AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Atmospheres

RESEARCH ARTICLE

10.1002/2015JD023073

Key Points:
- Current simulations and future projections over the Northeast U.S. are analyzed
- The GCM-RCM pairs may produce decreased combined biases compared to the driving GCMs
- Summer precipitation changes in GCM-RCM pairs are inconsistent with GCMs

Supporting Information:
- Texts S1–S5, Figures S1–S4, and Tables S1–S7

Correspondence to:
F. Fan,
fanfangxing@mail.iap.ac.cn

Citation:
Fan, F., R. S. Bradley, and M. A. Rawlins (2015), Climate change in the Northeast United States: An analysis of the NARCCAP multimodel simulations, *J. Geophys. Res. Atmos.*, *120*, doi:10.1002/2015JD023073.

Received 12 JAN 2015
Accepted 13 SEP 2015
Accepted article online 16 SEP 2015

Climate change in the Northeast United States: An analysis of the NARCCAP multimodel simulations

Fangxing Fan[1], Raymond S. Bradley[2], and Michael A. Rawlins[2]

[1]International Center for Climate and Environment Sciences, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China, [2]Department of Geosciences, University of Massachusetts, Amherst, Massachusetts, USA

Abstract Based on the outputs from multiple regional climate models (RCMs) forced by global climate models (GCMs) in phase II of the North American Regional Climate Change Assessment Program, current simulations and future projections of surface air temperature and precipitation over the Northeast U.S. region are analyzed. We address two questions in this study: How do the combined biases from the GCM-RCM pairs compare to that from their driving GCMs? Are the future responses projected by these GCM-RCM pairs robust and consistent with that predicted by their driving GCMs? The GCM-RCM pair may produce decreased combined biases compared to the inherent biases in the driving GCM, but the possibility of compensating errors suggest that this does not necessarily imply improved representations of current climate. An RCM simulation may also impart biases which exacerbate those from the driving GCM, leading to a large combined error. The substantial underestimates of winter and summer temperature produced by Geophysical Fluid Dynamics Laboratory (GFDL) Experimental Climate Prediction-2 and GFDL Regional Climate Model version 3 are partly attributed to moderate cold biases from their driving GFDL, whereas dynamically downscaling Hadley Centre Climate Model version 3 with Hadley Centre Regional Climate Model leads to mitigated combined biases. The "probability density functions" of temperature and precipitation are estimated for each individual model pair to illustrate the temporal and spatial distributions of these two climate variables. Under the Special Report on Emissions Scenarios A2 emission scenario, the ensemble averaged changes in winter and summer temperature for the mid-21st century (2041–2070) vary between 2.5 to 3.2°C across the subregions. These warming signals are consistent and statistically significant across the model pairs and far exceed the estimated natural variability. The projected future changes in precipitation indicate generally wetter winters and drier summers, but the magnitudes, directions, and spatial distributions of precipitation changes are model-dependent. Moreover, the ensemble average summer precipitation changes (0.6 to −7.9% as estimated by percentage of present-day values) fall within the estimated range of natural variability. Different changes in moisture flux convergence-divergence appear to contribute to the disagreement of the precipitation responses between some GCM-RCM pairs and their driving GCMs. The "reliability ensemble averaging" procedure is also applied and provides a complement to a simpler averaging method to estimate the average and uncertainty range of the simulated climate changes.

1. Introduction

Although climate models have reached a remarkable level of maturity and can reproduce many aspects of the climate system reasonably well [*Randall et al.*, 2007], it has proven difficult to reduce the model spread in projected future climate [*Knutti*, 2008]. Uncertainties in future climate projections arise mainly from (i) imperfect representations of key processes and feedback in the climate system by various climate models ("model uncertainty" or "science uncertainty") [*Meehl et al.*, 2007]; (ii) different estimates of future emissions of greenhouse gases and aerosols based on the assumptions regarding population, socio-economic development, and technological change ("scenario uncertainty"); and (iii) internal climate variability on interannual and decadal timescales, which may strengthen or offset future climate response to anthropogenic forcing ("natural variability uncertainty") [*Deser et al.*, 2012]. Downscaling projected future climate to the regional scale carries errors in the driving atmosphere-ocean general circulation model (AOGCM) to regional climate projections and introduces another uncertainty related to the application of different downscaling methods (i.e., statistical and dynamical downscaling) [*Pan et al.*, 2001; *Christensen et al.*, 2007]. However, while coarse resolution coupled AOGCMs may succeed in capturing large-scale climate features, finer-resolution regional climate models (RCMs) show better

capabilities in resolving climate change over regions of complex terrain, in simulating extreme events, and in representing mesoscale weather processes [*Jones et al*., 2004].

Although the purpose of embedding RCMs within GCMs is to provide finer-scale and hopefully more credible details over a domain of interest, biases in RCM simulations derive both from the lateral boundary forcing and the model formulation [*Noguer et al*., 1998]. On the one hand, the RCM simulations are consistent with their driving GCMs on large scales and consequently inherit the large-scale biases of their driving models. On the other, the RCMs can directly simulate the dynamics of the regional climate and generate additional biases depending on different representations of physical processes. Uncertainty assessments are critical in the evaluation of regional climate simulations produced through different dynamical downscaling (i.e., applying different sophisticated RCMs to simulate the dynamics of the regional climate). Phase I of the North American Regional Climate Change Assessment Program (NARCCAP) [*Mearns et al*., 2007, 2009] addresses this need and provides six RCM simulations driven by the same reanalysis boundary conditions from the National Centers for Environmental Prediction-Department of Energy (NCEP-DOE) Reanalysis II (R2) over North America for the contemporary 1980–2004 period. Systematic analyses of a variety of metrics regarding temperature and precipitation from these RCM simulations indicate that individual model performance varies seasonally and regionally, and different RCMs perform better than others in terms of different metrics [*Mearns et al*., 2012].

In this study, we evaluate the extent to which the various GCM-RCM pairs from the NARCCAP are able to reproduce present-day climate as represented by temperature and precipitation and to analyze the projected future changes in these two important climatic variables in response to anthropogenic greenhouse gas increases under the high emission scenario. The spatial domain for the NARCCAP simulations covers all of the lower 48 contiguous U.S. states, most of Canada, northern Mexico, and adjacent oceans. We focus on the entire Northeast U.S. region defined by the U.S. Geological Survey (USGS) for its Department of Interior Regional Climate Science Center, which is composed of 22 states and characterized by enormous climatic and biological diversity. In doing so, we extend the work of *Rawlins et al*. [2012], who presented a detailed assessment of biases in the estimates of current climate and uncertainties in future climate projections across the Greater New England region. The primary purposes of this study are to better understand the separate and combined sources of biases and uncertainties in regional climate simulations arising from applying different RCMs forced by multiple AOGCMs and to determine the robustness of projected future regional climate changes. The spatial distributions of biases and projected future changes in temperature and precipitation for individual model pairs are systematically examined. In addition to assessing specific metrics such as the mean and standard deviation, probability density functions of monthly temperature and precipitation are also provided to illustrate temporal and spatial distributions of these two climate variables. Moreover, biases in current simulations and uncertainties in future projections attributed solely to the driving AOGCMs are further analyzed and interpreted.

2. Models and Methodology

2.1. NARCCAP Simulations

The North American Regional Climate Change Assessment Program (NARCCAP) [*Mearns et al*., 2007, 2009] was developed for the purpose of systematically examining the combined uncertainties in regional climate change simulations from global and regional climate models. In phase II of the NARCCAP, four atmosphere-ocean general circulation models (AOGCMs) provide boundary conditions for six regional climate models (RCMs) to simulate climate of the recent past (1971–2000) and to project future climate (2041–2070). The future climate projections are forced using the Special Report on Emissions Scenarios (SRES) A2 emission scenario [*Nakićenović et al*., 2000]. The underlying assumptions for the SRES A2 scenario can be summarized as a very heterogeneous world of (i) continuously increasing population, (ii) regionally oriented economic development, and (iii) slow economic growth and technological change to preserve local identities. Under this relatively high emission scenario, the atmospheric carbon dioxide (CO_2) concentrations will reach approximately 530 ppm by midcentury and 850 ppm by the end of the 21st century (see Appendix II of *Houghton et al*. [2001]).

Instead of producing simulations from the full suite of 24 combinations (four AOGCMs × six RCMs), the NARCCAP applied a more balanced statistical design, wherein each AOGCM drives three different RCMs

Table 1. AOGCM-RCM Combinations (12 Pairings) Simulated in NARCCAP (Denoted by "X") and Nine Pairings Analyzed in This Study (Indicated by "Δ")

	RCM					
AOGCM	CRCM	MM5	WRFG	RCM3	ECP2	HRM3
CCSM	X Δ	X Δ	X Δ			
CGCM3	X Δ		X Δ	X Δ		
GFDL				X Δ	X Δ	X
HadCM3		X			X	X Δ

and each RCM is forced by two different AOGCMs for boundary conditions. The resulting half of the 4 × 6 AOGCM-RCM combinations (12 pairings) simulated in NARCCAP and nine pairings analyzed in this study are shown in Table 1.

2.2. Model and Observed Data

We employed high-resolution climate simulation outputs from multiple AOGCM-RCM pairings archived in the NARCCAP data set [*Mearns et al.*, 2007]. These NARCCAP simulation results, at 3-hourly intervals, are made available to the research community with a spatial resolution of approximately 50 km. Monthly means of 2 m air temperature and total precipitation rate, derived from the archived 3-hourly data, were analyzed to estimate model biases in current climate simulations and to examine uncertainties in their future climate projections. The four driving AOGCMs providing boundary conditions for the RCMs are the following: the National Center for Atmospheric Research Community Climate System Model version 3 [*Collins et al.*, 2006], the Canadian Global Climate Model version 3 (CGCM3) [*Flato et al.*, 2000], the Geophysical Fluid Dynamics Laboratory (GFDL) Climate Model version 2.1, and the Hadley Centre Climate Model version 3 (HadCM3) [*Pope et al.*, 2000].

The observed data set used here for model validation is version 3.0 of the Climatic Research Unit (CRU-TS 3.0) monthly time series of temperature and precipitation [*Mitchell and Jones*, 2005]. This database of monthly variations in climate was constructed from various sources of station records with an improved method of data homogenization. Thousands of station records around the world were included in this CRU database and interpolated onto a half-degree grid covering the global land surface. During database construction absolute values from stations were converted into anomalies prior to interpolation [*Jones*, 1994; *New et al.*, 2000]. This anomaly method has the weakness of excluding any station without the appropriate baseline period (1961–1990) normal. The process of absorbing different sources into a single database was refined to guard against unnecessary duplication and to ensure consistency. Despite the weaknesses in the approach of detecting inhomogeneities, this database is sufficient to provide best estimates of month-by-month climate variations. To facilitate the comparison, the model data were regridded from their native grid to the CRU half-degree grid over the Northeast U.S. domain.

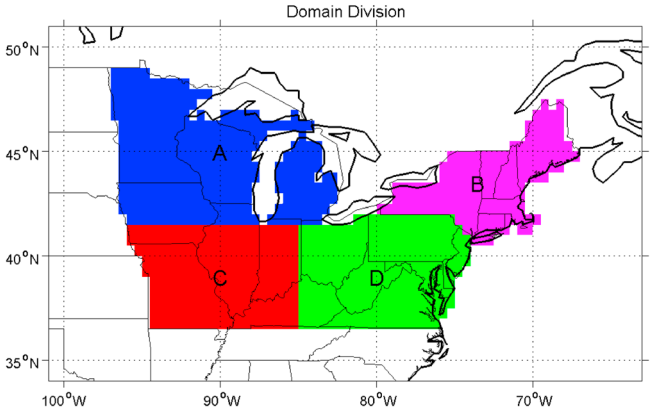

Figure 1. The Northeast U.S. domain divided into four subregions.

2.3. Analysis Methods

Based on the regionalization applied in a previous study [*Fan et al.*, 2014] using empirical orthogonal function analysis and K-means cluster analysis, we divided the Northeast U.S. domain into four subregions, representing relatively cold and dry (subregion A), cold and wet (subregion B), warm and dry (subregion C), and warm and wet (subregion D) climate zones (Figure 1). Besides analyzing mean climate, we also examined monthly mean surface air temperature and precipitation distributions for both contemporary (1971–2000) and future (2041–2070) periods across each of the

four subregions over the Northeast U.S. The probability density functions (PDFs) used to describe these distributions were calculated based on 30 year period monthly data for all the grid points within every subregion, and therefore reflect not only temporal but also spatial variability. The PDFs of monthly surface air temperature follow the Gaussian distribution as

$$f(x|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \tag{1}$$

where μ and σ represent the mean and standard deviation of the distribution, respectively. The gamma distribution, which is considered to be an appropriate fit to monthly precipitation, can be expressed as

$$f(x|k,\theta) = \frac{1}{\theta^k \Gamma} x^{k-1} e^{-\frac{x}{\theta}}, \tag{2}$$

where k and θ are the shape and scale parameters, respectively.

