Assessment of regional climate model simulation estimates over the northeast United States

M. A. Rawlins,[1] R. S. Bradley,[1] and H. F. Diaz[2]

Received 18 May 2012; revised 19 October 2012; accepted 22 October 2012; published 12 December 2012.

[1] Given the coarse scales of coupled atmosphere-ocean global climate models, regional climate models (RCMs) are increasingly relied upon for studies at scales appropriate for many impacts studies. We use outputs from an ensemble of RCMs participating in the North American Regional Climate Change Assessment Program (NARCCAP) to investigate potential changes in seasonal air temperature and precipitation between present (1971–2000) and future (2041–2070) time periods across the northeast United States. The models show a consistent modest cold bias each season and are wetter than observations in winter, spring, and summer. Agreement in spatial variability and pattern correlation is good for air temperature and marginal for precipitation. Two methods were used to evaluate robustness of the mid 21st century change projections; one which estimates model reliability to generate multimodel means and assess uncertainty and a second which depicts multimodel projections by separating lack of climate change signal from lack of model agreement. For air temperature we find changes of 2–3°C are outside the level of internal natural variability and significant at all northeast grid cells. Signals of precipitation increases in winter are significant region wide. Regionally averaged precipitation changes for spring, summer, and autumn are within the level of natural variability. This study raises confidence in mid 21st century temperature projections across the northeast United States and illustrates the value in comprehensive assessments of regional climate model projections over time and space scales where natural variability may obscure signals of anthropogenically forced changes.

Citation: 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.

1. Introduction

[2] Current trajectories in greenhouse gas concentrations are increasing the likelihood of significant impacts from climate change in coming decades. A recent study using a multithousand-member perturbed-physics ensemble showed global-mean temperature increases of 1.4 to 3 K by 2050, relative to 1961–1990 under a mid-range forcing scenario [*Rowlands et al.*, 2012]. However, the geographic pattern of change is not uniform. Understanding both the magnitude and uncertainty in climate change projections at the regional scale is critical, as uncertainties in the regional climate response can lead to uncertainties in associated climate impacts [*Mearns*, 2003; *Wood et al.*, 2004].

[3] Model simulations suggest the potential for future temperature and precipitation changes across the northeast United States. By the end of the 21st century, warming can be expected across all seasons, and winter precipitation is projected to increase by 11 to 14% depending on the emission scenario used in the simulations [*Hayhoe et al.*, 2007]. Climate change studies have traditionally been performed using atmosphere-ocean general circulation models (AOGCMs) with resolutions of 100 to 400 km [*Alley et al.*, 2007]. These coarse scales leave AOGCMs unable to capture the effects of local forcings such as complex topography which modulates the models' climate signal at local scales. To overcome this shortcoming, downscaling of climate model simulations has become common and has been shown to provide valuable information for impacts research and adaptation planning [*Mearns et al.*, 2009; *Wood et al.*, 2004]. Downscaling can be achieved through either dynamical or statistical methods. Statistical downscaling typically involves relating large scale climate features to local climate at a particular location. Dynamical downscaling involves the application of a regional climate model (RCM) forced by global climate model boundary conditions. More realistic parameterizations of surface processes in sophisticated RCM land-surface schemes can provide more realistic simulations of surface conditions and the associated physical processes critical to the simulation of climate extremes [*Roy et al.*, 2011]. Large international efforts such as PRUDENCE

[1]Climate System Research Center, Department of Geosciences, University of Massachusetts, Amherst, Massachusetts, USA.

[2]Climate Diagnostics Center, Office of Atmospheric Research, National Oceanic and Atmospheric Administration, Boulder, Colorado, USA.

Corresponding author: M. A. Rawlins, Climate System Research Center, Department of Geosciences, University of Massachusetts, 611 North Pleasant St., Amherst, MA 01002, USA. (rawlins@geo.umass.edu)

0148-0227/12/2012JD018137

(http://prudence.dmi.dk/), ENSEMBLES (http://ensembles-eu.metoffice.com/), and CORDEX have made use of high resolution downscaled model data for projections of future climates at regional scales [*Giorgi et al.*, 2009].

[4] The North American Regional Climate Change Assessment Program (NARCCAP) [*Mearns et al.*, 2007] is producing high resolution climate change simulations from different combinations of AOGCMs and RCMs, providing a rich set of data that facilitate investigations of uncertainties in regional scale projections of future climate across North America [*Mearns et al.*, 2009]. *Mearns et al.* [2012] describe results from Phase I, an evaluation component of the program, wherein the RCMs are nested within NCEP/DOE global reanalysis II. Their comparisons with observed data for the period 1980–2004 show that, for air temperature across North America, both positive (model overestimates) and negative biases occur. The Hadley Regional Climate Model 3 (HRM3) exhibits a large warm bias in winter and summer. Both positive and negative biases are noted with summer precipitation. For winter, precipitation biases are largely positive. Phase I results showed relatively low temperature and precipitation biases over the northeast. The RCMs replicate well the monthly frequency of precipitation extremes across coastal California, where the precipitation is largely topographic, while performing somewhat less well across the Upper Mississippi River watershed [*Gutowski et al.*, 2010]. NARCCAP data have been used to characterize potential future increases in the intensity of extreme winter precipitation across the western United States [*Dominguez et al.*, 2012] and to produce projections of seasonal climate across the southeast United States [*Sobolowski and Pavelsky*, 2012]. Simulations with the Abdus Salam Institute Theoretical Physics Regional Climate Model Version 3 (RegCM3), one of models used by NARCCAP, suggest an increased frequency of extreme hot events and a decreased frequency of extreme cold events across much of the northeast, by late century [*Diffenbaugh et al.*, 2005]. This RCM also simulated a future increases in mean annual and extreme precipitation event frequency. NARCCAP RCM simulations and data from each respective driving GCM were used to construct probabilistic projections of high-resolution monthly temperature over North America [*Li et al.*, 2012].

[5] Detailed assessments of individual model errors, often termed 'biases', are critical to determining model usefulness for understanding potential impacts of climate change. The ability to simulate the climate of a region depends largely on the quality of the forcing AOGCM and the degree to which it represents flow conditions at the boundary [*Christensen et al.*, 1998; *Giorgi et al.*, 2001]. For example, biases of only a few models can affect the multimodel mean and result in physically unrealistic results. One analysis [*Liepert and Previdi*, 2012] found that many of the studied AOGCMs had an unphysical and hence 'ghost' sink or source of atmospheric moisture. Several recent studies shed light on the nature of uncertainty in climate change projections. For instance, *Hawkins and Sutton* [2011] in an analysis of the CMIP3 multimodel ensemble found that for decadal means of seasonal mean precipitation, internal variability is the dominant uncertainty for predictions of the first decade everywhere, and for many regions until the third decade ahead. Model uncertainty is generally the dominant source of uncertainty for longer lead times. In an earlier review [*Hawkins and Sutton*, 2009] showed how the different contributions to climate projection uncertainty vary with lead-time over the 21st century (see their Figures 2 and 3). Beyond about 20 years, model uncertainty becomes greater than internal variability, and of course, emission scenario uncertainty becomes dominant after about mid-century. Most recently, *Deser et al.* [2012] found that the dominant source of uncertainty in the simulated climate response using a 40-member simulation ensemble with the NCAR Community Climate System Model Version 3 (CCSM3) under the SRES A1B for middle and high latitudes is internal atmospheric variability, and that uncertainties in the forced response are generally larger for sea level pressure than precipitation, and smallest for air temperature. Thus the implication from that study is that forced changes in air temperature can be detected earlier and with fewer ensemble members than those in atmospheric circulation and precipitation. The availability of multimodel simulations has helped to focus efforts on new approaches to synthesize climate change projections [*Giorgi and Mearns*, 2002; *Knutti et al.*, 2010; *Tebaldi and Knutti*, 2007; *Tebaldi et al.*, 2011]. This includes methods which weight models based on performance relative to present-day conditions and/or the deviation from the group mean [*Giorgi and Mearns*, 2002, 2003]. *Christensen et al.* [2010] describe a weighting scheme which incorporates six model performance metrics. *Tebaldi et al.* [2011] argue that assessments using multiple climate models should separate lack of climate change signal from lack of model agreement by assessing the degree of consensus on the significance of the change as well as the sign of the change.

[6] In this study we describe the sign, magnitude, and quantitative significance of precipitation and temperature changes across the northeast United States between the periods 2041–2070 and 1971–2000. We apply a method designed for calculating average, uncertainty range, and a measure of reliability of simulated climate changes at the regional scale from ensembles of different climate model simulations. A second method, which complements the first, is used to account for model performance and natural variability and, in turn, determine best estimates of likely changes by mid-century. The climate change analysis follows an assessment of model performance relative to fields derived from observed station data. Investigating the ability of the suite of RCMs to capture the magnitude and variability in current climate provides additional information on their potential for improving understanding of regional scale climate change impacts across the northeast United States.

