Does nudging squelch the extremes in regional climate modeling?

Tanya L. Otte[1], Christopher G. Nolte, Martin J. Otte, and Jared H. Bowden[2]

U.S. Environmental Protection Agency
National Exposure Research Laboratory
Research Triangle Park, North Carolina

Revision 1

Submitted to
Journal of Climate

25 April 2012

[1] *Corresponding author address*: Tanya L. Otte, U.S. EPA/ORD/NERL/AMAD, 109 T. W. Alexander Dr., MD-E243-01, Research Triangle Park, NC 27711.
E-mail: otte.tanya@epa.gov

[2] Current affiliation: Institute for the Environment, University of North Carolina at Chapel Hill

Abstract

An important question in regional climate downscaling is whether to constrain (nudge) the interior of the limited-area domain toward the larger-scale driving fields. Prior research has demonstrated that interior nudging can increase the skill of regional climate predictions originating from historical data. However, there is concern that nudging may also inhibit the regional model's ability to properly develop and simulate mesoscale features, which may reduce the value added from downscaling by altering the representation of local climate extremes. Extreme climate events can result in large economic losses and human casualties, and regional climate downscaling is one method for projecting how climate change scenarios will affect extreme events locally. In this study, the effects of interior nudging are explored on the downscaled simulation of temperature and precipitation extremes. Multi-decadal, continuous Weather Research and Forecasting model simulations of the contiguous United States are performed using coarse reanalysis fields as proxies for global climate model fields. The results demonstrate that applying interior nudging improves the accuracy of simulated monthly means, variability, and extremes over the multi-decadal period. The results in this case indicate that interior nudging does not inappropriately squelch the prediction of temperature and precipitation extremes and is essential for simulating extreme events that are faithful in space and time to the driving large-scale fields.

1. Introduction

Projecting climate change to local scales is important for understanding and mitigating the effects of climate change on society and the environment. Many of the current general circulation models (GCMs) for simulating climate are run with a horizontal resolution of about 1° × 1°. Although at this resolution the large-scale atmospheric features that drive weather and climate are well represented, mesoscale features and local topography are not resolved, and consequently the GCM may not accurately represent local changes in temperature and precipitation extremes (Dulière et al. 2011; Werth and Garrett 2011). To predict the local effects of climate change, the GCM's fields can be projected to local scales using a regional climate model (RCM) by applying dynamical downscaling techniques (e.g., Giorgi 1990). The RCM may then be used to inform problem-focused climate assessments that address community goals and values (Tryhorn and DeGaetano 2011).

To interpret climate change at the local level, there is great interest in characterizing changes in "extreme events". Extreme events are rare but important meteorological phenomena such as droughts, floods, extreme heat and cold, and strong wind events that are statistically associated with the tails of a probability distribution (e.g., Meehl et al. 2000b; Garrett and Müller 2008). Extreme weather events have significant societal impacts such as large economic costs and human casualties (e.g., Meehl et al. 2000a). Indices of climate extremes often involve tracking the exceedances of a critical threshold value (e.g., Karl et al. 1999), and may consider the frequency, duration, and areal extent of the exceedance. Changes in the duration and/or intensity of extreme events will impact air quality, water quantity and quality, agriculture (growing season, types of crops, water availability), energy demands and sources, urban infrastructure and building codes, and the overall economy. Because of the spatial heterogeneity in extreme precipitation and temperature events (e.g., Trenberth et al. 2007), RCMs that are used for projecting future changes in frequency and intensity of extreme events must reflect the state-of-the-science.

When using RCMs to downscale GCM fields, interior nudging may reduce errors in RCM predictions compared with applying a constraint only at the lateral boundaries (Miguez-Macho et al. 2004; Castro et al. 2005; Lo et al. 2008; Alexandru et al. 2009; Bowden et al. 2012a). Feser et al. (2011) indicate that constraint toward the atmospheric large scales (i.e., via nudging) when downscaling often increases mesoscale variability and “adds value” to the global climate model forecasts. The balance in the constraint toward the GCM fields against the RCM’s freedom to develop mesoscale features is difficult to determine objectively and has not yet been achieved (e.g., Kanamaru and Kanamitsu 2007; Alexandru et al. 2009; Bowden et al. 2012a). Arritt and Rummukainen (2011) juxtapose that nudging too weakly allows the RCM to diverge from the GCM fields, while nudging too strongly can suppress the development of the finer-scale processes that are sought with the RCM. Christensen et al. (2007) also caution that while nudging minimizes large-scale error in the RCM, it can also mask model biases. Pielke et al. (2012) argue that nudging can force the RCM to retain and potentially exacerbate errors that exist in the GCM. Although nudging is becoming increasingly common for regional climate modeling, using interior nudging techniques is not universally accepted as a standard practice for dynamical downscaling.