We also applied the "reliability ensemble averaging" (REA) [*Giorgi and Mearns*, 2002] procedure to estimate the weighted average change of temperature and precipitation from the multimodel ensemble members. This REA method takes into account both "model performance" criterion (i.e., a model's ability in reproducing observed current climate) and "model convergence" criterion (i.e., the degree of agreement of a model with the others for the future climate change) to weigh each individual model. The REA average temperature change is defined as a weighted average of the multimodel ensemble,

$$\widetilde{A}(\Delta T) = \frac{\sum_i R_i \Delta T_i}{\sum_i R_i}, \tag{3}$$

where ΔT_i is the simulated temperature change by an individual model, and $R_i \equiv R_{B,i} \times R_{D,i}$ is a model reliability factor calculated by the product of two factors measuring two respective criteria. The factor $R_{B,i}$ is a measure of model performance criterion. It is inversely proportional to the absolute value of the model bias in simulating present-day climate and has a fixed value for a given model in terms of a certain climatic variable. The other factor $R_{D,i}$ as a measure of the model convergence criterion is adjusted based on the distance of the change in each individual model from the REA average change through an iterative procedure. Higher weighting is given to a model with smaller bias and closer distance to the REA average change.

3. Model Validation

3.1. Multiannual Mean

The spatial patterns of multimodel mean surface air temperature biases are exhibited in Figure 2 for winter, spring, summer, and autumn seasons. The multimodel mean underestimates temperature over most of the Northeast U.S. region in all four seasons, and the underestimation is especially prominent during the winter. Despite the multimodel underestimation, individual model pairings exhibit quite different patterns of temperature biases. Figure 3 shows surface air temperature biases for each of the nine GCM-RCM pairs during the winter and summer seasons, respectively. Regions where the temperature biases are statistically significant at the two-sided $P = 0.05$ level are stippled.

During the winter, while Community Climate System Model-fifth-generation Penn State/NCAR Mesoscale Model (CCSM-MM5) tends to overestimate temperature over the northern part of the Northeast U.S., CCSM- Weather Research and Forecasting Model (WRFG) and the two RCMs driven by GFDL (i.e., GFDL-Experimental Climate Prediction-2 (ECP2) and GFDL-Regional Climate Model version 3 (RCM3)) underestimate temperature across the whole region. The driving models CCSM and CGCM3 exhibit comparable magnitudes of slightly overall underestimation of winter temperature (Figure 4a and Table S5 in the supporting information), but the WRFG forced by these two driving models produces substantial cold biases and moderate warm biases, respectively. Therefore, it is difficult to determine whether the strong cold biases in CCSM-WRFG are induced by the small cold biases from the driving model CCSM or additional biases generated by WRFG. GFDL-ECP2 has the largest winter temperature biases of any model pairings, locally more than 9°C below the observed temperature over the

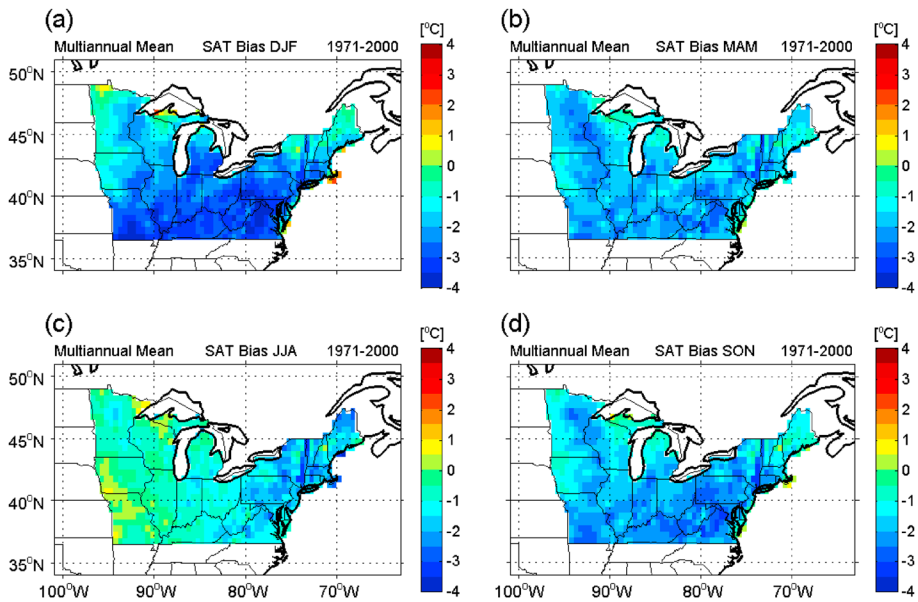

Figure 2. Biases (Model-CRU observed) in multimodel mean estimates of surface air temperature for (a) winter, (b) spring, (c) summer, and (d) autumn seasons during the contemporary 1971–2000 period.

southwestern part of the domain. As the driving model GFDL has moderate underestimation of winter temperature (−1.60 ~ −3.85°C), the strong cold biases in GFDL-ECP2 and GFDL-RCM3 may be partly attributed to their driving GCM. Moreover, the additional cold biases produced by ECP2 and RCM3 can explain the exaggerated cold biases in these two model pairings compared to their driving GFDL. It is evident that the RCM3 driven by the NCEP-DOE Reanalysis II (R2) as boundary conditions for the 1980–2004 period exhibits moderate winter temperature cold biases over the Northeastern U.S. [*Mearns et al.*, 2012], confirming that the additional cold biases generated by RCM3 are partly responsible for the pronounced underestimation in GFDL-RCM3. In contrast, the regional climate model HadRM3 driven by the NCEP-DOE Reanalysis II (R2) boundary conditions produces particularly large warm biases in winter. Consequently, slight underestimation of winter temperature shown by the driving HadCM3 is offset to some extent by HadRM3, resulting in decreased winter temperature biases for HadCM3-HRM3.

In summer, RCM simulations forced by CCSM show significant warm biases over the western part of the Northeast U.S. region, and the largest warm biases are exhibited by CCSM-CRCM over the southwestern region (3.0°C). More prominent summer warm biases are exhibited by the driving model CCSM over the western subregions (2.06~3.43°C) (Figure 4b and Table S5), making considerable contributions to the consistent overestimation of summer temperature in the three CCSM-driven model pairings. Statistically significant cold biases are still evident in two RCM simulations driven by GFDL in summer, although the magnitudes are much less than those in winter. Although the driving model HadCM3 overestimates summer temperature (0.87~2.18°C) over the entire region, the overestimation is generally cancelled out in HadCM3-HRM3.

Precipitation biases (expressed as percentage bias relative to the observed value) for the multimodel ensemble average in all four seasons and for the nine individual model pairs in winter and summer seasons are presented in Figures 5 and 6, respectively. For individual model pairs, regions where the precipitation biases are statistically significant at the two-sided $P = 0.1$ level are stippled. Multimodel mean estimates show different patterns of precipitation biases in different seasons, with predominantly wet biases in winter and spring, predominantly dry biases in autumn, and a mixture of both wet and dry biases during the summer. Individually, most of the model pairs mainly overestimate winter precipitation, especially CGCM3-RCM3, GFDL-ECP2, GFDL-RCM3, and HadCM3-HRM3 over the northwestern part of the domain, with regionally averaged overestimation exceeding 50%. The four driving AOGCMs produce even stronger overestimation of winter precipitation over the northwestern subregion, with percentage biases of 70.2%, 64.1%, 104.7%, and 112.8% for the CCSM, CGCM3, GFDL, and HadCM3, respectively (Figure 7a and Table S6). Given the dominance of large-scale features in winter precipitation [*Mearns et al.*, 2013], the particularly large wet biases in these several GCM-RCM pairs are

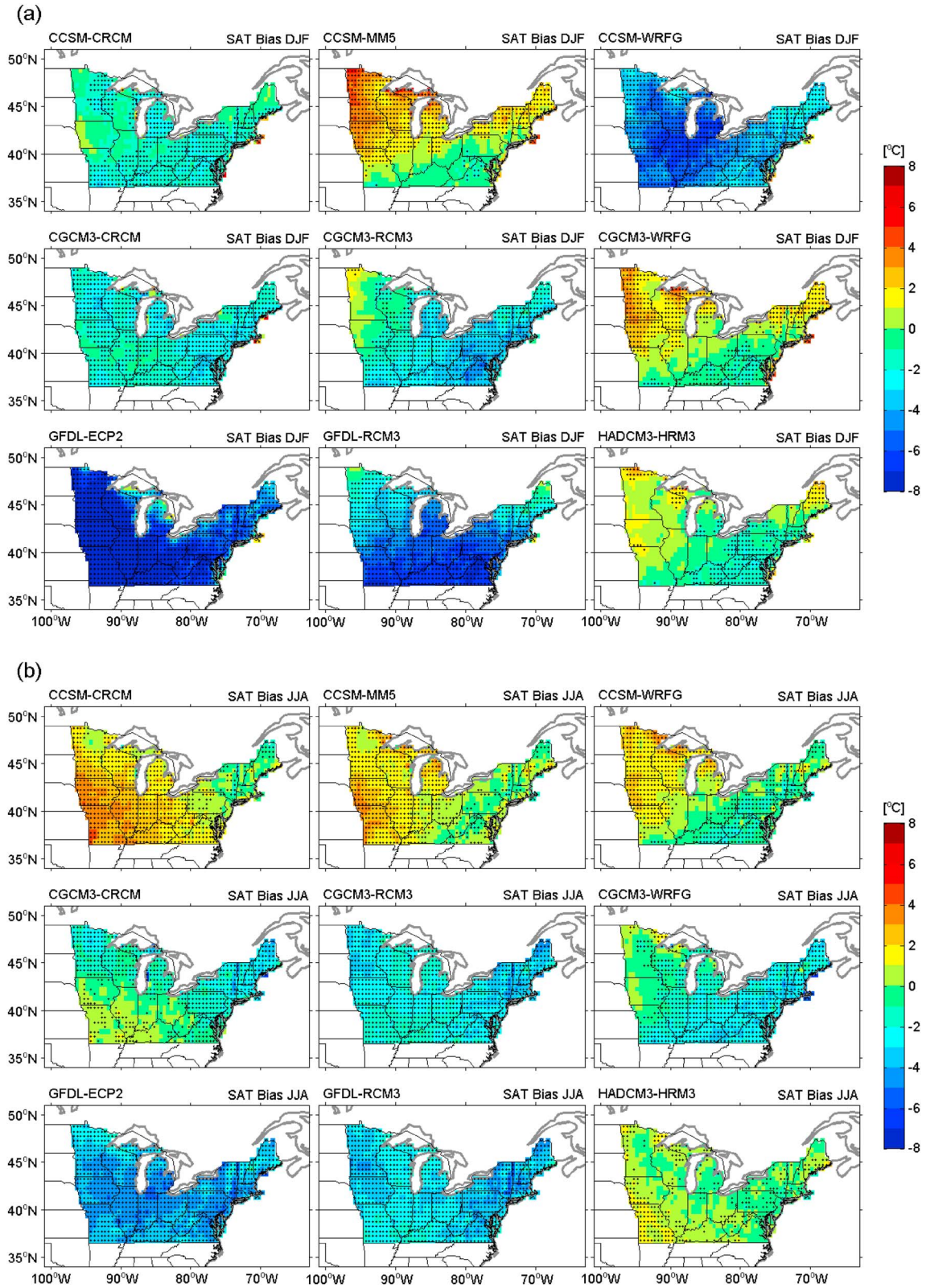

Figure 3. (a) Winter (December–February (DJF)) and (b) summer (June–August (JJA)) surface air temperature biases (Model-CRU observed) for each of the nine GCM-RCM pairs during the 1971–1999/2000 period. Stippled areas indicate that the biases are statistically significant at the two-sided $P = 0.05$ level.