2. Data and Methods

2.1. Model Data

[7] The NARCCAP [*Mearns et al.*, 2007] is archiving outputs from a set of regional climate model (RCM) simulations over a domain spanning North America. For the NARCCAP effort, each participating RCM is forced with boundary conditions from at least two atmosphere-ocean general circulation models (AOGCMs). Table 1 list the models and the respective modeling centers. Three hourly RCM outputs are available for the contemporary period 1971–2000 and for the future period 2041–2070. The NARCCAP effort involves the use of the SRES A2 emissions scenario [*Nakicenovic et al.*, 2000] by all modeling

Table 1. Models Used in This Study

Global Model	Model Center
CCSM	National Center for Atmospheric Research
CGCM3.1	Canadian Centre for Climate Modeling and Analysis, Canada
GFDL	Geophysical Fluid Dynamics Laboratory, USA
HadCM3	Hadley Centre for Climate Prediction and Research / Met Office, UK
Regional Model	**Model Center**
CRCM	OURANOS / UQAM, Canada
ECP2	UC San Diego / Scripps Institute of Oceanography, USA
HRM3	Hadley Centre for Climate Prediction and Research / Met Office, UK
MM5	Iowa State University, USA
RCM3	UC Santa Cruz, USA
WRFG	Pacific Northwest National Lab, USA

groups. Use of the high mid-century greenhouse gas concentrations of the A2 scenario may not be unreasonable given recent greenhouse gas concentration trajectories (P. Tan and R. Keeling, Trends in atmospheric carbon dioxide, 2012, Scripps Institution of Oceanography (scrippsco2.ucsd.edu/)) and the fact that the commonly used A1B scenario closely tracks A2 through mid century. We use 2 m air temperatures and precipitation data for a subset of GCM-RCM pairs available at the time of this writing. Table 2 shows the currently available model pairs among all planned NARCCAP combinations. When discussing a model simulation we use the convention GCM_RCM, for example CCSM_MM5. Spatial resolution for each model is approximately 50 km and each RCM has its own native grid. For this study we derived seasonal means and totals from the archived 3 hourly data. To facilitate the analysis we interpolated all native data values to a common 0.5 degree grid. Our analysis includes all 0.5 degree grid cells which fall within the 9 northeast U.S. states. A more complete discussion of the NARCCAP project is presented in *Mearns et al.* [2009].

2.2. Observed Data

[8] Bias assessments were made by comparing estimates from NARCCAP GCM-RCM pairings with data representing air temperature and precipitation observations from station records. We use monthly 2 m air temperatures and precipitation on a 0.5 degree grid available from the University of Delaware (UDel) database (K. Matsuura and C. J. Willmott, Terrestrial air temperature: 1900–2008 gridded monthly time series, version 2.01, 2009, http://climate.geog.udel.edu/~climate/; C. J. Willmott and K. Matsuura, Terrestrial precipitation: 1900–2008 gridded monthly time series, version 2.01, 2009, http://climate.geog.udel.edu/~climate/). The UDel data set was developed through interpolations of meteorological station data which account for the lapse rate in temperature with increasing elevation [*Willmott and Matsuura*, 1995] and makes use of spatially high-resolution air temperature and precipitation climatologies [*Willmott and Robeson*, 1995]. Monthly values over the period 1971–2000 are used here to construct fields of seasonal mean air temperatures and precipitation at each grid cell over the northeast.

2.3. Analysis

[9] This paper presents an assessment of biases and uncertainties in seasonal air temperatures and precipitation fields among the available NARCCAP GCM-RCM pairs. We focus on mean climate across the northeast United States for each season over the two periods, hereafter the "present" (1971–2000) and "future" (2041–2070). Seasons are defined using monthly values as follows: winter (DJF), spring (MAM), summer (JJA), and autumn (SON). Through this analysis we quantify regional averages from the models and observations and describe changes in the projected mean seasonal climate. We also examine biases between RCM and observed data fields and changes in statistical properties of the fields.

[10] Uncertainties in future climate projections can be estimated through use of not only multimodel ensembles, but also through application of statistical methods which take into account the natural climate variability (ϵ) and the performance of individual models in relation to the ensemble group. To improve our assessments we apply here the reliability ensemble averaging (REA) method [*Giorgi and Mearns*, 2002] to determine average, uncertainty range, and a measure of reliability of simulated changes from the ensemble of available NARCCAP GCM-RCM pairings. *Sobolowski and Pavelsky* [2012] applied the REA method together with NARCCAP data to estimate likely future air temperature and precipitation changes across the southeast United States.

[11] The REA method defines a change over two time periods as a weighted average of ensemble climate model members. Here each climate model value is the regionally averaged seasonal temperature. Two model reliability factors contribute to the weighting for each model; a factor based on a model's ability to reproduce current climate and a factor based on the distance of the model's change estimate from the REA average. A simple multimodel mean of ensemble members for, as an example, season temperature T is

$$\overline{\Delta T} = \frac{1}{N}\sum_{i=1,N}\Delta T_i \tag{1}$$

where N is the number of models, the overbar indicates ensemble averaging, and Δ indicates the simulated change.

Table 2. Regional Climate Model (RCM) and Forcing Atmosphere-Ocean General Circulation Model (AOGCM) Combinations Used in This Study[a]

	Regional Models					
Global Models	CRCM	ECP2	HRM3	MM5	RCM3	WRFG
CCSM	1			2		3
CGCM3	4				5	6
GFDL		7	X		8	
HADCM3		X	9	X		

[a]We use model pairings that have data for both historical and future periods. Numbers are references for model pairs examined. An 'X' indicates a model pair for which data are being produced, but are unavailable at time of writing. See section 2.

With the REA method, the average change $\widetilde{\Delta T}$ is a weighted average of the ensemble members

$$\widetilde{\Delta T} = \widetilde{A}(\Delta T) = \frac{\sum_i R_i \Delta T_i}{\sum_i R_i} \quad (2)$$

where the operator $\widetilde{A}$ indicates REA averaging and R_i is a model reliability factor, defined

$$R_i = \left[\left(R_{B,i}\right)^m \times \left(R_{D,i}\right)^n \right]^{[1/(m \times n)]} \quad (3)$$

The factor $R_{B,i}$ is a measure representing model reliability as a function of model bias ($B_{T,i}$) in simulating contemporary temperature. The factor $R_{D,i}$ is a measure which reflects model reliability in terms of the distance ($D_{T,i}$) of the change calculated by a given model from the REA average. Parameters m and n are user defined weights for each reliability factor. Here we choose a value of 1 for each weight. The uncertainty range around the REA change is estimated using the root-mean square difference (rmsd) of the changes, $\tilde{\delta}_{\Delta T}$. The total uncertainty range is $\pm\tilde{\delta}_{\Delta T}$ or $2\tilde{\delta}_{\Delta T}$. Natural variability, ϵ_T, for regionally averaged seasonal air temperature (ϵ_P for precipitation) is estimated using the UDel data observed fields. For temperature, the 100 year time series of regionally averaged seasonal values is de-trended and then smoothed using a 30 year running mean. The difference between the maximum and minimum values in the 100 year smoothed series becomes ϵ_T. Natural variability estimation and the other details of the REA method are described in *Giorgi and Mearns* [2002].

[12] More recently, *Tebaldi et al.* [2011] introduced a method appropriate for studies involving multiple models. The method separates lack of signal from lack of information due to model disagreement. It accomplishes this objective by assessing the degree of consensus on the significance of the change as well as the sign of the change. With this method, weather noise in the simulations is the variability over which significance in the model change signal is compared. Here, grids where at least 7 of the 9 models (X = 78%) show significant change and at least Y = 80% of those models agree on sign of change are shaded by magnitude and stippled. Grids are shaded by the multimodel mean but not stippled where less than 7 of the models show significant change. It is in these areas that *Tebaldi et al.* [2011] argue that a signal of change lies within the noise of weather variability, and that the models still contain useful information. Areas where at least 7 of the models show significant change but less than 80% of those model agree in sign are left unshaded. For these grids the models are presenting conflicting information. Hence, confidence there is low. The above X and Y percentages are a subjective choice and can be set differently based on the desired level of confidence. Our choices for X and Y are similar but not identical to the 66% and 90% levels adopted in the Intergovernmental Panel on Climate Change's Fourth Assessment Report [*Alley et al.*, 2007]. We apply both techniques described by *Giorgi and Mearns* [2002] and *Tebaldi et al.* [2011] to the NARCCAP model suite in order to answer questions regarding the magnitude and significance of projected climate changes across the northeast United States. Narrowing down the uncertainty range in climate model projections with the help of observations is an important challenge in climate research [*Mearns et al.*, 2012]. As *Lenderink* [2012] points out, while "it makes sense to weight model results according to the model's ability to represent pertinent aspects of the observed present-day climate … such an evaluation of simulations is far from trivial", and furthermore, while "… a model whose simulations of the present-day climate are close to observations [it] may well contain a set of errors that compensate each other today, but may strongly distort the response to climate warming as the balance between errors changes."

