 193–212, doi:10.3354/cr019193.

Goodess, C. M., and Coauthors, 2005: STARDEX downscaling climate extremes. STARDEX Tech. Rep., 21 pp. [Available online at http://www.cru.uea.ac.uk/projects/stardex/reports/STARDEX_FINAL_REPORT.pdf.]

Harding, B. L., A. W. Wood, and J. R. Prairie, 2012: The implications of climate change scenario selection for future streamflow projection in the Upper Colorado River Basin. *Hydrol. Earth Syst. Sci.,* **16,** 3989–4007, doi:10.5194/hess-16-3989-2012.

Hawkins, E., and R. Sutton, 2009: The potential to narrow uncertainty in regional climate predictions. *Bull. Amer. Meteor. Soc.,* **90,** 1095–1107, doi:10.1175/2009BAMS2607.1.

Hayhoe, K., and Coauthors, 2008: Regional climate change projections for the northeast USA. *Mitigation Adapt. Strategies Global Change,* **13,** 425–436, doi:10.1007/s11027-007-9133-2.

Haylock, M. R., G. C. Cawley, C. Harpham, R. L. Wilby, and C. M. Goodess, 2006: Downscaling heavy precipitation over the United Kingdom: A comparison of dynamical and statistical methods and their future scenarios. *Int. J. Climatol.,* **26,** 1397–1415, doi:10.1002/joc.1318.

Hewitson, B. C., and R. G. Crane, 2006: Consensus between GCM climate change projections with empirical downscaling: Precipitation downscaling over South Africa. *Int. J. Climatol.,* **26,** 1315–1337, doi:10.1002/joc.1314.

Horton, R., and Coauthors, 2014: Northeast. *Climate Change Impacts in the United States: The Third National Climate Assessment,* J. M. Melillo, Terese (T.C.) Richmond, and G. W. Yohe, Eds., U.S. Global Change Research Program, 371–395, doi:10.7930/JOSF2T3P.

Maraun, D., and Coauthors, 2010: Precipitation downscaling under climate change: Recent development to bridge the gap between dynamical models and the end user. *Rev. Geophys.,* **48,** RG3003, doi:10.1029/2009RG000314.

Maurer, E. P., 2007: Uncertainty in hydrologic impacts of climate change in the Sierra Nevada, California, under two emissions scenarios. *Climatic Change,* **82,** 309–325, doi:10.1007/s10584-006-9180-9.

——, A. W. Wood, J. C. Adam, D. P. Lettenmaier, and B. Nijssen, 2002: A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States. *J. Climate,* **15,** 3237–3251, doi:10.1175/1520-0442(2002)015<3237:ALTHBD>2.0.CO;2.

——, T. Das, and D. R. Cayan, 2013: Errors in climate model daily precipitation and temperature output: Time invariance and implications for bias correction. *Hydrol. Earth Syst. Sci.,* **17,** 2147–2159, doi:10.5194/hess-17-2147-2013.

Meehl, G. A., and C. Tebaldi, 2004: More intense, more frequent, and longer lasting heat waves in the 21st century. *Science,* **305,** 994–997, doi:10.1126/science.1098704.

——, and Coauthors, 2007: Global climate projections. *Climate Change 2007: The Physical Science Basis,* S. Solomon et al., Eds., Cambridge University Press, 747–846. [Available online at http://www.ipcc.ch/pdf/assessment-report/ar4/wg1/ar4-wg1-chapter10.pdf.]

Melillo, J. M., T. C. Richmond, and G. W. Yohe, Eds., 2014: *Climate Change Impacts in the United States: The Third National Climate Assessment.* U.S. Global Change Research Program, 841 pp., doi:10.7930/J0Z31WJ2.

Moss, R. H., and Coauthors, 2010: The next generation of scenarios for climate change research and assessment. *Nature,* **463,** 747–756, doi:10.1038/nature08823.

Negri, D. H., N. R. Gollehon, and M. P. Aillery, 2005: The effects of climatic variability on US irrigation adoption. *Climatic Change,* **69,** 299–323, doi:10.1007/s10584-005-1817-6.

Ning, L., and Y. Qian, 2009: Interdecadal change in extreme precipitation over South China and its mechanism. *Adv. Atmos. Sci.,* **26,** 109–118, doi:10.1007/s00376-009-0109-x.

——, and R. S. Bradley, 2014: Winter precipitation variability and corresponding teleconnections over the northeastern United States. *J. Geophys. Res. Atmos.,* **119,** 7931–7945, doi:10.1002/2014JD021591.