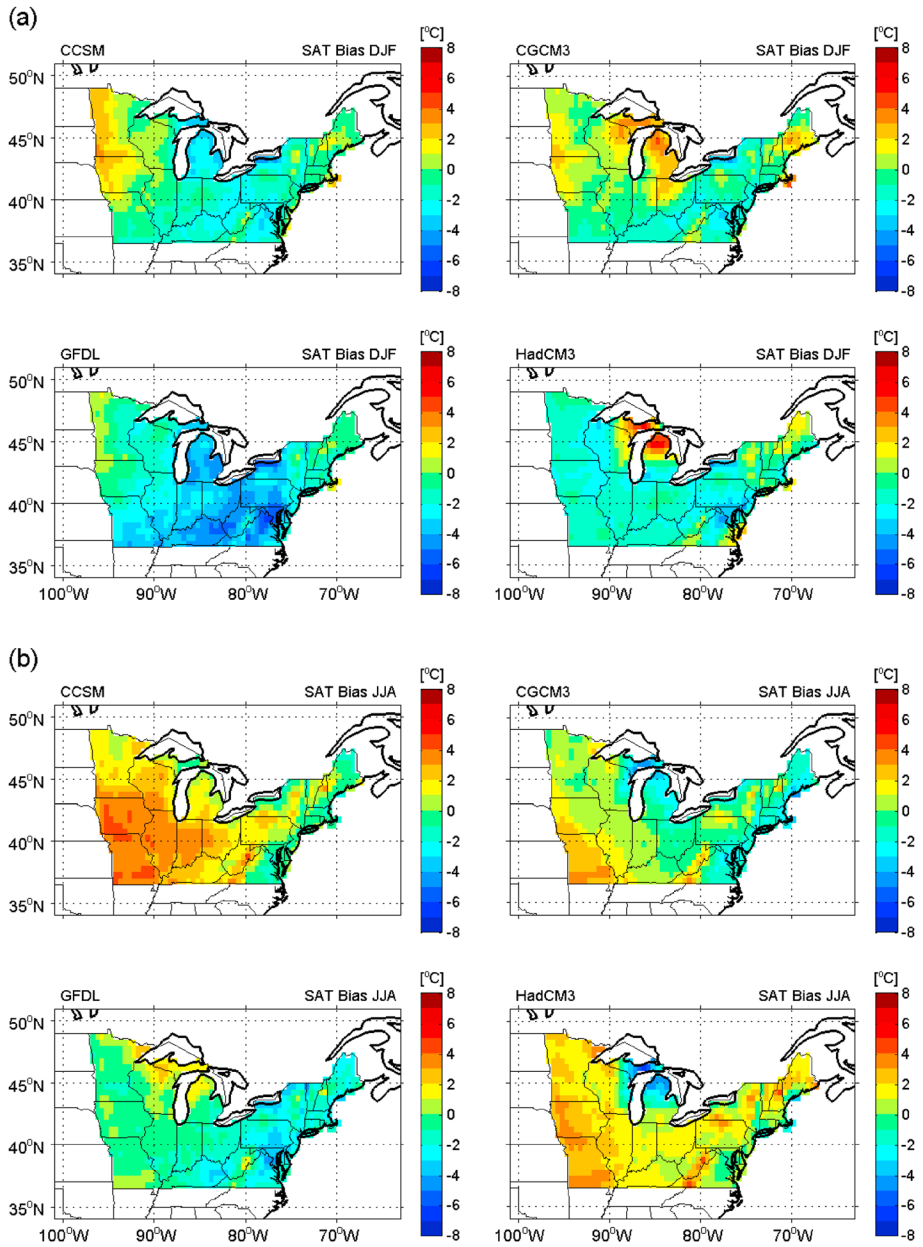

Figure 4. (a) Winter (December–February (DJF)) and (b) summer (June–August (JJA)) surface air temperature biases (Model-CRU observed) for the four driving AOGCMs during the 1971–1999/2000 period. To facilitate the comparison, the driving model data was interpolated onto the CRU 0.5° × 0.5° grid over the Northeast U.S. domain.

very likely inherited from their driving AOGCMs. However, CCSM-WRFG is an exception, producing substantial winter dry biases over most parts of the domain.

Summer precipitation biases are model dependent, ranging from substantial underestimation in CCSM-CRCM and CCSM-WRFG to significant overestimation in CGCM3-RCM3, GFDL-ECP2, and HadCM3-HRM3. The pronounced dry biases especially over the western subregions in CCSM-CRCM and CCSM-WRFG are inherited directly from their driving CCSM, whereas the prominent wet biases in HadCM3-HRM3 are ascribable to HadCM3 (Figure 7b and Table S6). Due to the greater importance of mesoscale processes (i.e., convective precipitation) in controlling summer precipitation, multiple RCM simulations (e.g., CGCM3-RCM3 and CGCM3-WRFG) driven by the same GCM simulation may exhibit contrasting summer precipitation mean biases. Comparing spatial patterns of summer precipitation biases (Figure 6b) to those of summer temperature

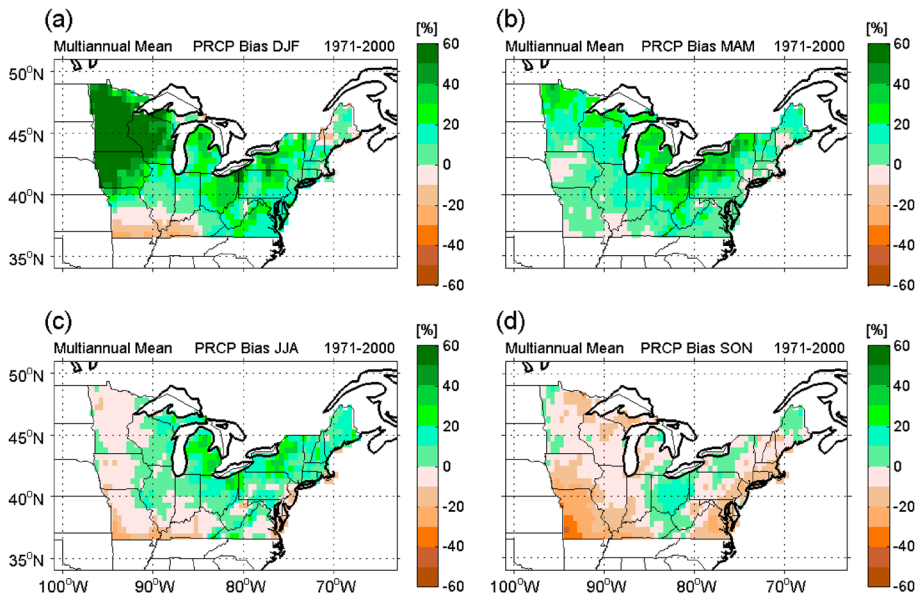

Figure 5. Percentage biases ((Model-CRU observed)/(CRU observed) × 100%) in multimodel mean estimates of total precipitation rate for (a) winter, (b) spring, (c) summer, and (d) autumn seasons during the contemporary 1971–2000 period.

biases (Figure 3b) from individual model pairs, we find that negative precipitation biases coincide with positive temperature biases over the western part of the domain for the three RCM simulations driven by the CCSM. One possible explanation is that less precipitation is associated with drier soils and less cloud cover, which further induce less evaporative cooling and enhanced solar radiation at the surface (Figures S1a and S1b in the supporting information). These effects result in increased sensible heat flux from the surface and thus higher surface air temperature (Figure S1c).

3.2. Probability Density Functions (PDFs)

The PDFs of monthly surface air temperature and precipitation for each of the nine model pairs in four distinct subregions of the Northeast U.S. domain are presented in Figures 8–11. These analyses indicate how well the models reproduce temporal and spatial distributions of these two important climatic variables. Model biases in the mean and variability of temperature and precipitation relative to the CRU-observed data in four subregions are recorded in Tables S1–S4. Histogram-based measures of temperature and precipitation biases are also provided to illustrate the models' skills in reproducing current climate (Text S4 in the supporting information).

For winter temperature distributions, GFDL-ECP2 displays the coldest mean temperature and relatively large variability, whereas CCSM-MM5 shows the warmest mean temperature and relatively small variability (Figure 8a). Relative to the CRU observed data, there is a large spread in the mean and variability biases of winter temperature in different individual model pairs (Figures S2a and S2b). The largest underestimations of winter temperature are exhibited in GFDL-ECP2, which are −6.7, −4.6, −9.3, and −6.3°C averaged over the northwestern (A), northeastern (B), southwestern (C), and southeastern (D) subregions, respectively. CCSM-MM5 produces the largest overestimations of winter temperature in three subregions (A, B, and C) but a slight underestimation (−0.1°C) in the southeastern subregion D (Table S1). The winter temperature variability biases over the northern regions range from the largest negative biases in GFDL-RCM3 (−24.9 ~ −30.7%) to the largest positive biases in GFDL-ECP2 (14.2~16.1%). In the southern regions, the model biases in winter temperature variability fall in the range between negative biases in CGCM3-WRFG (−12.4 ~ −32.3%) and positive biases in CCSM-WRFG (19.3~24.7%), although GFDL-RCM3 (GFDL-ECP2) still shows the second largest negative (positive) variability biases (Table S1).

In summer, GFDL-ECP2 shows not only the coldest mean temperature but also the smallest temperature variability in almost all subregions. In contrast, CCSM-CRCM shows the warmest mean temperature and the largest temperature variability (Figures 9a, S3a, and S3b). The widths of summer temperature distributions

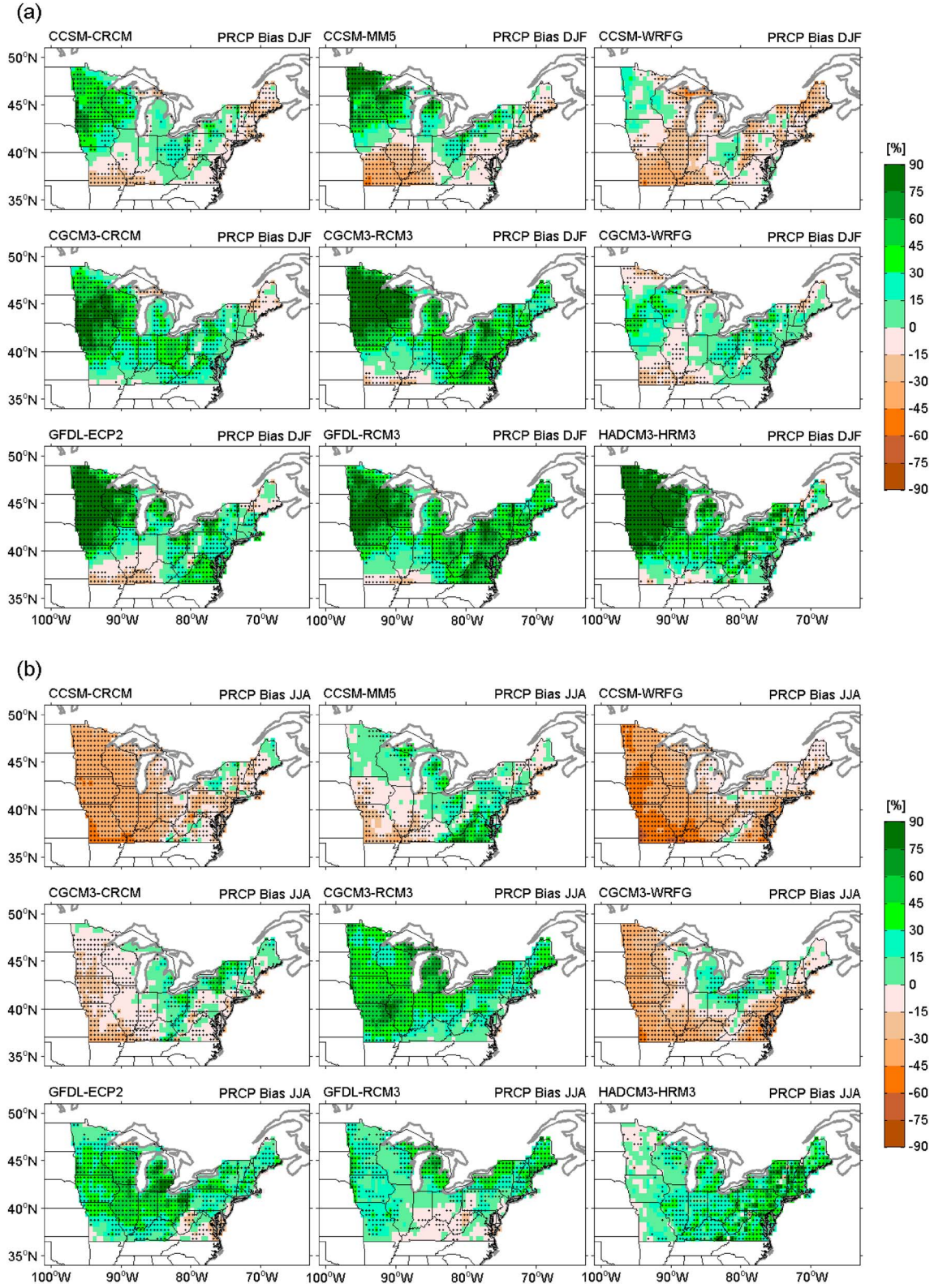

Figure 6. (a) Winter (December–February (DJF)) and (b) summer (June–August (JJA)) total precipitation rate biases ((Model-CRU observed)/(CRU observed) × 100%) for each of the nine GCM-RCM pairs during the 1971–1999/2000 period. Stippled areas indicate that the biases are statistically significant at the two-sided $P = 0.1$ level.