3. Results and Discussion

3.1. Air Temperature

[13] The ability of models to reproduce contemporary climate is important, as large biases limit our confidence in projections of future climate impacts such as extreme weather events. Regional climate models should capture well the spatial variability in climatic fields along with the regionally averaged mean climate. For seasonal temperatures (1971–2000), the models as a group perform well in capturing spatial variability across the northeast United States as measured by standard deviation in the observed and model fields. Figure 1 shows Taylor diagrams which capture statistics of the observed and model seasonal temperature field across the 211 northeast grid cells. The models tend to exhibit spatial variability which is similar to the observational fields, as reflected by the agreement in standard deviations. The models slightly underestimate spatial variability in winter. Grid cell correlations are also relatively high, generally greater than 0.85, and highest in winter.

[14] Across the northeast United States, the NARCCAP models tend to underestimate seasonal temperatures (Figure 2). No strong spatial pattern exists in the temperature bias fields. Biases are significant (95% confidence level based on student t test) for 88, 62, 73, and 87% of the grid cells in winter, spring, summer, and autumn, respectively. The outlier over western Maine each season is a grid cell with a large amount of open water. An anomaly with the grid multimodel mean 2 m temperature appears to be the source of the temperature bias. Regional mean bias are −2.1, −1.8, −1.9, −2.3°C for winter, spring, summer, autumn seasons, respectively (Table 3). Individually, all model pairings underestimate spatially averaged 1971–2000 mean temperature in summer and autumn (Figure 3). All but two model pairs (CCSM_MM5 and CGCM3_WRFG) underestimate temperature in winter, while all but two (CCSM_MM5 and CCSM_WRFG) underestimate it in spring.

[15] Boxplots of seasonal air temperature distributions for the gridded fields of present observations (O) and NARCCAP models present (P) and future (F) are shown in Figure 4. Comparing the distribution for the future period alongside those for the present provides a sense of how projected future changes compare with model biases. It also reveals if and how temperatures might change, along with other properties of the distribution. Temperature distributions are nearly normal. In autumn, for example, while 25% of the grid cells have observed temperature exceeding 10.9°C, none of the RCM (multimodel mean) grid cells exceed that value. In summer a similarly shaped distribution is biased colder by

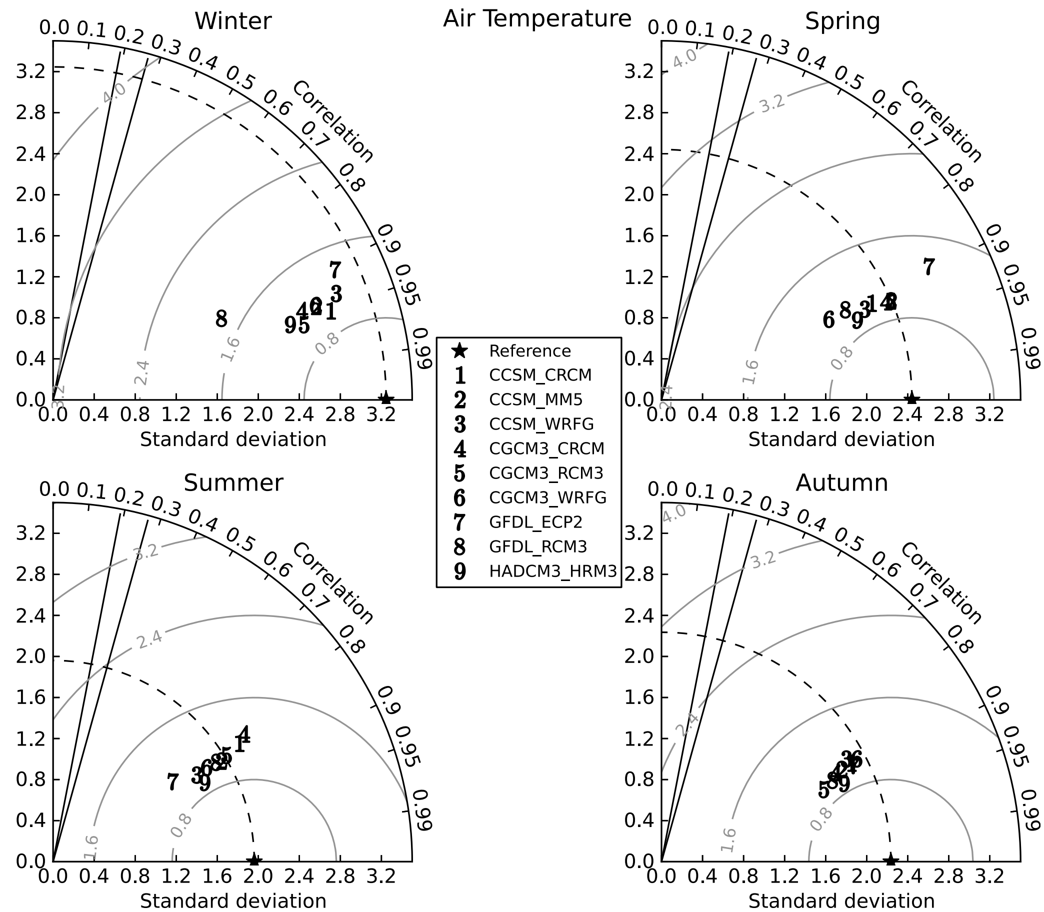

Figure 1. Taylor diagrams showing standard deviation (°C), RMSE (°C), and correlation for the observed and for each model seasonal (1971–2000) air temperature field across the northeast United States. Seasons throughout the analysis and in the subsequent figures are define winter (DJF); spring (MAM); summer (JJA); autumn (SON). GCMs and RCMs are listed in Table 1 and Table 2. Statistics are calculated over all 211 grid cells of the observed field and from the nine GCM-RCM fields. The star indicates the statistics for the observed field. Contour of the reference standard deviation (from observed field) is show by the dashed line. RMSE contours are in gray. Correlation rays are the (left) 95th and (right) 99th significance levels.

approximately 2°C. No change occurs with the shape of temperature distributions between the present and future periods. Comparing the multimodel means for the present and future periods shows seasonal temperature increases of 3.0, 2.0, 2.6, 2.9°C in winter, spring, summer, and autumn, respectively (Figure 4). Thus, based on simple multimodel means for each season, the mean change exceeds the mean bias. The greater winter changes are consistent with expected global trends assumed to be related to the ice-albedo feedback [*Dickinson et al.*, 1987; *Hall*, 2004]. A recent analysis of projected changes for the period 2035–2064 (relative to 1961–1990) using nine AOGCMs and the A2 scenario [*Hayhoe et al.*, 2007], however, found lower change magnitudes across the northeast and, interestingly, a larger projected increase in summer than winter. The study authors speculated that the larger change in summer versus winter was attributable to feedbacks between evaporation and temperature along with a declining effect from the ice-albedo feedback.

[16] Applying the REA method [*Giorgi and Mearns*, 2002] provides additional information regarding uncertainties in the magnitude of future seasonal temperature changes (Figure 5). First, each REA average change differs from the ensemble average change by a few tenths of a degree C, with winter temperature changes across the northeast lower by 0.3°C (2.7 versus 3.0°C). In comparison, natural variability (ϵ_T) is 0.5°C or less in spring, summer, and autumn, and approximately 1°C in winter (Table 3). Thus, the magnitude of change each season is well outside the range of natural variability. The uncertainty range calculated from the REA method ($\pm\delta_{\Delta T}$) is largest in winter and smallest in spring (Table 3).

[17] Results from the model agreement mapping method [*Tebaldi et al.*, 2011] reveal interesting differences across the region. Temperature changes are significant (95% level) over the entire region in each season (Figure 6). As noted above, the ensemble mean change is largest in winter, where the projected temperature increase exceeds 3°C across more than

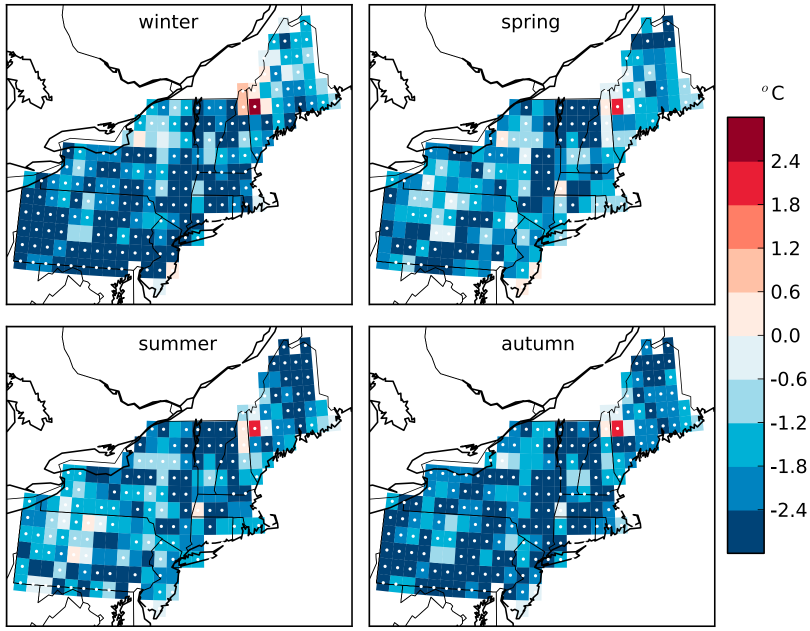

Figure 2. Bias (°C) in multimodel mean estimate of seasonal 2 m air temperature (1971–2000) at northeast United States grid cells. Bias is defined T_{models} - T_{UDel}, where T_{models} is the multimodel mean air temperature for the grid. Stipling indicates a significant bias at 95% confidence level based on student t test. Data for observations are taken from the UDel 0.5 degree data set.