Despite improving the means and retaining large-scale consistency with the driving model, there is some concern that using interior nudging techniques may dampen the extremes and variability. Using the Canadian RCM (CRCM), Alexandru et al. (2009) found that increasing the strength of spectral nudging by initiating spectral nudging closer to the surface decreased the intensity of precipitation during a summer period. Cha et al. (2011), using the Weather Research and Forecasting (WRF) model with a version of spectral nudging that follows von Storch et al. (2000) and is similar to the CRCM, found that while spectral nudging reduced errors in the tracks of tropical cyclones, it artificially weakened tropical cyclone intensities. Bowden et al. (2012a) showed that spatial variability with analysis nudging in WRF is decreased as the nudging timescale is decreased.

There are few studies that investigate the ability of RCMs to simulate extreme events, particularly over North America, and none of the following explicitly mention using nudging. Using the Pennsylvania State University–National Center for Atmospheric Research mesoscale model (MM5), Lynn et al. (2007) showed that correctly predicting the surface energy balance was essential for predicting extreme summer temperatures over the eastern U.S. Dulière et al. (2011) showed that both WRF and the Hadley Centre Regional Model (HadRM) adequately represented local extremes of temperature and precipitation in the Northwest U.S. over a recent 5-yr period. Caldwell et al. (2009) found that WRF driven by 40-yr climate simulations overpredicted precipitation extremes over California and underpredicted the frequency of precipitation events. By contrast, Mladjic et al. (2011) found that the CRCM underpredicted precipitation extremes across Canada for an historical 30-yr period.

This study addresses two relevant questions for dynamical downscaling for the contiguous United States (CONUS): how well can a RCM simulate temperature and precipitation means and extremes for a multi-decadal period, and how does nudging affect the frequencies and intensities of those extreme events? Colin et al. (2010) created a 23-yr simulation with ALADIN-Climate and found that spectral nudging did not adversely affect the prediction of extreme precipitation events over Europe. This study investigates the effects of nudging techniques on predictions of extreme temperatures and precipitation with the WRF model as a RCM to simulate an historical 20-yr period. We evaluate the results against high-resolution analyses and examine the impacts of nudging on simulated extremes across the CONUS to determine whether interior nudging in WRF inappropriately squelches the extremes.

2. Model description

The WRF model version 3.2.1 (WRFv3.2.1) was initialized at 0000 UTC 2 December 1987 and run for a 1-month spin-up, then run continuously for 20 years through 0000 UTC 1 January 2008. The two-way-nested modeling domains (108- and 36-km horizontal grid spacing; see Fig. 1) covered North America and the CONUS, respectively. WRF was run with a 34-layer configuration that extended to a model top at 50 hPa. The physics options included the Rapid Radiative Transfer Model for global climate models (RRTMG; Iacono et al. 2008) for longwave and shortwave radiation, the WRF single-moment 6-class microphysics scheme (Hong and Lim 2006), the Grell ensemble convective parameterization scheme (Grell and Dévényi 2002), the Yonsei University planetary boundary layer (PBL) scheme (Hong et al. 2006), and the Noah land-surface model (Chen and Dudhia 2001). The input data are 2.5° × 2.5° analyses from the NCEP-Department of Energy Atmospheric Model Intercomparison Project (AMIP-II) Reanalysis data (Kanamitsu et al. 2002) (hereafter, R-2), which are at comparable spatial and temporal resolutions as GCM fields. Since the data are from an historical period, the downscaled runs can be evaluated against higher-resolution reanalysis products. The R-2 fields provide initial, lateral, and surface boundary conditions, and they serve as the constraints when interior nudging is used. No further observational data are assimilated into the WRF simulation.

Three 20-yr runs are performed with WRF. One simulation includes nudging only through the lateral boundaries (Davies and Turner 1977) using a 5-point sponge zone (i.e., no nudging, NN). The other simulations additionally use one of the two forms of grid-based nudging that are available in public versions of WRF: analysis nudging (AN) and spectral nudging (SN). Both forms of interior nudging can reduce errors in the means in regional climate modeling with WRF (e.g., Lo et al. 2008; Bowden et al. 2012a).

The analysis nudging technique in WRF (Stauffer and Seaman 1990; Deng et al. 2007) is theorized to be most useful when the input data fields are not significantly coarser than the model resolution. In WRF, analysis nudging adds a non-physical term to the prognostic equations that is proportional to the difference between the model state and a value that is interpolated in time and space from the reference analysis. Analysis nudging is applied toward horizontal wind components, potential temperature, and water vapor mixing ratio. The analysis nudging coefficients (Table 1) are set to the default values in WRF for wind and temperature for the 108-km domain, but reduced for moisture (e.g., Otte 2008) and reduced for all coefficients for the 36-km domain (e.g., Stauffer and Seaman 1994). The analysis nudging is only applied above the PBL to maximize WRF's freedom to develop mesoscale circulation in the PBL.