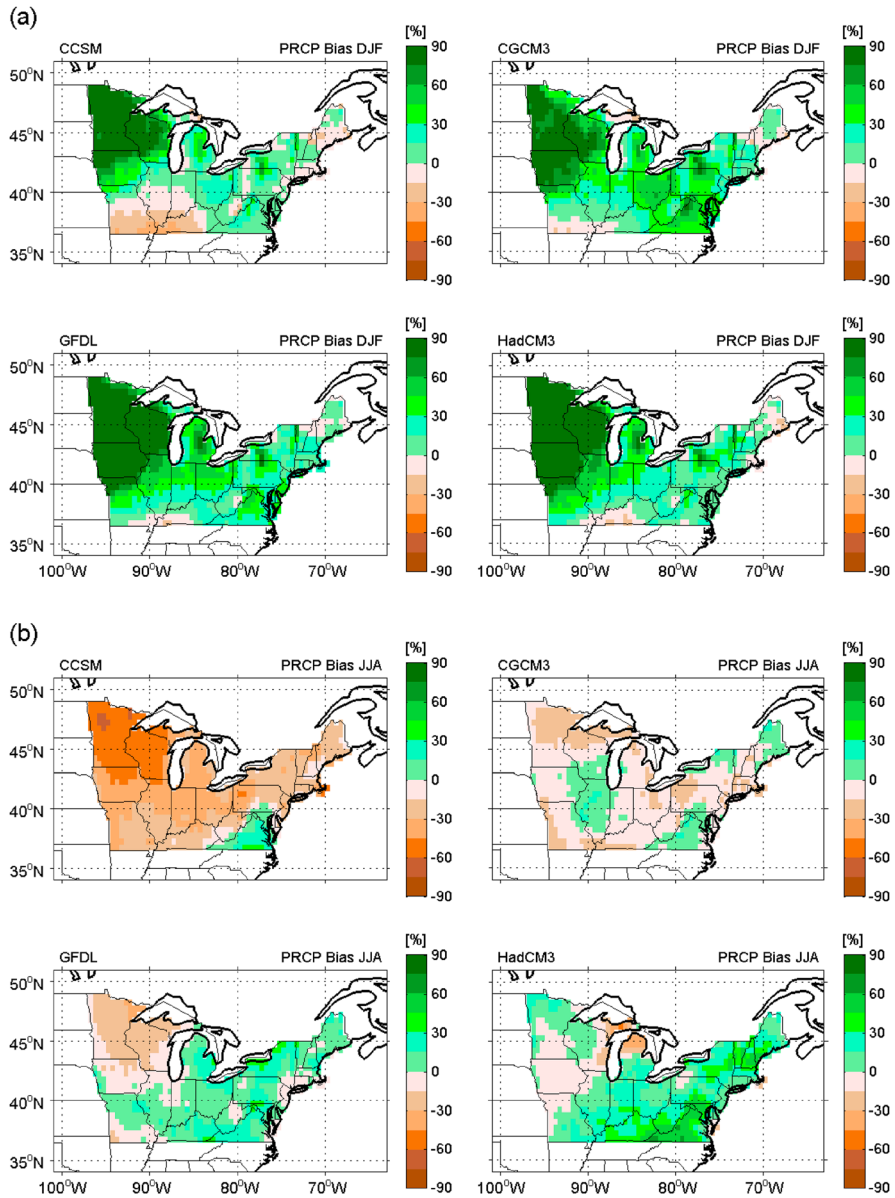

Figure 7. (a) Winter (December–February (DJF)) and (b) summer (June–August (JJA)) total precipitation rate biases ((Model-CRU observed)/(CRU observed) × 100%) for the four driving AOGCMs during the 1971–1999/2000 period.

are much narrower than those of winter temperature distributions in both observed data and model simulations, indicating that the models tend to exhibit diminished temperature variability during summer. As winter temperature biases, GFDL-ECP2 also produces the largest cold biases of summer mean temperature (−3.8 ~ −4.2°C). The temperature variability biases in this model pair change from positive biases in winter (14.2~20.4%) to negative biases in summer (−0.1 ~ −16.7%) (except for the southwestern subregion C). CCSM-MM5, which shows the largest warm biases in winter mean temperature and negative winter temperature variability biases, exhibits the largest overestimates of summer mean temperature (summer temperature variability) over the northeastern subregion B (southwestern subregion C) and the second largest warm biases in summer mean temperature in the other three subregions (Table S1).

The PDFs of winter precipitation indicate relatively dry conditions with small variability over the northwestern subregion (A), where the observed (ensemble averaged) mean and variability of winter precipitation are only 0.91 mm day^{-1} and 0.62 mm day^{-1} (1.27 mm day^{-1} and 0.78 mm day^{-1}) (Figure 10a). In this dry region,

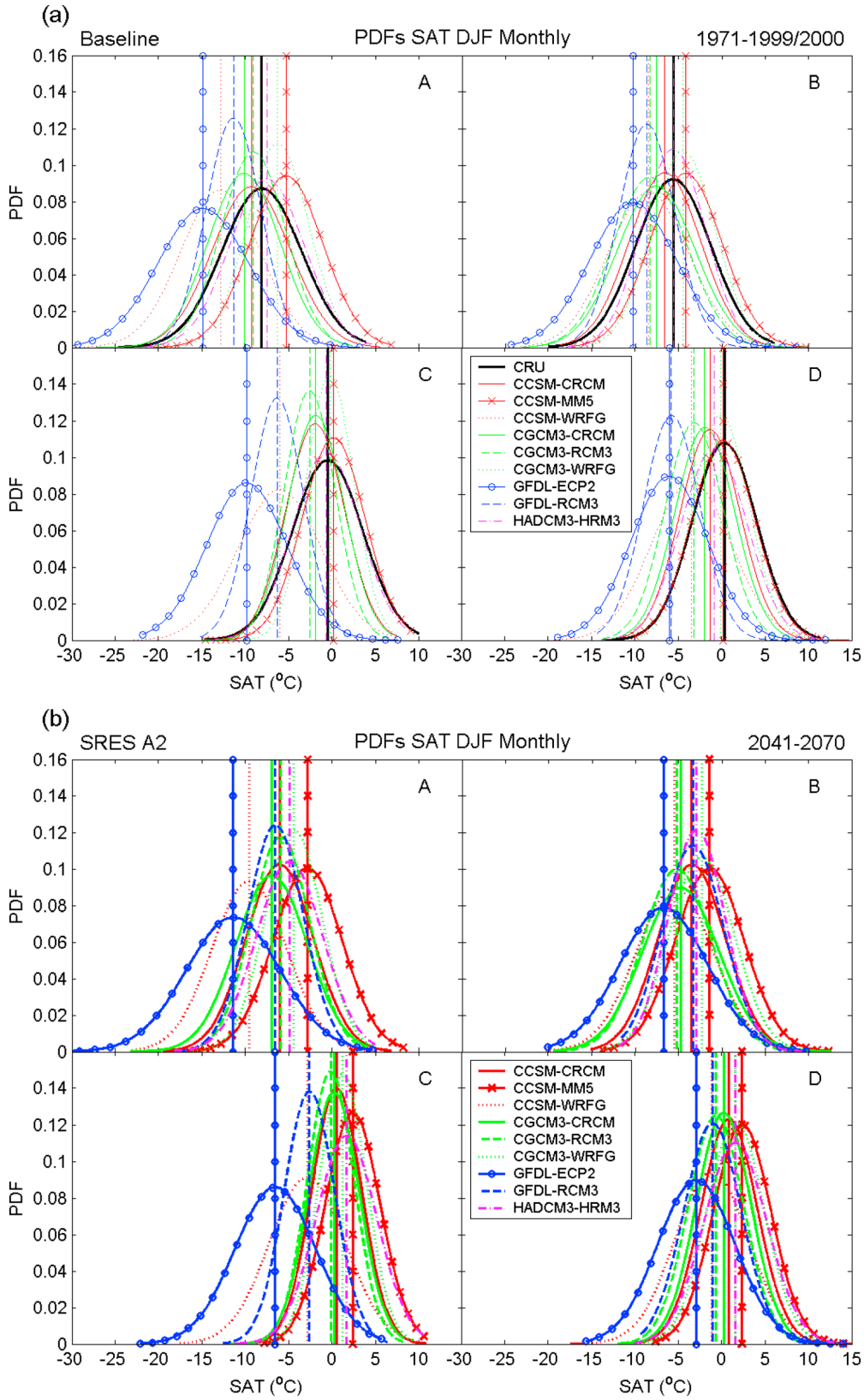

Figure 8. Probability density functions of monthly surface air temperature in winter season for the CRU-observed data and each of the nine GCM-RCM pairs in the four Northeast U.S. subregions (A, B, C, and D) during (a) the present 1971–1999/2000 period and (b) future 2041–2070 period. Vertical lines denote the means of respective distributions.

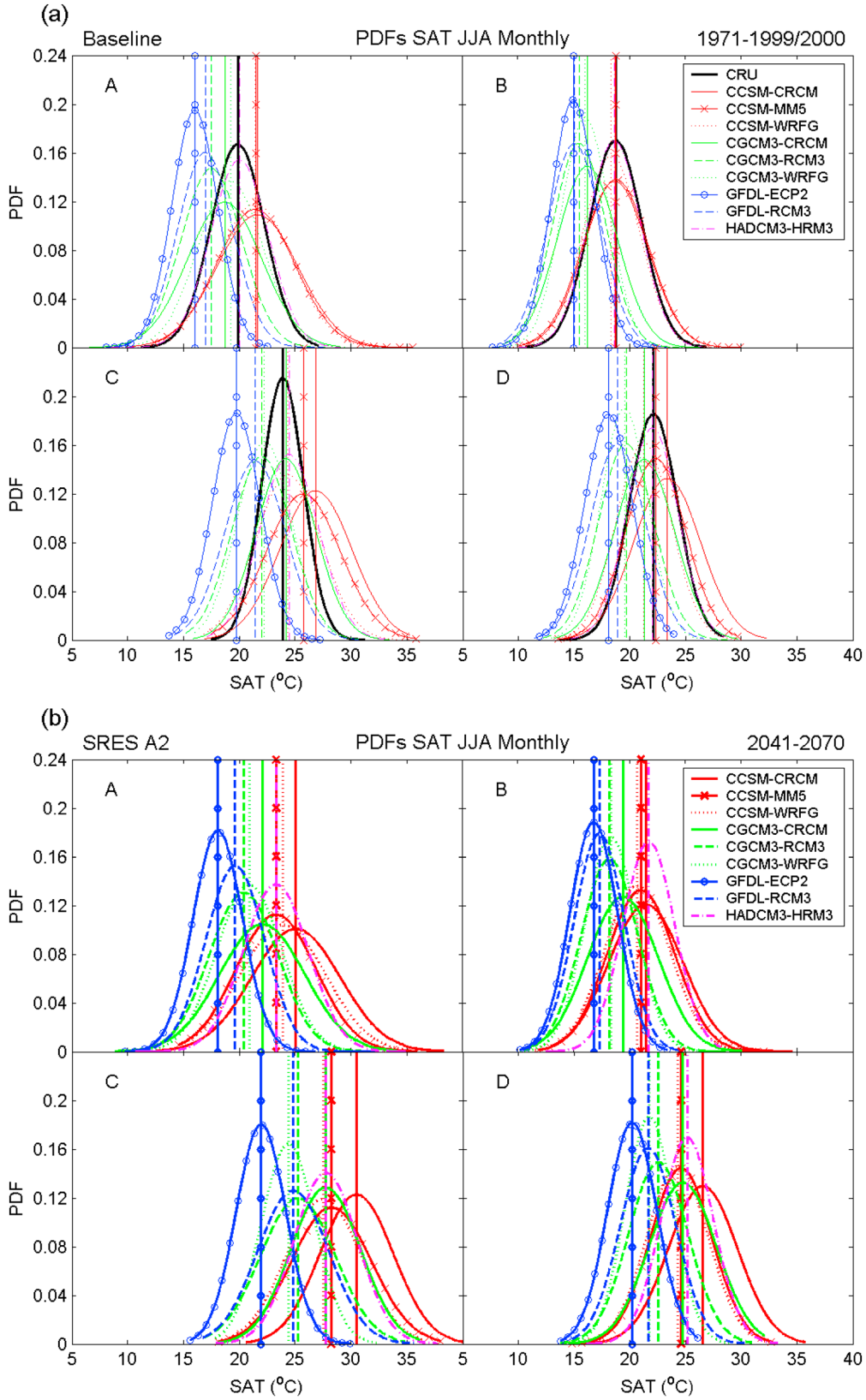

Figure 9. Probability density functions of monthly surface air temperature in summer season for the CRU-observed data and each of the nine GCM-RCM pairs in the four Northeast U.S. subregions (A, B, C, and D) during (a) the present 1971–1999/2000 period and (b) future 2041–2070 period. Vertical lines denote the means of respective distributions.