60% of the region. Spatially, the largest changes are found across northern areas. In summer the pattern is reversed, with changes approaching 3°C across the southwest and declining northward. Relatively smaller changes are found across coastal areas in summer. It is clear that across the northeast United States, the anthropogenically forced signal of change by mid-century rises above the level of internal weather noise for each season. The pattern of winter temperature change suggests a reduction in cold air outbreaks affecting the northern part of the region and a reduction in snow cover. For summer the greater increases in temperature toward the south suggest increased warm air inflow from the continental interior.

3.2. Precipitation

[18] In contrast to the performance for temperature, RCM estimates for precipitation show lower agreements with the observed fields. Precipitation correlations for many of the GCM-RCM pairings (Figure 7) are much lower than those for air temperature (Figure 1) and in many cases are not significant. Spatial correlations are highest in winter. This result is consistent with studies which suggest that climate models are better at capturing stratiform than convective precipitation [*Mearns et al.*, 1995; *Christensen et al.*, 1998; *Giorgi et al.*, 1998; *Frei et al.*, 2003]. With the exception of summer, standard deviations of the seasonal model fields are consistently low relative to the observed field, illustrating the more narrow precipitation distributions generated by the models.

[19] Biases in multimodel mean seasonal precipitation range from overestimates to underestimates, with a noted spatial pattern across the region; lower biases toward the coast, and greater biases inland (Figure 8). All biases are significant. Biases tend to be higher across the northwest part of the domain and decrease to the southeast. With the exception of autumn, biases are generally positive, i.e. model precipitation exceeds observed precipitation. In autumn the largest negative biases occur across the east coast. The reason for these negative biases is unclear and warrants further investigation, particularly into the question of whether the models adequately capture tropical and/or extratropical cyclone formation and progression through the region. Individually, most model pairs overestimate precipitation in winter, spring, and summer (Figure 9). The CCSM_WRFG model shows the lowest estimates and is the only model which underestimates precipitation each season.

[20] Seasonal precipitation distributions illustrate the reduced spatial variability exhibited by the RCMs compared to the observed data, particularly in spring (Figure 10), for

Table 3. Ensemble Mean Bias, Projected Change (2041–2070 Minus 1971–2000), Uncertainty Range, and Estimated Natural Variability for Air Temperature (T) and Precipitation (P) Each Season Across the Northeast as an Average Across the Nine GCM-RCM Pairings[a]

	Temperature				Precipitation			
	T bias	T Change	$\pm\delta_{\Delta T}$	ϵ_T	P bias	P Change	$\pm\delta_{\Delta P}$	ϵ_P
Winter	−2.1	3.0[2.7]	0.5	1.1	18	13[12]	7.0	12
Spring	−1.8	2.0[2.1]	0.2	0.3	16	9[3]	7.9	6
Summer	−1.9	2.6[2.7]	0.4	0.5	12	−2[−3]	6.1	8
Autumn	−2.3	2.9[2.7]	0.2	0.4	−2	3[2]	6.5	6

[a]Temperature changes are in °C and precipitation changes are in percentage of present model amounts. Values in brackets are the regional averages from the REA method.

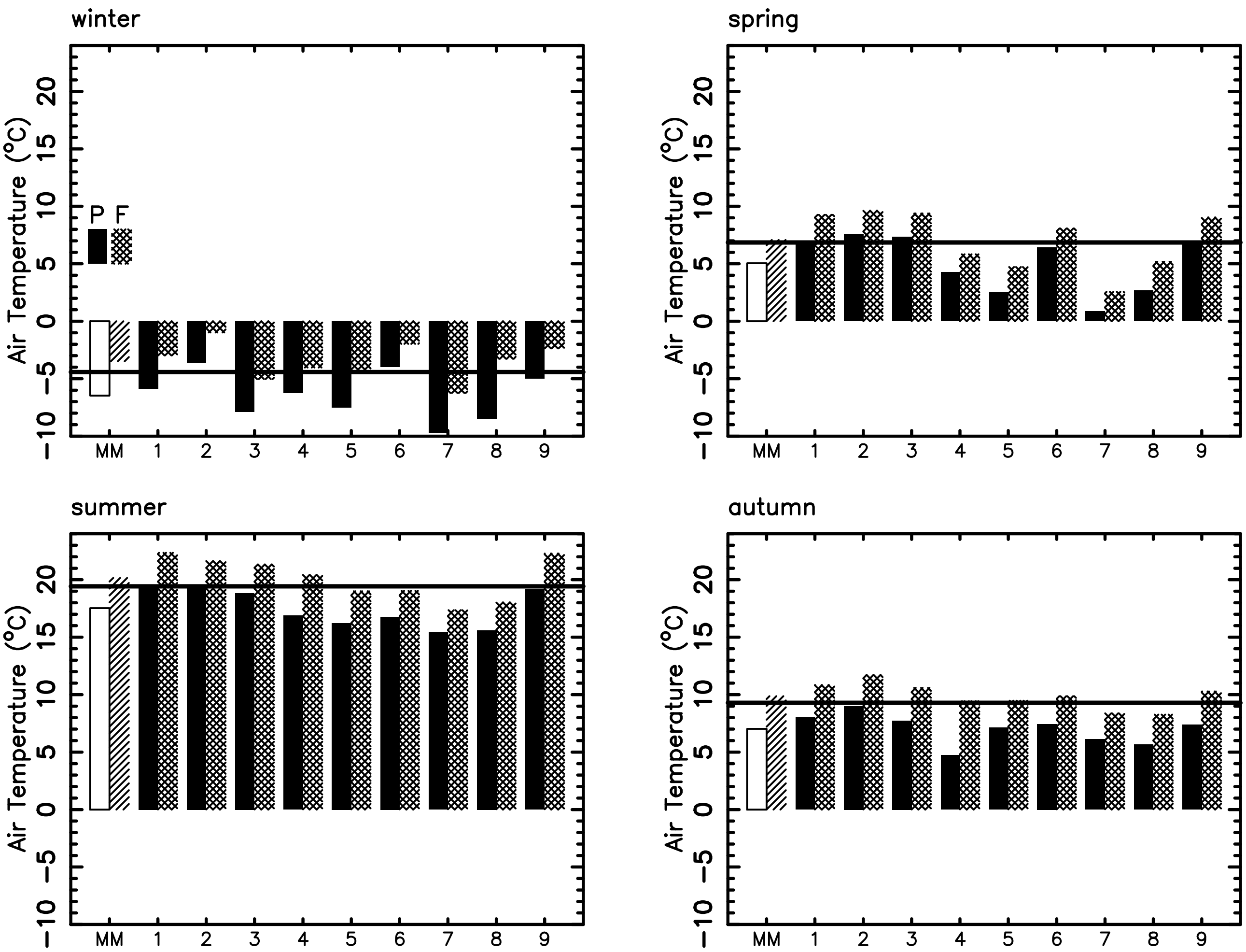

Figure 3. Seasonal mean air temperature (°C) averaged across the northeast United States for each of the nine GCM-RCMs pairs and the multimodel mean over the present period 1971–2000 (solid rectangles) and future period 2041–2070 (hatched). The multimodel means are simple averages of the nine model estimates (no weighting). The value for seasonal air temperature from the observed field is indicated by the horizontal line. Models 1–9 are listed in Table 2.

which 90% of the RCM grid estimates span a range of 72 mm season^{-1} (280 to 352 mm season^{-1}), while 90% of the observed grids span 116 mm season^{-1} (214 to 331 mm season^{-1}). No significant change is evident in the shape of future precipitation distributions relative to the present distributions. For winter many of the individual model pairs project wetter conditions by mid-century (Figure 9). No consensus exists in the other seasons. While a change across the winter and spring multimodel distribution is apparent (Figure 10), the change in the mean (future - present) is less than the mean (present - obs) bias, raising questions as to the robustness of the projected seasonal precipitation changes across the northeast United States.