——, and ——, 2015: Winter climate extremes over the northeastern United States and southeastern Canada and teleconnections with large-scale modes of climate variability. *J. Climate,* **28,** 2475–2493, doi:10.1175/JCLI-D-13-00750.1.

——, M. E. Mann, R. Crane, and T. Wagener, 2012a: Probabilistic projections of climate change for the mid-Atlantic region of the United States: Validation of precipitation downscaling during the historical era. *J. Climate,* **25,** 509–526, doi:10.1175/2011JCLI4091.1.

——, ——, ——, ——, R. G. Najjar Jr., and R. Singh, 2012b: Probabilistic projections of anthropogenic climate change impacts on precipitation for the mid-Atlantic region of the

United States. *J. Climate,* **25,** 5273–5291, doi:10.1175/JCLI-D-11-00565.1.

Palmer, R., M. Cook, R. Horton, L. Deegan, A. W. D'Amato, F. R. Thompson III, and L. Gilbert, 2011: A proposal to the Department of the Interior to establish the Northeast Climate Science Center (NE-CSC). [Available online at https://necsc.umass.edu/sites/default/files/NECSC-proposal-May2011.pdf.]

Pierce, D. W., T. P. Barnett, B. D. Santer, and P. J. Gleckler, 2009: Selecting global climate models for regional climate change studies. *Proc. Natl. Acad. Sci. USA,* **106,** 8441–8446, doi:10.1073/pnas.0900094106.

Plummer, D. A., and Coauthors, 2006: Climate and climate change over North America as simulated by the Canadian RCM. *J. Climate,* **19,** 3112–3132, doi:10.1175/JCLI3769.1.

Prudhomme, C., N. Reynard, and S. Crooks, 2002: Downcaling of global climate models for flood frequency analysis: Where are we now? *Hydrol. Processes,* **16,** 1137–1150, doi:10.1002/hyp.1054.

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. [Available online at http://www.ipcc.ch/pdf/assessment-report/ar4/wg1/ar4-wg1-chapter8.pdf.]

Riemann-Campe, K., K. Fraedrich, and F. Lunkeit, 2009: Global climatology of Convective Available Potential Energy (CAPE) and Convective Inhibition (CIN) in ERA-40 reanalysis. *Atmos. Res.,* **93,** 534–545, doi:10.1016/j.atmosres.2008.09.037.

Schmidli, J., C. M. Goodess, C. Frei, M. R. Haylock, Y. Hundecha, J. Ribalaygua, and T. Schmith, 2007: Statistical and dynamical downscaling of precipitation: An evaluation and comparison of scenarios for the European Alps. *J. Geophys. Res.,* **112,** D04105, doi:10.1029/2005JD007026.

Sillmann, J., V. V. Kharin, X. Zhang, F. W. Zwiers, and D. Bronaugh, 2013: Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate. *J. Geophys. Res. Atmos.,* **118,** 1716–1733, doi:10.1002/jgrd.50203.

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, doi:10.1175/BAMS-D-11-00094.1.

Trenberth, K. E., and Coauthors, 2007: Observations: Surface and atmospheric climate change. *Climate Change 2007: The Physical Science Basis,* S. Solomon et al., Eds., Cambridge University Press, 235–336. [Available online at http://www.ipcc.ch/pdf/assessment-report/ar4/wg1/ar4-wg1-chapter3.pdf.]

van Vuuren, D. P., and Coauthors, 2011: The representative concentration pathways: An overview. *Climatic Change,* **109,** 5–31, doi:10.1007/s10584-011-0148-z.

Walsh, J., and Coauthors, 2014: Climate science supplement. *Climate Change Impacts in the United States: The Third National Climate Assessment,* J. M. Melillo, Terese (T.C.) Richmond, and G. W. Yohe, Eds., U.S. Global Change Research Program, 371–395, doi:10.7930/JOKS6PHH.

Wilby, R. L., and T. M. L. Wigley, 1997: Downscaling general circulation model output: A review of methods and limitations. *Prog. Phys. Geogr.,* **21,** 530–548, doi:10.1177/030913339702100403.

Wood, A. W., E. P. Maurer, A. Kumar, and D. P. Lettenmaier, 2002: Long-range experimental hydrologic forecasting for the eastern United States. *J. Geophys. Res.,* **107,** 4429, doi:10.1029/2001JD000659.

——, L. R. Leung, V. Sridhar, and D. P. Lettenmaier, 2004: Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic Change,* **62,** 189–216, doi:10.1023/B:CLIM.0000013685.99609.9e.