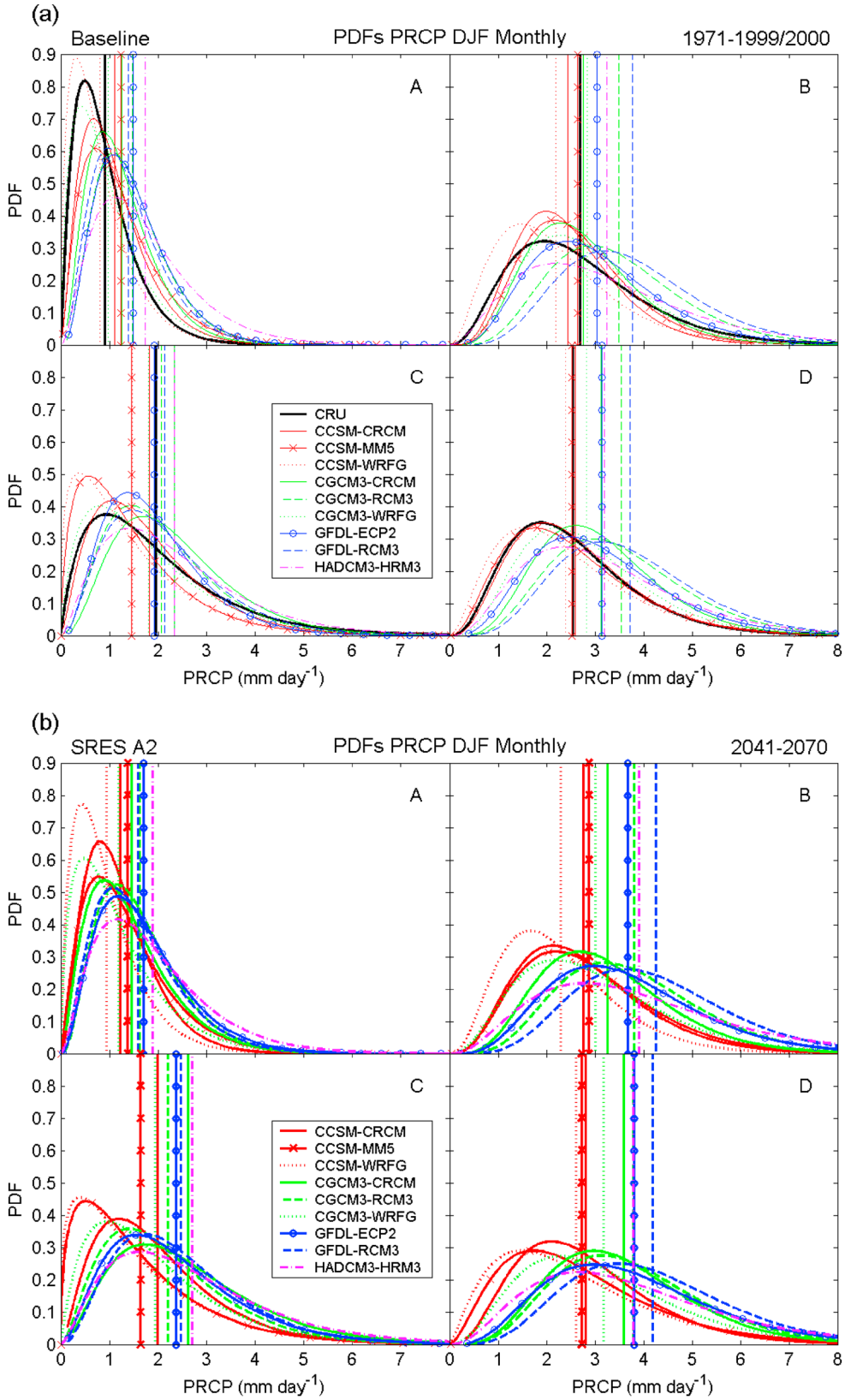

Figure 10. Probability density functions of monthly total precipitation rate in winter season for the CRU-observed data and each of the nine GCM-RCM pairs in the four Northeast U.S. subregions (A, B, C, and D) during (a) the present 1971–1999/2000 period and (b) future 2041–2070 period. Vertical lines denote the means of respective distributions.

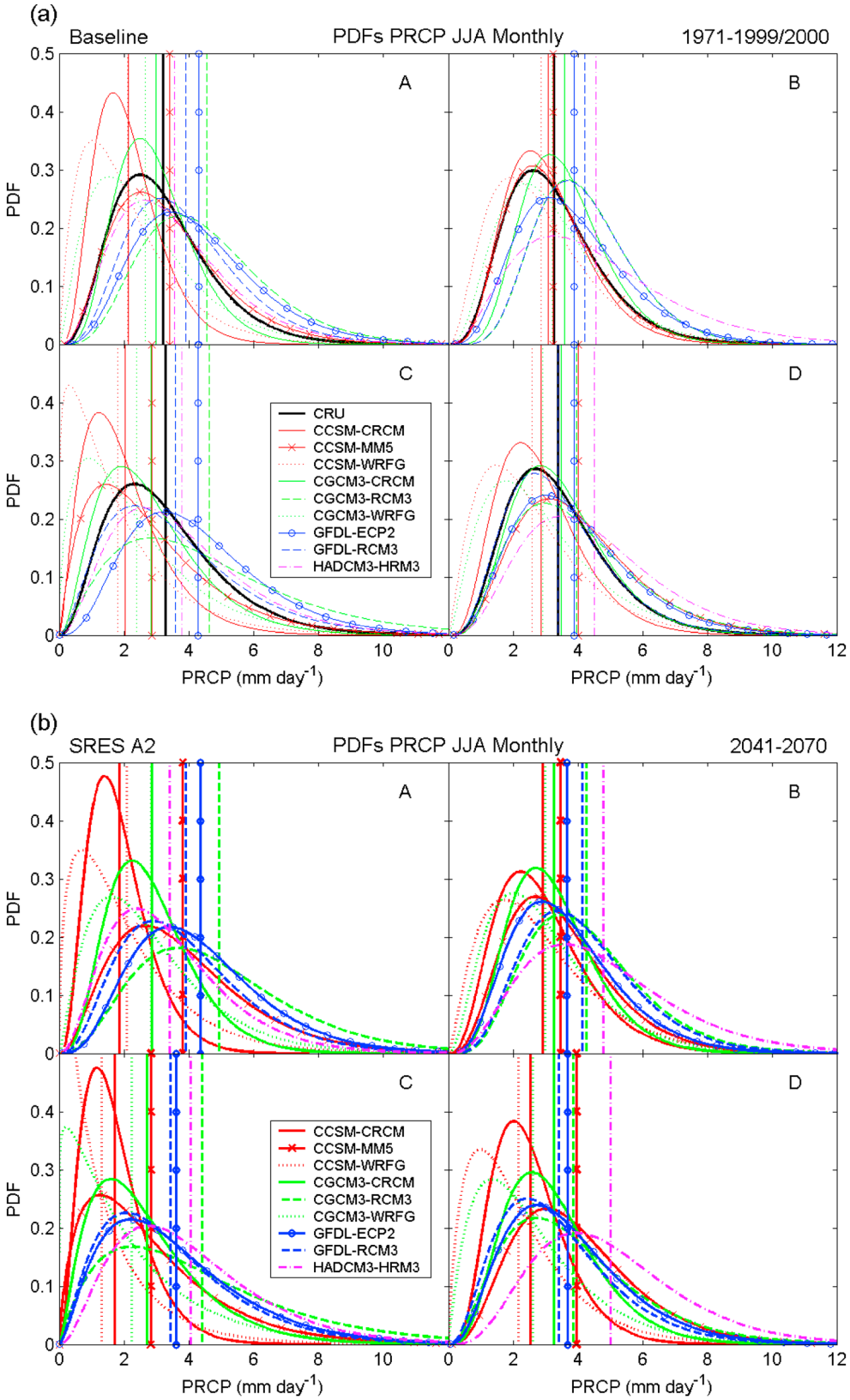

Figure 11. Probability density functions of monthly total precipitation rate in summer season for the CRU observed data and each of the nine GCM-RCM pairs in the four Northeast U.S. subregions (A, B, C, and D) during (a) the present 1971–1999/2000 period and (b) future 2041–2070 period. Vertical lines denote the means of respective distributions.

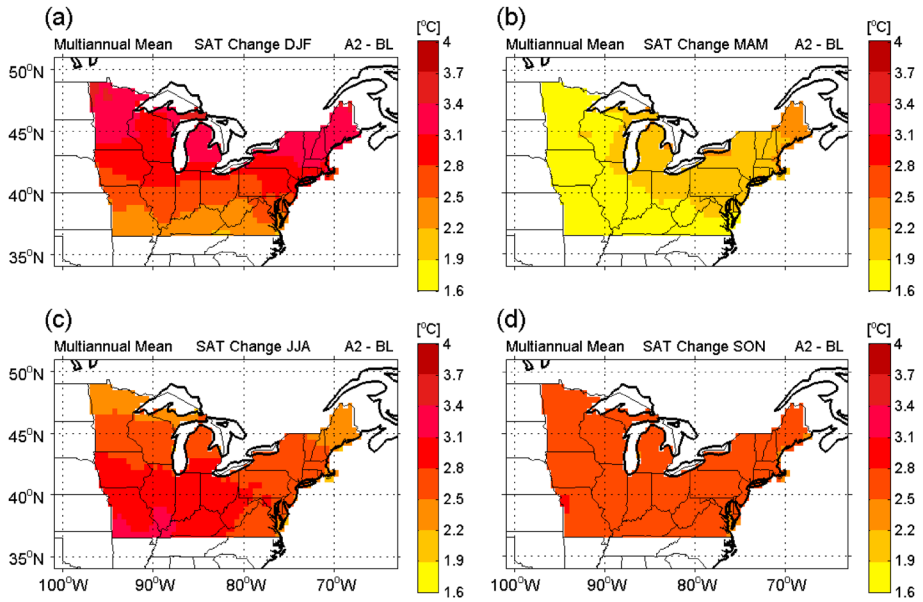

Figure 12. Projected future changes (SRES A2-Baseline) in multimodel mean surface air temperature during (a) winter, (b) spring, (c) summer, and (d) autumn seasons.

although the absolute biases in winter mean precipitation are fairly small, most model pairs exhibit pronounced overestimations of winter precipitation as measured by percentage biases, with the HadCM3-HRM3 exhibiting an extremely large wet bias exceeding 90%. In contrast, CCSM-WRFG is the only model pair that exhibits a regionally averaged dry bias in winter mean precipitation over the northwestern subregion A (−11.0%). Moreover, as mentioned earlier, this model (CCSM-WRFG) also produces substantial overall winter dry biases, with the underestimates being −18.5%, −24.8%, and −5.7% over the other three subregions (B, C, and D), respectively (Figure S2). The largest wet biases in winter mean precipitation over the eastern parts of the domain are exhibited by GFDL-RCM3, exceeding 40% in the northeastern (B) and southeastern (D) subregions (Table S2). While all nine model pairs overestimate winter precipitation variability over the northwestern subregion A, all but one (HadCM3-HRM3) display underestimations of variability over the southwestern subregion C (Figure 10a). In contrast to its performance in reproducing winter temperature, HadCM3-HRM3 not only exhibits substantially large wet bias in winter mean precipitation but also produces the largest positive bias in winter precipitation variability over all four subregions (Table S2 and Figure S2).

Compared to the winter precipitation distributions, the PDFs of summer precipitation illustrate larger means and wider widths for both the observations and all model simulations, indicating more abundant water supply and enhanced precipitation variability during summer (Figure 11a). Of all nine model pairs, CCSM-WRFG has the largest dry biases in all four subregions, with the underestimations more significant over relatively dry subregions (A and C) (−34.6 ~ −44.9%), whereas CGCM3-RCM3 and HadCM3-HRM3 show the largest wet biases over the western (41.7~42.2%) and eastern (33.5~39.8%) subregions, respectively (Table S2 and Figure S3).

4. Climate Change Response

4.1. Surface Air Temperature Response

The projected future changes in surface air temperature show consistent warming trends for all four seasons, with more prominent temperature increases occurring over the northern subregions (A and B) during the winter and the southwestern subregion (C) during the summer (Figure 12). The ensemble averaged winter temperature increases are 3.1, 3.2, 2.6, and 2.6°C over the northwestern, northeastern, southwestern, and southeastern subregions, respectively (Table S3). Compared to winter warming, the multimodel means for summer temperature changes show lower increases over the northern subregions (2.6 and 2.5°C) but higher increases over the southern subregions (3.0 and 2.8°C). Different individual model pairs exhibit different