[21] For precipitation, the difference between the REA average change and ensemble average change is less than 2% in winter, summer, and autumn (Figure 11 and Table 3). The REA average change is 6% lower than the ensemble average change in spring. Regarding uncertainty in the future precipitation change projections, the REA average changes are well within the bounds of natural variability (ϵ_P) in spring, summer, and autumn, and in winter both the REA change and natural variability are comparable, with values around 12% of present-day precipitation (Table 3). The REA uncertainty range ($\pm\delta_{\Delta P}$) is approximately 6–8% around the mean change each season.

[22] Applying the model agreement mapping method reveals differing signs and magnitudes of change in both space and time. It also provides important information regarding our confidence in anticipated future trends. For winter, the majority of models agree on precipitation increases across the entire northeast, with the change magnitudes highest (>15%) over interior areas and lowest along the coast (Figure 12). The changes are statistically significant for all grids. As noted above, the multimodel mean winter change from the REA method is ~12% (Table 3). For spring changes are positive and significant across most of the region. The exception being the state of Maine and the eastern half of Pennsylvania where white shading indicates that several of the models simulate significant changes of opposite sign. Where there is model agreement future changes approach 10% in the lee of Lake Ontario. For summer, precipitation changes are soon significant across the southwest

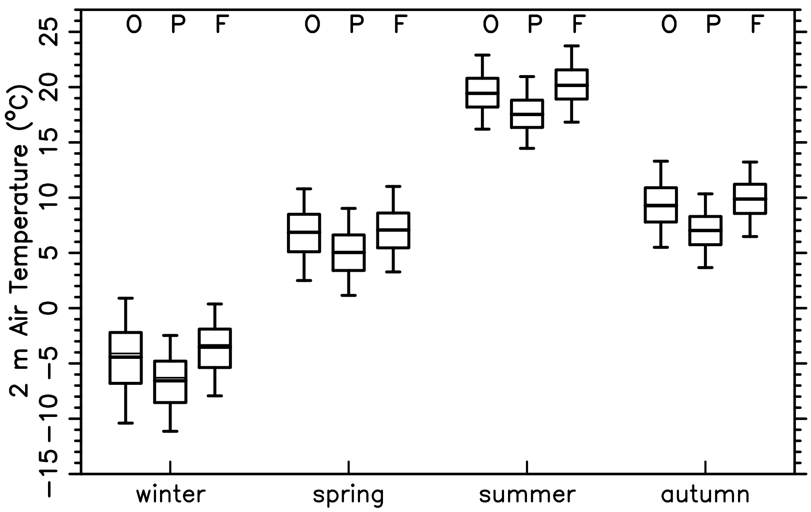

Figure 4. Distributions of air temperature (°C) for the observed (O) and RCM present period (P) fields for period 1971–2000, and for the future period (F). Each distribution consists of 211 0.5 degree grid cells spanning the northeast United States. Heavy line in each box is the distribution mean. Thin line (nearly identical to mean in most cases) is the distribution median. Boxes bracket the 25th and 75th percentiles. Whiskers show the 5th and 95th percentiles.

section of the northeast—much of Pennsylvania—where the projected declines locally approach −10%. Across northern areas the colored grids depict small change magnitudes, both positive and negative, that are within the level of natural variability. A similar result is apparent for autumn when the models agree, over 90% of the region, that the signal is small (positive change of around 1–5%) and has not emerged from the noise.

4. Discussion and Conclusions

[23] Correlations between the observed and RCM gridded seasonal temperature fields across the northeast are generally good, and the models closely capture the standard deviation present in the observed field. This suggest that the NARCCAP RCMs adequately represent the spatial variations across the northeast United States. Cold biases, however, are common, with ensemble mean biases of approximately −2°C each season. It is not clear if this is largely attributable to a cold bias in the forcing AOGCMs or to physical parameterizations in the RCMs. Largest model biases range from +4°C to −2°C and +10°C to −2°C in winter and summer, respectively. Investigations of biases in the NARCCAP RCMs is ongoing. Studies which quantify errors in both model parameterizations and AOGCM forcing data sets will improve assessment efforts.

[24] Temperature distributions of the observed and present RCM (ensemble mean) fields are nearly normal, as is the future ensemble mean field. The REA average change

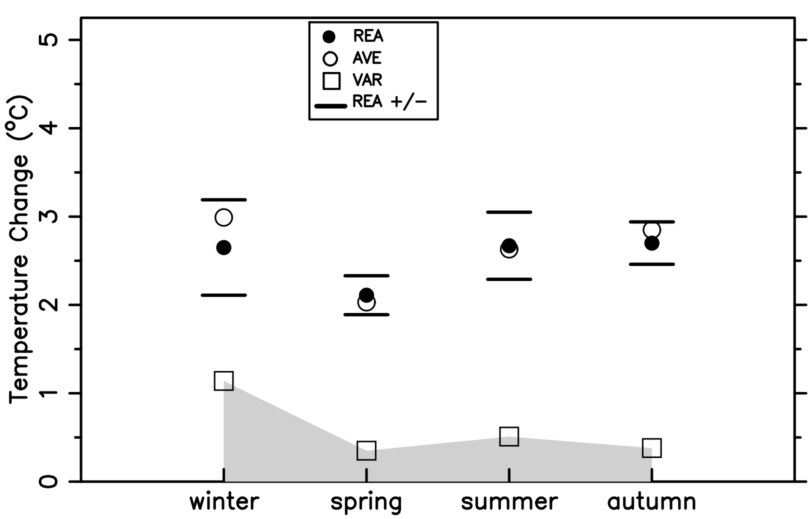

Figure 5. REA change in seasonal 2 m air temperature (°C, 2041–2070 minus 1971–2000) across the northeast United States (solid circles); corresponding upper and lower REA uncertainty limits [horizontal lines]; ensemble average changes (open circles); and estimated natural variability values (squares).

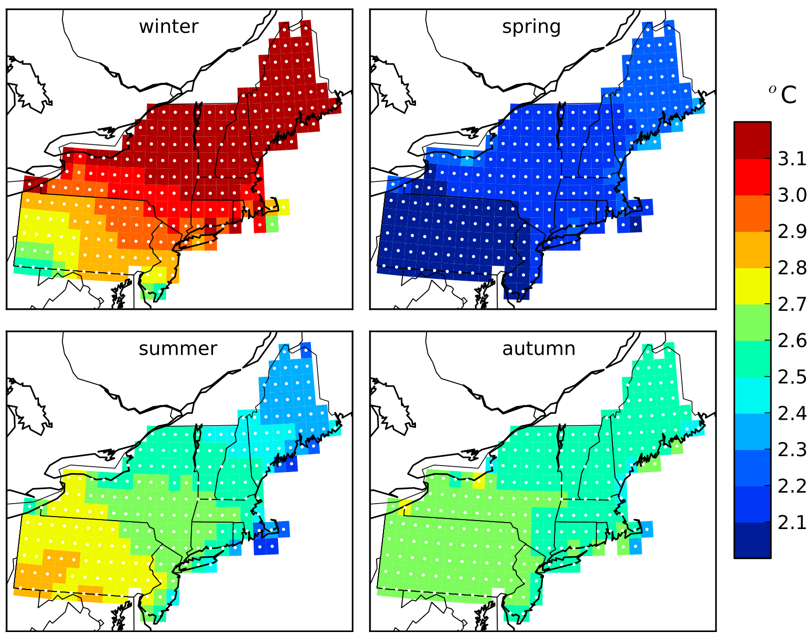

Figure 6. Change (°C, 2041–2070 minus 1971–2000) in seasonal temperature from the ensemble mean of the nine model pairs. Significance determined following criteria described by *Tebaldi et al.* [2011]. Changes are significant at the 95% level for all grids (shown stippled) across the northeast. See text for details on meaning behind the uncertainty and significance logic.

differs from the ensemble average change by only a few tenths of a degree C. This is likely a result of no large outliers among the seasonal model temperatures (Figure 3). Change magnitudes are more than double the level of natural variability.

[25] Results from the model agreement mapping method lends confidence to the mean projected mid-century seasonal temperature changes and their spatial patterns. Changes exceed the 95% confidence level at all grid points each season. We speculate that greater changes across more northern areas in winter are attributable to losses in snow cover through the ice-albedo feedback. Winter changes exceed 3°C over more than half of the northeast. The winter regional average 3°C increase here contrasts with a recent study of nine IPCC AOGCM forced with the A2 scenario which found a projected ensemble average increase (2035–2064 minus 1961–1990) across the northeast of less than 2°C [*Hayhoe et al.*, 2007].