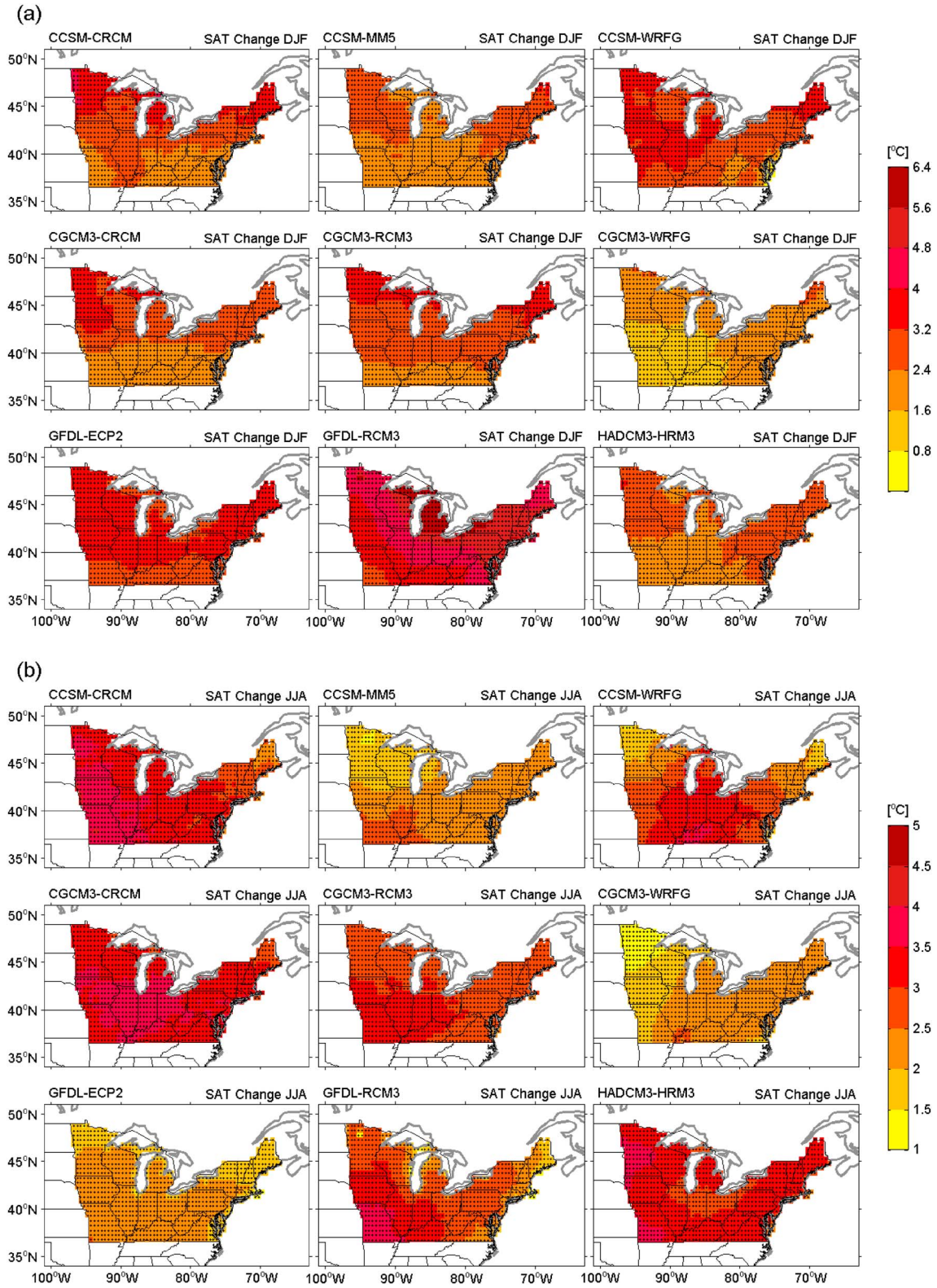

Figure 13. Projected future changes (SRES A2—Baseline) in surface air temperature for each of the nine GCM-RCM pairs during (a) winter and (b) summer seasons. Stippled areas indicate that the changes are statistically significant at the one-tailed $P = 0.05$ level.

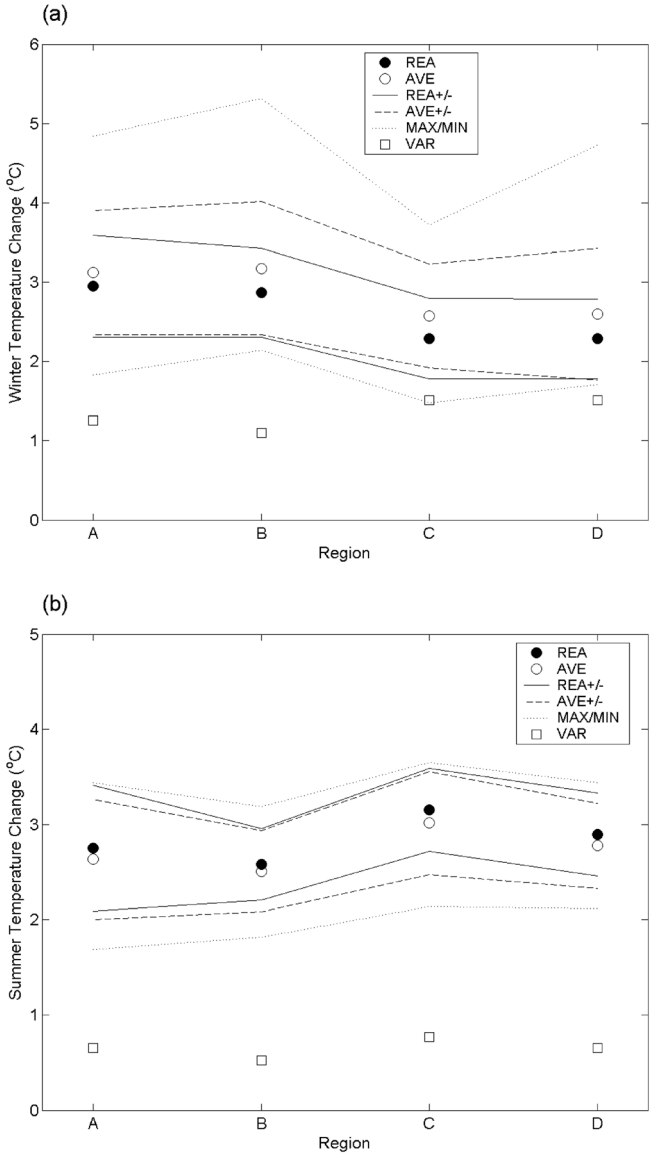

Figure 14. REA average changes (REA, dark circles) and corresponding upper and lower REA uncertainty limits (REA+/−, continuous lines), ensemble average changes (AVE, open circles) and corresponding uncertainty limits (AVE+/−, dashed lines), maximum and minimum changes simulated by individual model pairs in the ensemble (MAX/MIN, dotted lines), and estimated natural variability values (VAR, squares) for surface air temperature over the four Northeast U.S. subregions during (a) winter and (b) summer seasons.

spatial patterns of warming during winter and summer seasons (Figure 13). Based on Student's *t* test, winter temperature increases are statistically significant at the one-tailed $P = 0.001$ level across the whole Northeast U.S. domain in all nine model pairs, except for a very few grid points along the eastern coast in CCSM-WRFG. Similarly, all nine model pairs project substantially significant summer warming at $P = 0.001$ level by one-sided *t* test, except for a small area over the northwestern corner in three model pairs (i.e., CCSM-MM5, CCSM-WRFG, and CGCM3-WRFG), where the warming is still significant at the one-sided $P = 0.05$ level.

Changes in the PDFs of winter temperature indicate decreased variability in most model pairs, with the ensemble averaged variability changing by −2.4~−8.0% (Figure 8b and Table S3). The largest decreases in variability arise in the CCSM-CRCM over the western subregions (A and C) (−13.9% and −15.3%). GFDL-ECP2 and GFDL-RCM3, which exhibit the largest positive and negative variability biases over the northern subregions, are the only two models projecting increased winter temperature variability. In contrast, with the exception of a few cases, the models show increased summer temperature variability over all four subregions (Figure 9b and Table S3).

Compared to the ensemble average temperature changes, the REA average temperature changes are lower for winter but higher for summer across the four subregions (Figure 14). The differences between these two average changes are on the order of a few tenths of a degree (0.2~0.3°C) in winter and only on the order 10^{-2}°C to 0.1°C in summer. The reason for the lower REA temperature change relative to the ensemble average change during the winter is that models with high future winter temperature increases (i.e., GFDL-ECP2, GFDL-RCM3, and CCSM-WRFG) are accompanied by strong present-day biases. As the REA method tends to give higher weights to models with lower bias and higher convergence, these three models receive relatively low weights. The REA winter temperature change therefore is dominated by the other models with more moderate temperature increases. The natural variability estimates for winter temperature (1.10~1.51°C) are twice as large as those for summer temperature (0.52~0.77°C) over the four subregions. The snow-ice albedo feedback mechanism may be responsible for relatively large winter temperature variability by further increasing (decreasing) the temperature during relatively warm (cold) periods through positive feedback process. It is important to

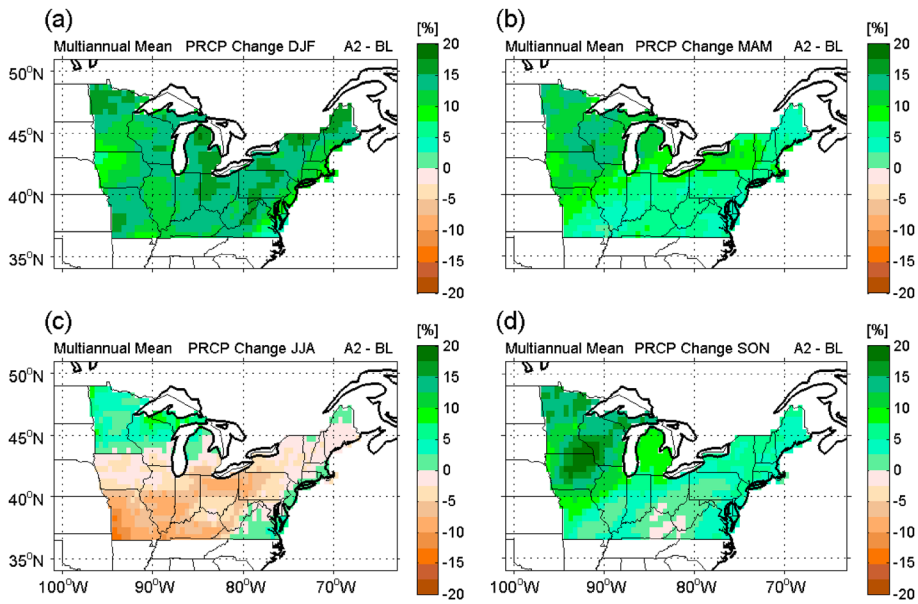

Figure 15. Projected future changes ((SRES A2-Baseline)/(Baseline) × 100%) in multimodel mean total precipitation rate during (a) winter, (b) spring, (c) summer, and (d) autumn seasons.

note that the natural variability estimates for summer temperature are remarkably less than the REA and ensemble average summer warming. Although the winter natural variability estimates are still well below the average warming, their magnitudes are comparable to the lower uncertainty limits over the southern subregions (C and D). The distance between the maximum and minimum temperature changes, reflecting the full range of warming simulated by individual model pairs, is more than twice as large as the REA uncertainty range during the winter and slightly larger than either REA or ensemble uncertainty range during the summer. The large intermodel spread in winter warming (2.2~3.2°C), especially over the northern regions (3.0~3.2°C), indicates that different representations of snow-ice albedo feedback in these model pairs may have a great influence on their sensitivity to the greenhouse gas forcing.

4.2. Precipitation Response

The ensemble averaged winter precipitation change is positive over the study domain (12.5~13.6%) (Figure 15a and Table S4), with all nine individual model pairs exhibiting positive changes over all four subregions (Figure 16a). The increases in winter precipitation projected by these GCM-RCM pairs are consistent with that predicted by their driving GCMs (Figure 17a and Table S7). The most prominent winter precipitation increases are shown in CGCM3-WRFG over the northwestern subregion (A) (22.1%) and in GFDL-ECP2 over the other three subregions (21.0~23.0%). The winter precipitation increases projected by CGCM3-CRCM and GFDL-ECP2 are statistically significant (two-sided $P < 0.1$ confidence level) for almost the entire Northeast U.S. domain, whereas the positive and significant changes in the RCM simulations driven by CCSM are restricted to small amounts of grid points over the northern region (Figure 16a). Associated with enhanced winter mean precipitation, winter precipitation variability averaged across the models is also projected to increase by 15.7~19.4% over the four Northeast U.S. subregions (Table S4). Individually, projected future changes in the widths of the PDFs suggest increases in winter precipitation variability in all nine model pairs (Figure 10b), ranging from the lowest increase in CCSM-WRFG over the northeastern subregion (B) (1.6%) to the highest increase in GFDL-ECP2 over the southwestern subregion (C) (37.4%).

Summer precipitation as estimated by the multimodel mean is projected to experience a slight increase over the northwestern subregion (A) (0.64%) but a reduction across the other three subregions (−1.5~−7.9%) (Figure 15c and Table S4). Unlike the consistently wetter winters, different individual model pairs project quite different spatial patterns of summer precipitation changes (Figure 16b). It is particularly noteworthy that the changing directions of summer precipitation in GCM-RCM pairs may be not consistent with that in their driving GCMs. While CCSM-WRFG projects substantially significant decreases in summer precipitation

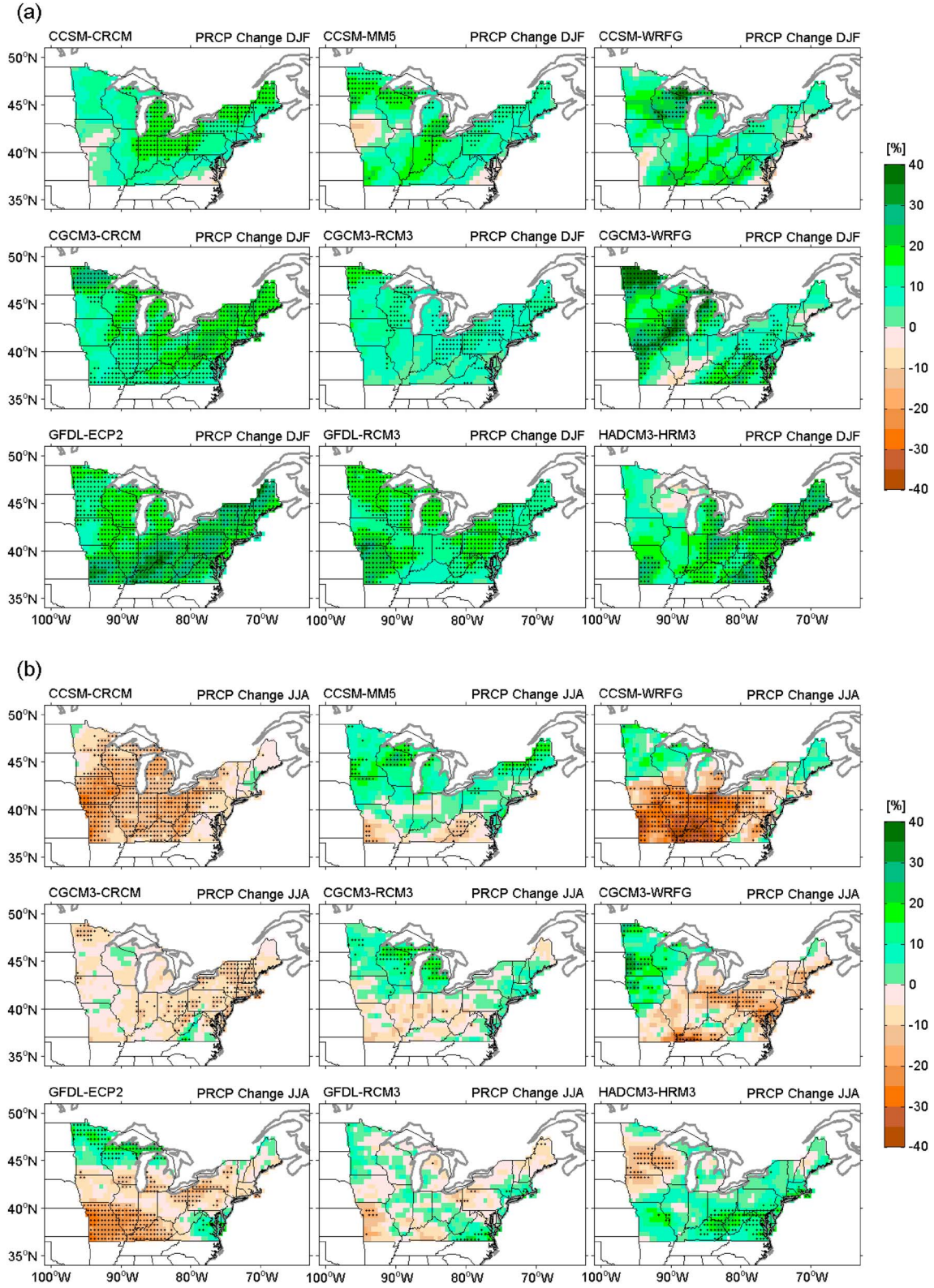

Figure 16. Projected future changes [(SRES A2-Baseline)/(Baseline) × 100%] in total precipitation rate for each of the nine GCM-RCM pairs during (a) winter and (b) summer seasons. Stippled areas indicate that the changes are statistically significant at the two-sided $P = 0.1$ level.

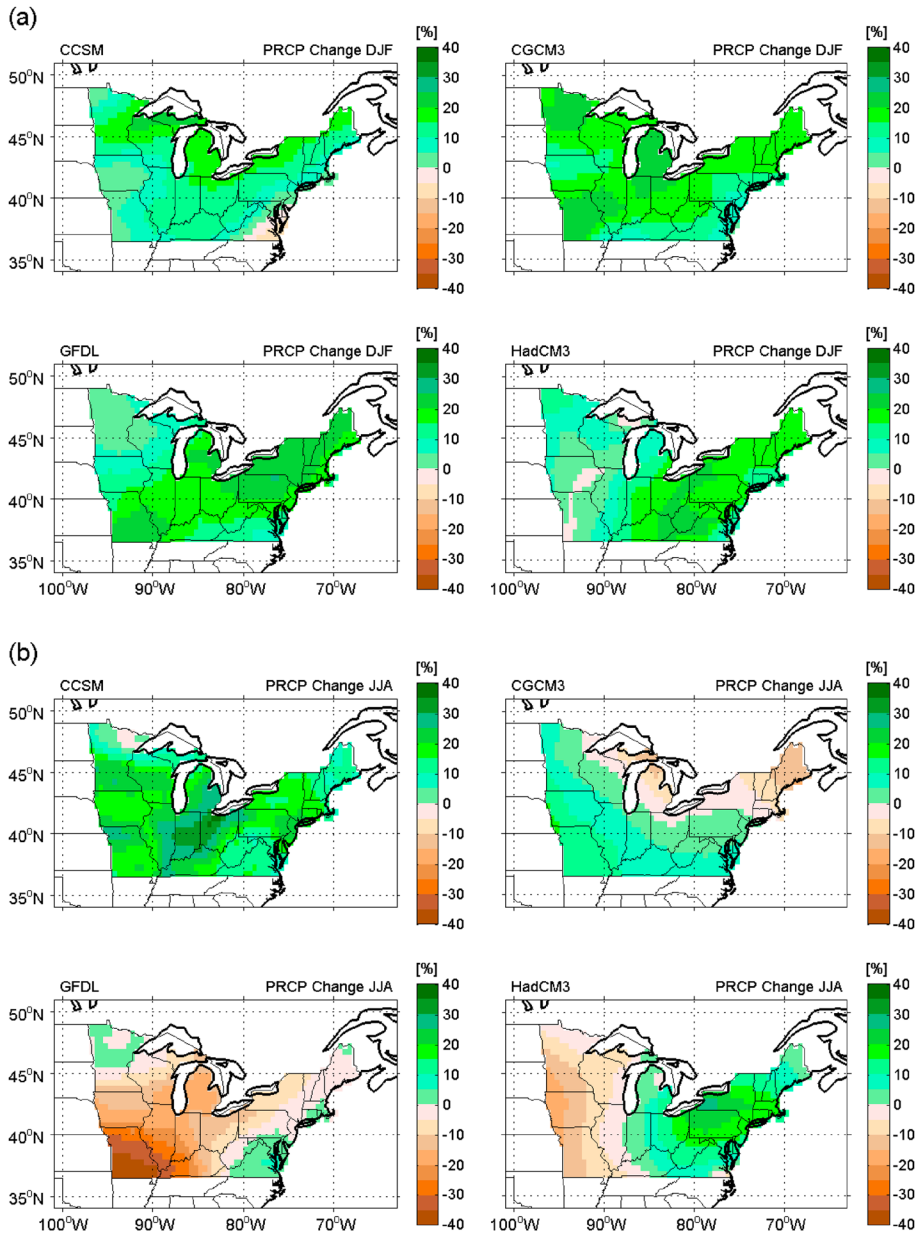

Figure 17. Projected future changes [(SRES A2-Baseline)/(Baseline)×100%] in total precipitation rate for the four driving AOGCMs during (a) winter and (b) summer seasons.

(−16.3 ~ −27.0%) over the southern subregions (C and D) (Figure 16b), its driving CCSM predicts intensified summer precipitation (14.4~23.2%) in almost the entire Northeast U.S. region (Figure 17b and Table S7). For the purpose of better understanding the opposite changing directions of summer precipitation, we further examined moisture transport at 850 hPa from the two GCM-RCM pairs (i.e., CCSM-CRCM and CCSM-WRFG) and their driving CCSM (Figure 18 and Text S5). Consistent with the findings by *Bukovsky and Karoly* [2011], the difference in direction of change between the CCSM-WRFG and its driving CCSM for summer precipitation can be explained by the difference of change in net moisture flux convergence-divergence. Although both the two GCM-RCM pairs and their driving CCSM project strengthened moisture transport at 850 hPa, the changes in moisture transport exhibit distinctive patterns. Changes in moisture transport projected by the CCSM-CRCM and CCSM-WRFG are characterized by anticyclonic flows in the southern subregions, indicative of increased moisture divergence. In contrast, the driving CCSM projects substantially larger increases in moisture transport into the western boundary of the Northeast U.S. region, suggesting enhanced net moisture convergence. The

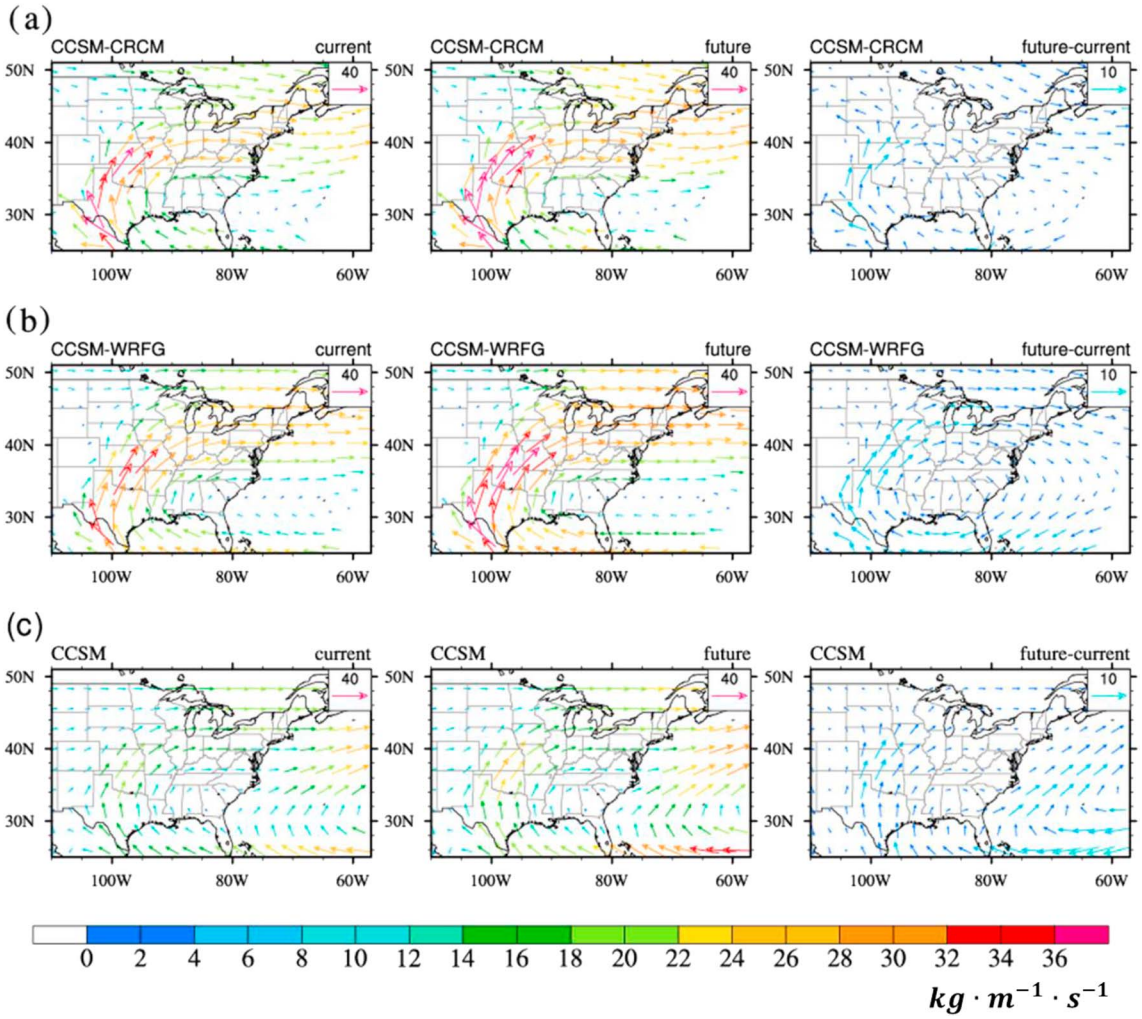

Figure 18. Multiannual mean moisture transport at 850 hPa in summer season for (left column) the current 1971–1999 period and (middle column) future 2041–2070 period simulated by (a) CCSM-CRCM, (b) CCSM-WRFG, and (c) the driving CCSM; (right column) projected future changes (SRES A2-Baseline) in moisture transport at 850 hPa for summer season from these three model simulations.

comparison between the future and current summer precipitation distributions suggests both increases and decreases in variability across different individual model pairs (Figures 11a and 11b). The largest decreases in summer precipitation variability over the northwestern subregion (A) (−6.6%) and two southern subregions (C and D) (−22.9% and −16.6%) are projected by CCSM-CRCM (Table S4), which produces the strongest negative variability biases (−7.6 ~ −27.5%).

The differences between the REA and ensemble average changes in winter precipitation are less than 3%, and in the northeastern (B) and southwestern (C) subregions these two average changes are almost equal to each other (Figure 19a). Relatively large REA average change compared to ensemble average change in winter precipitation over the northwestern subregion (A) is due to the fact that the models (i.e., HadCM3-HRM3 and CGCM3-RCM3) with relatively small precipitation increases contain relatively large wet biases, and thus, their contributions to the average are reduced in the REA method. The lower uncertainty limits of the REA and ensemble average winter precipitation changes are well above the estimated natural variability over the northwestern subregion (A). Although the REA and ensemble average winter precipitation changes are also above the natural variability in two southern subregions (C and D), their lower uncertainty limits are within the range of natural variability. Region B has the largest natural variability of winter precipitation (12.5%), which is very close to the REA (12.9%) and ensemble average changes (12.7%). The magnitudes of the REA (ensemble) average changes in summer precipitation are less than 5% (8%), which are within the range of natural variability, except for the ensemble average change in the southwestern subregion (C) (Figure 19b).