[26] Our assessment of precipitation biases and uncertainties in the future change projections stand in contrast with those for air temperature. Correlations between the observed and RCM precipitation fields are low and often insignificant. The models also tend to underestimate spatial variability, thus the likelihood that spatial patterns of change will be close to those shown in this study is higher for seasonal temperature than seasonal precipitation, as noted below. The frequent inability of climate models to simulate regional precipitation patterns is due to inherent smaller spatial scales of variability in precipitation compared to air temperature. Precipitation decorrelation scales are at least one order of magnitude smaller. Simulations in summer are more locally controlled than for other seasons [*Plummer et al.*, 2006]. In turn, winter precipitation is controlled more by the large-scale flow. Summer precipitation is more closely tied to model parameterizations and winter precipitation to the driving data [*Caya and Biner*, 2004]. We note that precipitation biases are generally positive (model overestimates) and exhibit a southeast to northwest gradient across the region. The regionally averaged bias is negative in autumn, with a mean bias of −2% but local differences in excess of −20%. NARCCAP Phase I simulations with forcing from NCEP DOE II reanalysis show generally small biases across the northeast, relative to other regions of North America [*Mearns et al.*, 2012]. Studies focused on RCM precipitation processes are needed to determine how well tropical and extra-tropical cyclones, which transport considerable moisture into the northeast in autumn, are simulated. Mean seasonal biases for both precipitation and air temperature are similar when the RCM estimates are compared with Global Historical Climatology Network data (http://www.ncdc.noaa.gov/temp-and-precip/ghcn-gridded-products.php). This suggests that the UDel data is adequate for assessing model performance.

[27] The ensemble mean precipitation change estimated using the REA method is approximately equal to natural variability in winter, while changes in other seasons fall within the estimated range of variability. Thus, the REA method results suggest a lower level of confidence in the future precipitation changes for spring, summer, and autumn, but modest confidence for winter. In essence, the regional, mid 21st century winter change signal from a weighted multimodel mean can just be detected above the noise of weather variability. For the remaining seasons, changes are within the level of natural variability. Under the model

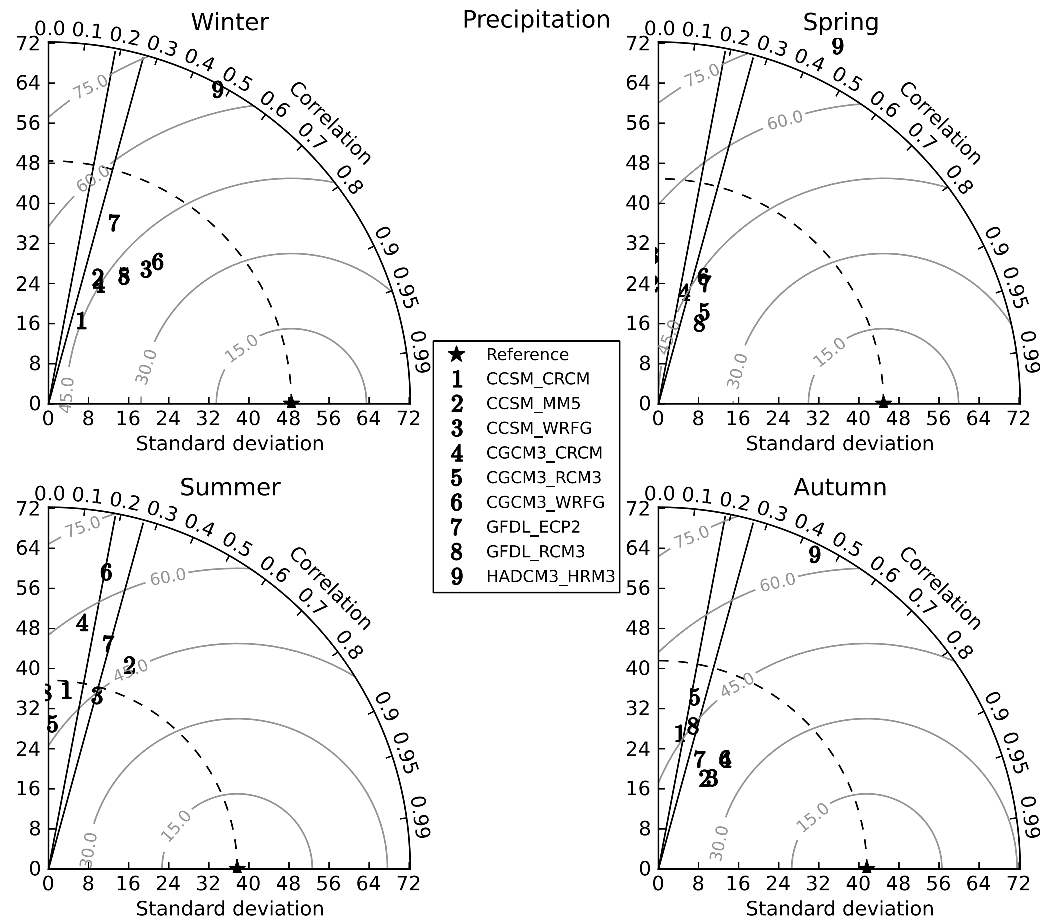

Figure 7. Taylor diagrams showing standard deviation (mm season^{-1}), RMSE (mm season^{-1}), and correlation for the observed and for each model seasonal (1971–2000) precipitation field across the northeast United States. GCMs and RCMs are listed in Tables 1 and 2. Statistics are calculated over all 211 grid cells of the observed field and from the nine GCM-RCM fields. The star indicates the statistics for the observed field. Contour of the reference standard deviation (from observed field) is show by the dashed line. RMSE contours are in gray. Correlation rays are the (left) 95th and (right) 99th significance levels.

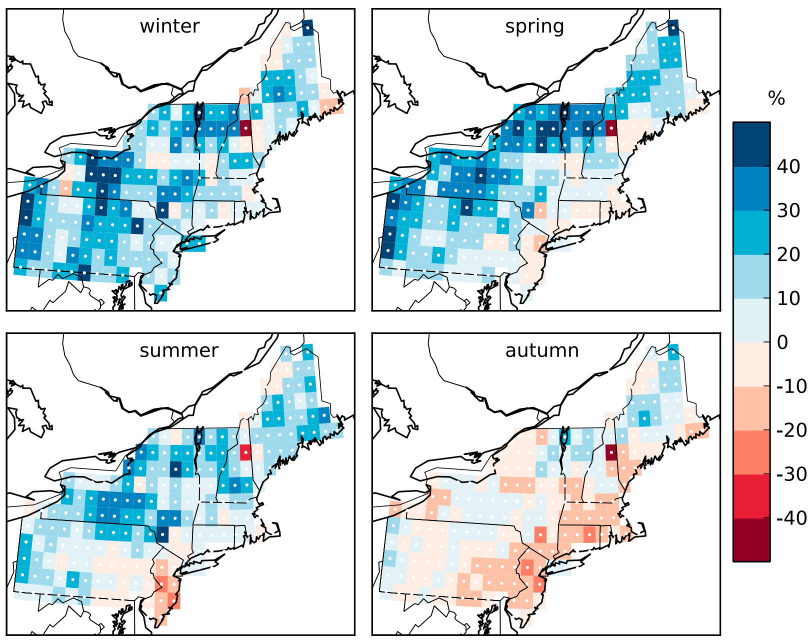

Figure 8. Bias (mm season^{-1}) in multimodel mean estimate of seasonal precipitation (1971–2000) at northeast United States grid cells. Bias is defined (P_{models} - P_{UDel}) / P_{UDel} * 100%, where P_{models} is the multimodel mean precipitation for the grid. Stipling indicates a significant bias at 95% confidence level based on student t test. Data for observations are taken from the UDel 0.5 degree data set.