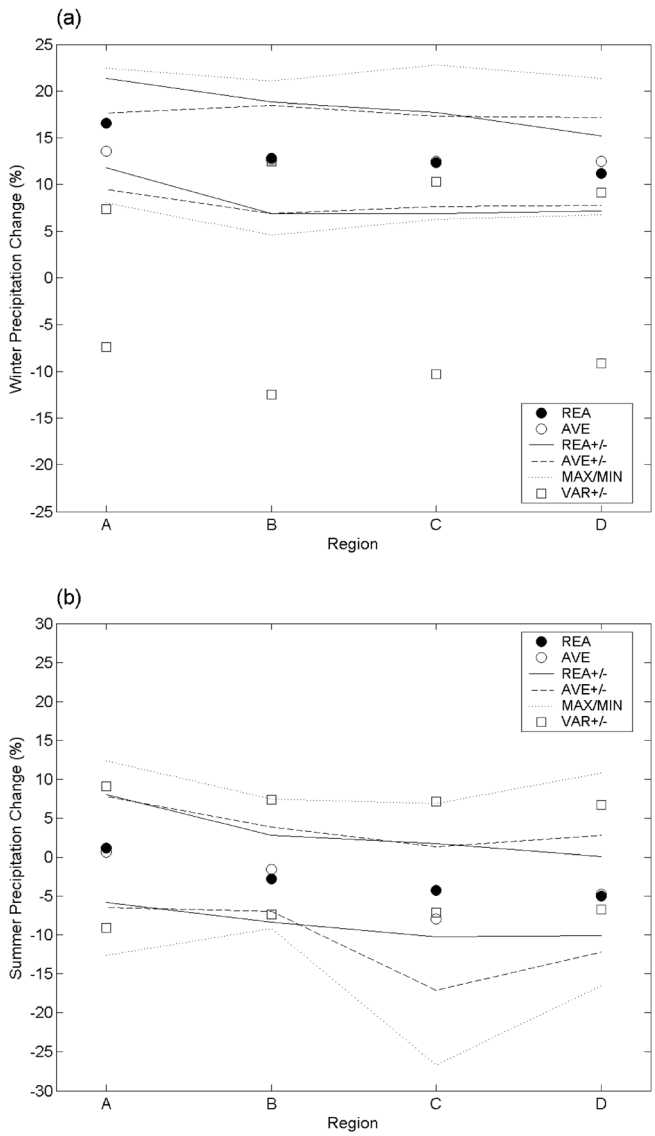

Figure 19. REA average changes (REA, dark circles) and corresponding upper and lower REA uncertainty limits (REA+/−, continuous lines), ensemble average changes (AVE, open circles) and corresponding uncertainty limits (AVE+/−, dashed lines), maximum and minimum changes simulated by individual model pairs in the ensemble (MAX/MIN, dotted lines), and estimated natural variability values (VAR, squares) for total precipitation rate over the four Northeast U.S. subregions during (a) winter and (b) summer seasons.

5. Conclusions

The availability of the model outputs from the NARCCAP Phase II simulations affords an opportunity to examine the model performance in reproducing current climate (1971–2000) and to evaluate the projections of future climate changes (2041–2070) at the regional scale over North America. Here we have focused on assessing surface air temperature and precipitation over the entire Northeast U.S. region. Although some model pairs may perform better than others in simulating current climate in terms of particular metrics over specific subregions, it is difficult to discern a "best performing" model that succeeds in reproducing the observed mean state and variability of temperature and precipitation for all four seasons across the entire domain.

The comparison of biases between the GCM-RCM pair and its driving GCM leaves two possibilities: (1) the sign of the RCM biases is in agreement with that of its driving GCM biases, thereby further increasing mean biases in the GCM-RCM pairing, or (2) the additional biases generated by the RCM may counteract the driving GCM biases to result in decreased combined biases in the GCM-RCM pairing. It is important to note that the mitigated combined biases do not necessarily imply notable improvements of the GCM-RCM pairing in reproducing current climate. The possibility of compensating errors can give the appearance of improvement, although the models may have serious issues with particular physical processes. The most prominent winter temperature biases are produced by GFDL-ECP2, especially over the southwestern subregion (−9.3°C). By providing boundary forcing, the moderate winter cold biases produced by the driving GFDL are inherited to the GFDL-ECP2 and GFDL-RCM3. Moreover, the stronger underestimation of winter temperature exhibited in GFDL-ECP2 and GFDL-RCM3 than their driving GFDL is ascribable to the additional cold biases generated by these two RCMs. The pronounced summer warm biases shown by the CCSM-driven models (i.e., CCSM-CRCM, CCSM-MM5, and CCSM-WRFG) over the western subregions may be greatly attributed to their driving model, which produces more intense overestimation of summer temperature. By counteracting the driving HadCM3 biases, dynamically downscaling HadCM3 with HadRM3 produces decreased combined biases of winter and summer mean temperature. Regarding precipitation mean biases, we found that winter wet biases are common among most of the model pairs, whereas different model pairs exhibit distinctive spatial patterns of summer precipitation biases. While the consistent winter wet biases are inherited from

the driving GCMs, different summer precipitation biases exhibited by multiple RCM simulations driven by the same GCM simulation indicate the dominance of mesoscale processes in summer precipitation. The consistency between the summer warm biases and dry biases exhibited by the three CCSM-driven model pairs implies that deficient precipitation and the associated drier land and less cloud cover lead to enhanced sensible heat flux from the surface and increased surface air temperature.

The projected future changes in temperature indicate overall warming, statistically significant at the one-tailed $P = 0.05$ confidence level in all nine individual model pairs during winter and summer. Winter warming is enhanced over the northern subregions, with ensemble averaged temperature increases exceeding 3°C. The increases in summer temperature are greater over the southern than northern subregions, and the maximum warming occurs over the southwestern subregion. Associated with warming trends in winter temperature, all the GCM-RCM pairs and their driving GCMs project consistently wetter winter conditions. In contrast to consistent and significant increases in summer temperature, different model pairs project quite different regional changes in summer precipitation. To complement the simple ensemble averaging method, the reliability ensemble averaging (REA) procedure was applied to estimate the temperature and precipitation changes based on both the model performance and model convergence criteria. The comparisons between the average changes and natural variability emphasize the robustness in the future warming, but the uncertainty in the direction of precipitation changes.

The wide variations in the current simulations and future projections of the Northeast U.S. climate among the NARCCAP GCM-RCM pairs point out the importance of further investigations of key processes and parameterizations in individual models. It is worth noting that the role of the temperature-dependent bias has not been estimated in this study. If the temperature bias slope increases with the degree of warming as suggested by *Boberg and Christensen* [2012], the projected temperature changes for the Northeast U.S. region may be overestimated. Despite the fact that statistically significant decreases in summer precipitation are projected by several model pairs over specific regions, the direction of summer precipitation changes is model-dependent. The inconsistency in the changing directions of summer precipitation between the GCM-RCM pairs and their driving GCMs indicates different changes in net moisture flux convergence-divergence. Exploring the root cause of different changes requires thorough understanding of issues related to the low-level jet and convective parameterization schemes and is beyond the scope of this study.

Acknowledgments
This research was supported by the Department of Interior's Northeast Climate Science Center, under USGS funding. We wish to thank the North American Regional Climate Change Assessment Program (NARCCAP) for providing the data used in this paper. NARCCAP is funded by the National Science Foundation, the U.S. Department of Energy (DOE), the National Oceanic and Atmospheric Administration, and the U.S. Environmental Protection Agency Office of Research and Development. NARCCAP data are available on the Earth System Grid data portal under the "NARCCAP" section (https://www.earth-systemgrid.org/project/NARCCAP.html). We also greatly appreciate Renping Lin for assistance in computing moisture transport from the model simulations.

References

Boberg, F., and J. H. Christensen (2012), Overestimation of Mediterranean summer temperature projections due to model deficiencies, *Nat. Clim. Change*, *2*, 433–436, doi:10.1038/NCLIMATE1454.

Bukovsky, M. S., and D. J. Karoly (2011), A regional modeling study of climate change impacts on warm-season precipitation in the central United States, *J. Clim.*, *24*, 1985–2002.

Christensen, J. H., et al. (2007), Regional climate projections, in *Climate Change 2007: The Physical Science Basis*, edited by S. Solomon et al., pp. 847–940, Cambridge Univ. Press, Cambridge, U. K., and New York.

Collins, W. D., et al. (2006), The Community Climate System Model version 3: CCSM3, *J. Clim.*, *19*, 2122–2143.

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

Fan, F., R. S. Bradley, and M. A. Rawlins (2014), Climate change in the northeastern US: Regional climate model validation and climate change projections, *Clim. Dyn.*, *43*, 145–161, doi:10.1007/s00382-014-2198-1.

Flato, G. M., G. J. Boer, W. G. Lee, N. A. McFarlane, D. Ramsden, M. C. Reader, and A. J. Weaver (2000), The Canadian Centre for Climate Modelling and Analysis global coupled model and its climate, *Clim. Dyn.*, *16*, 451–467.

Giorgi, F., and L. O. Mearns (2002), Calculation of average, uncertainty range, and reliability of regional climate changes from AOGCM simulations via the "reliability ensemble averaging" (REA) method, *J. Clim.*, *15*, 1141–1158.

Houghton, J. T., Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell, and C. A. Johnson (2001), *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, 881 pp., Cambridge Univ. Press, Cambridge, U. K.

Jones, P. D. (1994), Hemispheric surface air temperature variations: A reanalysis and update to 1993, *J. Clim.*, *7*, 1794–1802.

Jones, R. G., M. Noguer, D. C. Hassell, D. Hudson, S. S. Wilson, G. J. Jenkins, and J. F. B. Mitchell (2004), *Generating High Resolution Climate Change Scenarios Using PRECIS*, 40 pp., Met Office Hadley Centre, Exeter, U. K.

Knutti, R. (2008), Should we believe model predictions of future climate change?, *Philos. Trans. R. Soc.*, *366*, 4647–4664, doi:10.1098/rsta.2008.0169.

Mearns, L. O., et al. (2007), The North American Regional Climate Change Assessment Program dataset, National Center for Atmospheric Research Earth System Grid data portal, Boulder, Colo. [Available at http://www.earthsystemgrid.org/project/NARCCAP.html.]

Mearns, L. O., W. J. Gutowski, R. Jones, L. R. Leung, S. McGinnis, A. Nunes, and Y. Qian (2009), A regional climate change assessment program for North America, *Eos Trans. AGU*, *90*(36), 311–312, doi:10.1029/2009EO360002.

Mearns, L. O., et al. (2012), The North American Regional Climate Change Assessment Program: Overview of phase I results, *Bull. Am. Meteorol. Soc.*, *93*, 1337–1362, doi:10.1175/BAMS-D-11-00223.1.

Mearns, L. O., et al. (2013), Climate change projections of the North American Regional Climate Change Assessment Program (NARCCAP), *Clim. Change*, *120*, 965–975, doi:10.1007/s10584-013-0831-3.

Meehl, G. A., et al. (2007), Global climate projections, in *Climate Change 2007: The Physical Science Basis*, edited by S. Solomon et al., pp. 747–845, Cambridge Univ. Press, Cambridge, U. K., and New York.

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.

Nakićenović, N., et al. (2000), *Special Report on Emissions Scenarios: A Special Report of Working Group III of the Intergovernmental Panel on Climate Change*, 599 pp., Cambridge Univ. Press, Cambridge, U. K.

New, M., M. Hulme, and P. D. Jones (2000), Representing twentieth century space-time climate variability. Part 2: Development of 1901–96 monthly grids of terrestrial surface climate, *J. Clim.*, *13*, 2217–2238.

Noguer, M., R. G. Jones, and J. Murphy (1998), Sources of systematic errors in the climatology of a nested Regional Climate Model (RCM) over Europe, *Clim. Dyn.*, *14*, 691–712.

Pan, Z., J. H. Christensen, R. W. Arritt, W. J. Gutowski Jr., E. S. Takle, and F. Otieno (2001), Evaluation of uncertainties in regional climate change simulations, *J. Geophys. Res.*, *106*(D16), 17,735–17,751, doi:10.1029/2001JD900193.

Pope, V. D., M. L. Gallani, P. R. Rowntree, and R. A. Stratton (2000), The impact of new physical parameterizations in the Hadley Centre climate model: HadAM3, *Clim. Dyn.*, *16*, 123–146.

Randall, D. A., et al. (2007), Climate models and their evaluation, in *Climate Change 2007: The Physical Science Basis*, edited by S. Solomon et al., pp. 589–662, Cambridge Univ. Press, Cambridge.

Rawlins, M. A., R. S. Bradley, and H. F. Diaz (2012), Assessment of regional climate model simulation estimates over the northeast United States, *J. Geophys. Res.*, *117*, D23112, doi:10.1029/2012JD018137.