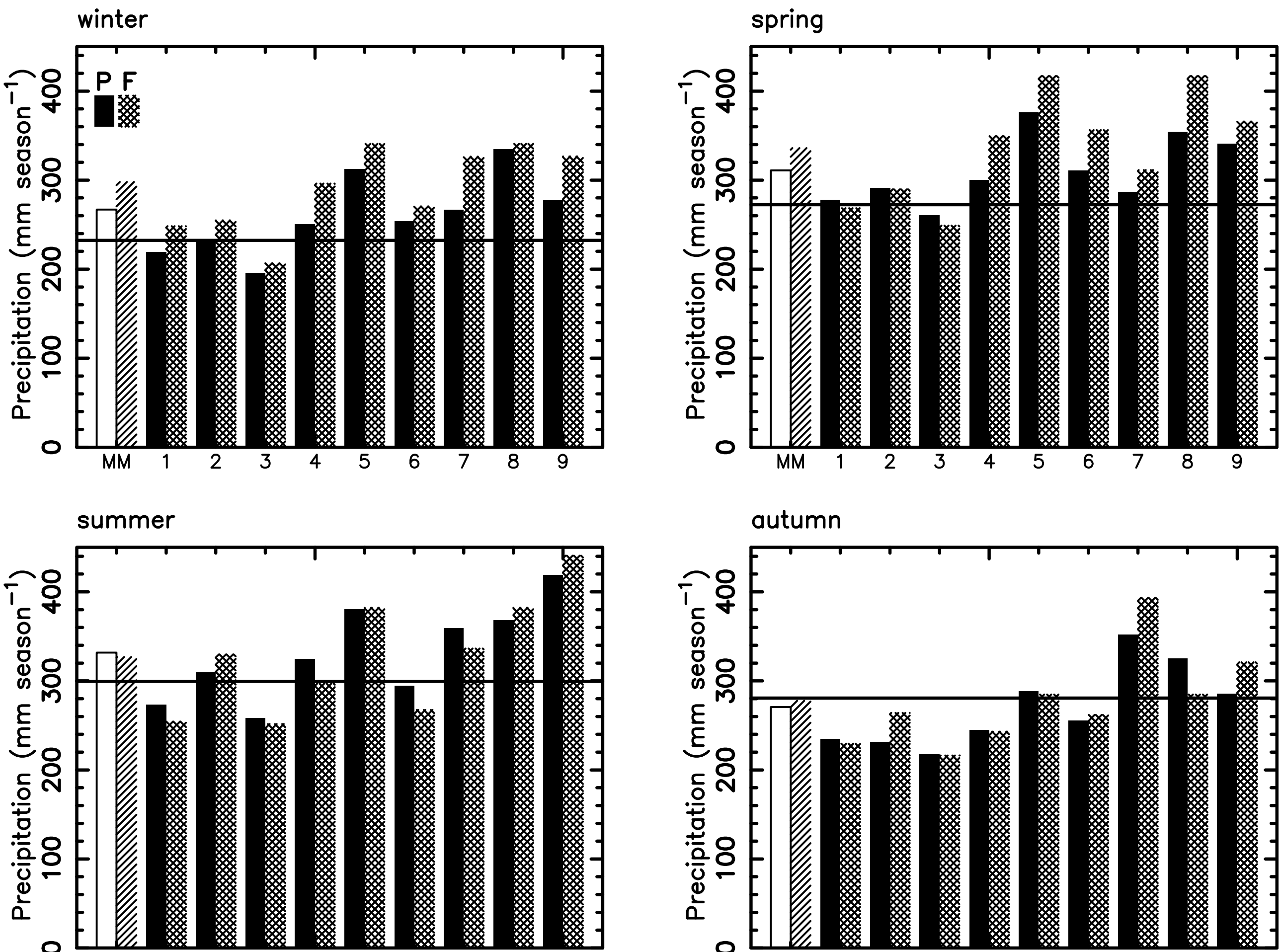

Figure 9. Seasonal mean total precipitation (mm season^{-1}) averaged across the northeast United States for each of the nine GCM-RCMs pairs and the multimodel mean over the present period 1971–2000 (solid rectangles) and future period 2041–2070 (hatched). The multimodel means are simple averages of the nine model estimates (no weighting). The value for seasonal precipitation from the observed field is indicated by the horizontal line. Models 1–9 are listed in Table 2.

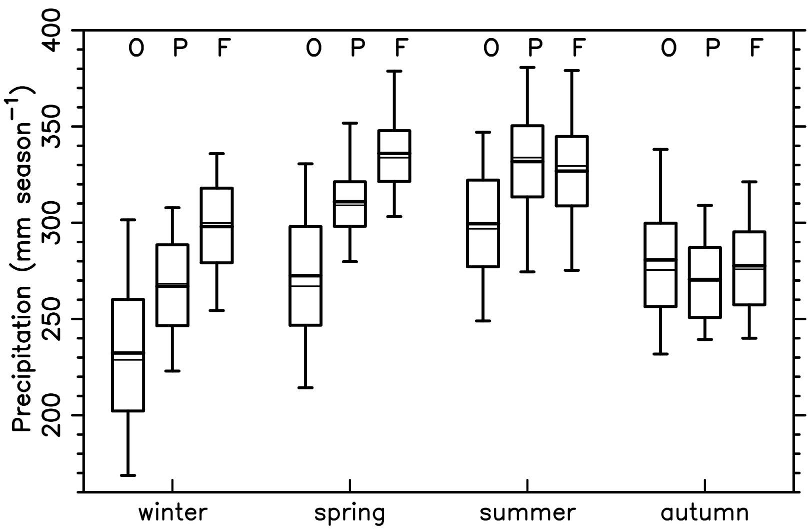

Figure 10. Distributions of precipitation (mm season^{-1}) for the observed (O) and RCM present period (P) fields for period 1971–2000, and for the future period (F). Each distribution consists of 211 0.5 degree grid cells spanning the northeast United States. Heavy line in each box is the distribution mean. Thin line (nearly identical to mean in most cases) is the distribution median. Boxes bracket the 25th and 75th percentiles. Whiskers show the 5th and 95th percentiles.

agreement mapping method [*Tebaldi et al.*, 2011] winter precipitation changes are significant and highest across interior areas. For spring and autumn models agree on small positive changes, which are significant over much of the region in spring and within the level of natural variability in autumn. The spatial change pattern in summer, with moderate precipitation declines across Pennsylvania and lower changes, relative to internal variability, to the north, presents an interesting case for attribution study.

[28] The magnitude and sign of seasonal precipitation changes are broadly consistent with a recent study using nine IPCC AOGCMs which also showed projected increases in winter precipitation and no change to a decrease in summer rainfall [*Hayhoe et al.*, 2007]. Future increases in winter temperature and precipitation would extend recent observed decreases in the snowfall-to-precipitation ratio [*Huntington et al.*, 2004]. Our results raise confidence in expectations for mid-century winter precipitation increases across the northeast. Given the robust projections of winter air temperature increases, a continuation of the recent trend toward wetter and warmer winters [*Hayhoe et al.*, 2007; *Keim et al.*, 2005] appears likely. For summer the combination of 2–3°C warming and precipitation decreases approaching 10% across much of Pennsylvania would likely create severe

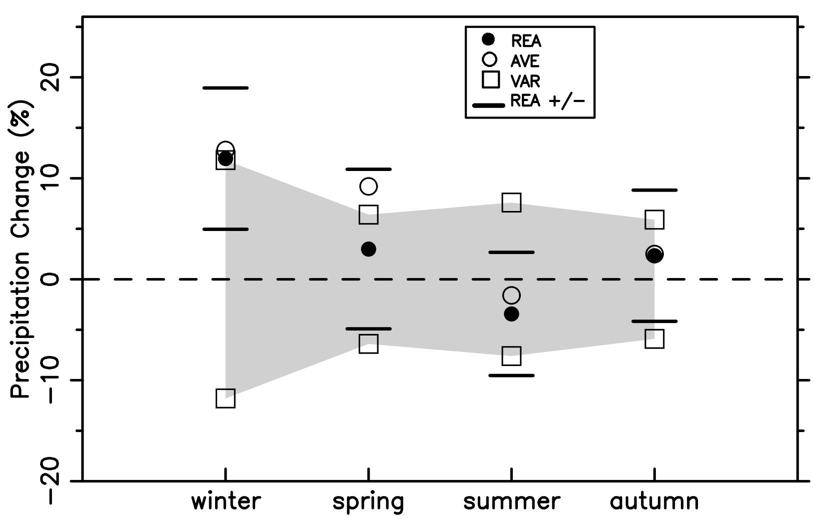

Figure 11. REA change (%, $\Delta = P_{2041–2070} - P_{1971–2000}/P_{1971–2000} * 100\%$) in seasonal precipitation across the northeast United States (solid circles); corresponding upper and lower REA uncertainty limits (horizontal lines); ensemble average changes (open circles); and estimated natural variability values (square). Units are percent of observed precipitation.

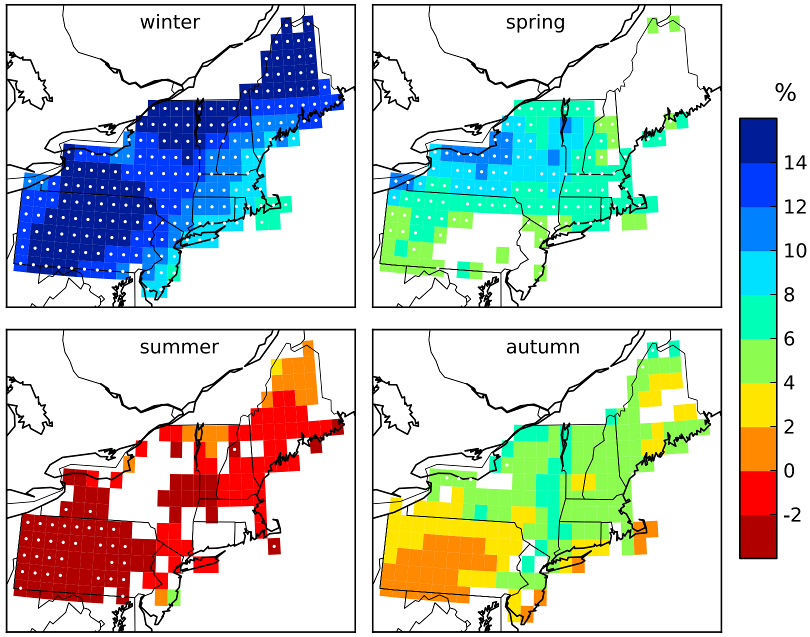

Figure 12. Relative percent change (%, $\Delta = P_{2041\text{–}2070} - P_{1971\text{–}2000}/P_{1971\text{–}2000} * 100\%$) in seasonal precipitation from the ensemble mean of the nine model pairs. Units are percent of present-day precipitation. Significance determined following criteria described by *Tebaldi et al.* [2011]. See text for details on meaning behind the uncertainty and significance logic.

water stress to ecosystems. This study illustrates the importance of using multiple model simulation estimates to gain understanding of likely changes in climate at decadal timescales and smaller spatial scales.

[29] **Acknowledgments.** This research was supported by the Consortium for Climate Risk in the Urban Northeast (award NA10OAR4310212 from the NOAA RISA program). The authors 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 (NSF), the U.S. Department of Energy (DoE), the National Oceanic and Atmospheric Administration (NOAA), and the U.S. Environmental Protection Agency Office of Research and Development (EPA). They also thank Fangxing Fan for comments on an earlier version of the manuscript.

References

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

Caya, D., and S. Biner (2004), Internal variability of RCM simulations over an annual cycle, *Clim. Dyn.*, *22*, 33–46, doi:10.1007/s00382-003-0360-2.

Christensen, J. H., E. Kjellström, F. Giorgi, G. Lenderink, and M. Rummukainen (2010), Weight assignment in regional climate models, *Clim. Res.*, *44*, 179–194.

Christensen, O. B., J. H. Christensen, B. Machenhauer, and M. Botzet (1998), Very high-resolution regional climate simulations over Scandinavia—Present climate, *J. Clim.*, *11*(12), 3204–3229, doi:http://dx.doi.org/10.1175/1520-0442(1998)011<3204:VHRRCS>2.0.CO;2.

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.

Dickinson, R. E., G. A. Meehl, and W. M. Washington (1987), Ice-albedo feedback in a CO_2 doubling simulation, *Clim. Change*, *10*, 241–248, doi:10.1007/BF00143904.

Diffenbaugh, N. S., J. S. Pal, R. J. Trapp, and F. Giorgi (2005), Fine-scale processes regulate the response of extreme events to global climate change, *Proc. Natl. Acad. Sci. U. S. A.*, *102*(44), 15,774–15,778, doi:10.1073/pnas.0506042102.

Dominguez, F., E. Rivera, D. P. Lettenmaier, and C. L. Castro (2012), Changes in winter precipitation extremes for the western United States under a warmer climate as simulated by regional climate models, *Geophys. Res. Lett.*, *39*, L05803, doi:10.1029/2011GL050762.

Frei, C., J. Hesselbjerg Christensen, M. Déqué, D. Jacob, R. G. Jones, and P. L. Vidale (2003), Daily precipitation statistics in regional climate models: Evaluation and intercomparison for the European Alps, *J. Geophys. Res.*, *108*(D3), 4124, doi:10.1029/2002JD002287.

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*(10), 1141–1158.

Giorgi, F., and L. O. Mearns (2003), Probability of regional climate change based on the reliability ensemble averaging (REA) method, *Geophys. Res. Lett.*, *30*(12), 1629, doi:10.1029/2003GL017130.

Giorgi, F., L. O. Mearns, C. Shields, and L. McDaniel (1998), Regional nested model simulations of present day and 2 x CO_2 climate over the Central Plains of the US, *Clim. Change*, *40*, 457–493.

Giorgi, F., et al. (2001), *Regional Climate Information—Evaluation and Projections*, edited by J. T. Houghton et al., 881 pp., Cambridge Univ. Press, Cambridge, U. K.

Giorgi, F., J. Coln, and A. Ghassem (2009), Addressing climate information needs at the regional level. the cordex framework, *WMO Bull.*, *58*(3), 175–183.

Gutowski, W. J., et al. (2010), Regional extreme monthly precipitation simulated by NARCCAP RCMs, *J. Hydrometeorol.*, *11*, 1373–1379, doi:10.1175/2010JHM1297.1.

Hall, A. (2004), The role of surface albedo feedback in climate, *J. Clim.*, *17*, 1550–1568, doi:http://dx.doi.org/10.1175/1520-0442(2004)017<1550:TROSAF>2.0.CO;2.

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

Hawkins, E., and R. Sutton (2011), The potential to narrow uncertainty in projections of regional precipitation change, *Clim. Dyn.*, *37*, 407–418, doi:10.1007/s00382-010-0810-6.

Hayhoe, K., et al. (2007), Past and future changes in climate and hydrological indicators in the US Northeast, *Clim. Dyn.*, *28*, 381–407, doi:10.1007/s00382-006-0187-8.

Huntington, T. G., G. A. Hodgkins, B. D. Keim, and R. W. Dudley (2004), Changes in the proportion of precipitation occurring as snow in New England (1949 to 2000), *J. Clim.*, *17*, 2626–2636.

Keim, B. D., M. R. Fischer, and A. M. Wilson (2005), Are there spurious precipitation trends in the United States Climate Division database?, *Geophys. Res. Lett.*, *32*, L04702, doi:10.1029/2004GL021985.

Knutti, R., R. Furrer, C. Tebaldi, J. Cermak, and G. A. Meehl (2010), Challenges in combining projections from multiple climate models, *J. Clim.*, *23*(18), 2739–2758, doi:10.1175/2009JCLI3361.1.

Lenderink, G. (2012), Climate change: Tropical extremes, *Nat. Geosci.*, *5*, 689–690, doi:10.1038/ngeo1587.

Li, G., X. Zhang, F. Zwiers, and Q. H. Wen (2012), Quantification of uncertainty in high-resolution temperature scenarios for North America, *J. Clim.*, *25*(9), 3373–3389, doi:10.1175/JCLI-D-11-00217.1.

Liepert, B. G., and M. Previdi (2012), Inter-model variability and biases of the global water cycle in CMIP3 coupled climate models, *Environ. Res. Lett.*, *7*(1), 014006, doi:10.1088/1748-9326/7/1/014006.

Mearns, L. O. (2003), Issues in the impacts of climate variability and change on agriculture, *Clim. Change*, *60*, 1–7.

Mearns, L. O., F. Giorgi, L. McDaniel, and C. Shields (1995), Results from the model evaluation consortium for climate assessment, *Global Planet. Change*, *10*(1–4), 55–78, doi:10.1016/0921-8181(94)00020-E.

Mearns, L. O., et al. (2007), The North American Regional Climate Change Assessment Program dataset, accessed 2 December 2011, http://www.earthsystemgrid.org/project/NARCCAP.html, NCAR, Boulder, Colo.

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.

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

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

Plummer, D. A., D. Caya, A. Frigon, H. Côté, M. Giguère, D. Paquin, S. Biner, R. Harvey, and R. de Elia (2006), Climate and climate change over North America as simulated by the Canadian RCM, *J. Clim.*, *19*, 3112–3132, doi:10.1175/JCLI3769.1.

Rowlands, D. J., et al. (2012), Broad range of 2050 warming from an observationally constrained large climate model ensemble, *Nat. Geosci.*, *5*, 256–260, doi:10.1038/ngeo1430.

Roy, P., P. Gachon, and R. Laprise (2011), Assessment of summer extremes and climate variability over the north-east of North America as simulated by the Canadian Regional Climate Model, *Int. J. Climatol.*, *32*(11), 1615–1627, doi:10.1002/joc.2382.

Sobolowski, S., and T. M. Pavelsky (2012), Evaluation of present and future North American Regional Climate Change Assessment Program (NARCCAP) regional climate simulations over the southeast United States, *J. Geophys. Res.*, *117*, D01101, doi:10.1029/2011JD016430.

Tebaldi, C., and R. Knutti (2007), The use of the multi-model ensemble in probabilistic climate projections, *Philos. Trans. R. Soc. A*, *365*, 2053–2075, doi:10.1098/rsta.2007.2076.

Tebaldi, C., J. M. Arblaster, and R. Knutti (2011), Mapping model agreement on future climate projections, *Geophys. Res. Lett.*, *38*, L23701, doi:10.1029/2011GL049863.

Willmott, C. J., and K. Matsuura (1995), Smart interpolation of annually averaged air temperature in the United States, *J. Appl. Meterol.*, *34*, 811–816.

Willmott, C. J., and S. M. Robeson (1995), Climatologically aided interpolation (CAI) of terrestrial air temperature, *Int. J. Climatol.*, *15*, 221–229.

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