Spectral nudging is attractive as a scale-selective interior constraint for regional climate downscaling because it can restrict nudging toward the longer wavelengths. Similar to analysis nudging, spectral nudging affects the model solution through a non-physical term in the prognostic equations, but instead the term is based on the difference between the spectral decompositions of the model solution and the reference analysis. The spectral nudging in WRFv3.2.1 follows Miguez-Macho et al. (2004) and can be applied toward horizontal wind components, potential temperature, and geopotential. As in the analysis nudging simulation, spectral nudging is only applied above the PBL. Spectral nudging is used to constrain WRF toward synoptic-scale wavelengths and is applied in WRF to wavelengths longer than a threshold that is a function of domain size and a specified cutoff wavenumber. The threshold wavelength for spectral nudging should not be less than the shortest wavelength resolved by the input fields, which is at least $4\Delta x$ (Pielke 1984) of the R-2 analyses, or ~1100 km in midlatitudes. Nudging coefficients, threshold wavenumbers used for spectral nudging, and their corresponding wavelengths are given in Table 1.

3. Analysis

The three WRF simulations on the 36-km domain are analyzed for the historical period 1988–2007. We seek to determine how nudging affects the representation of 2-m temperature and precipitation extremes over the 20-yr period. Since no interior nudging occurs within the PBL, neither 2-m temperature nor precipitation is directly assimilated.

For a variable with a given statistical distribution, the frequency of extreme events (as measured by threshold exceedances) changes if the mean of the distribution shifts and/or if the variance (width) of the distribution changes (Meehl et al. 2000a). A change in the mean will cause an increase in threshold exceedances on one end (e.g., the number of hot days) and a decrease on the other side of the distribution (e.g., the number of cold days). A change in the variance will affect the frequency and magnitude of extremes on both sides of the distribution, and according to Katz and Brown (1992) it may be more important for changes in extreme outliers (i.e., events more than one standard deviation from the mean). Since the representation of the mean and variance is important for the frequency and severity of extreme events, we first examine how the three downscaling strategies influence the mean 2-m temperature and precipitation from the RCM. Then, to investigate the effects of nudging on the variability in the RCM we compare spatial spectra from the RCM fields with those from the reanalysis fields. Finally, we examine the extremes of 2-m temperature and precipitation in the downscaled runs.

The WRF simulations are compared to the R-2 fields to determine the extent to which the large-scale variability is preserved in the WRF simulation. For near-surface fields, where mesoscale detail is expected to be gained by using a RCM, the WRF simulations are compared to high-resolution reanalyses from the North American Regional Reanalysis (NARR; Mesinger et al. 2006) and the Climate Forecast System Reanalysis (CFSR; Saha et al. 2010). Both the NARR and the CFSR should include mesoscale detail that is comparable to what could be produced in the 36-km WRF simulations. The NARR is a 32-km limited-area reanalysis that has 3-h fields and is often used for understanding regional climate and for validation of regional climate modeling studies over North America (e.g., Ruiz-Barradas and Nigam 2006; Bukovsky and Karoly 2007; Lo et al. 2008; Becker et al. 2009; Bowden et al. 2012a). The CFSR is a 0.31° (~35-38-km at midlatitudes) global reanalysis that consists of 6-h analyses supplemented with hourly forecasts. Here, CFSR is used for comparisons of 2-m temperature, and NARR is used for precipitation, as explained below.

Several of the extremes that are examined in this paper are comparisons of 2-m temperature against threshold values. With 3-h temporal sampling, the NARR is inadequate for counting temperature exceedances. We instead use the hourly gridded fields from the CFSR. Saha et al. (2010) show that the multi-year mean and trend of 2-m temperature from CFSR match well with comparable fields used in the climate change community to estimate global warming trends. Wang et al. (2011) show that 2-m temperature from CFSR is more highly correlated with observations than either R-2 or its predecessor R-1 is.

To ensure that the fields from CFSR are qualitatively and quantitatively consistent with a validated source, the mean 2-m temperature for 1988–2007 (i.e., the 20-yr period of the WRF simulations) is computed for both NARR and CFSR interpolated to the 36-km WRF domain at their highest temporal resolutions (i.e., 3-h and 1-h fields for the NARR and CFSR, respectively) using WRF preprocessing software. Outside of regions with complex terrain, the 20-yr mean 2-m temperature is consistent between NARR and CFSR (Fig. 2). East of the Rocky Mountains (excluding the southern Appalachian Mountains), the differences in the 20-yr mean 2-m temperature between NARR and

CFSR are typically within ±1.5 K. Differences between NARR and CFSR in the 20-yr mean 2-m temperature typically exceed ±2.5 K in areas of complex terrain in the CONUS. Although both NARR and CFSR are reanalysis products that are strongly influenced by observations, neither model assimilates 2-m temperature directly.

Precipitation comparisons are made against NARR fields that have been interpolated to the 36-km WRF domain. Over the CONUS, precipitation fields from the NARR are influenced by assimilating hourly precipitation derived from 1/8° daily analyses of rain gauge data, which are then converted to latent heat to constrain the NARR precipitation (Mesinger et al. 2006). The amplitude of the annual cycle of precipitation is well-depicted by NARR (Ruiz-Barradas and Nigam 2006) and, overall, NARR precipitation is “virtually indistinguishable” from observations (Nigam and Ruiz-Barradas 2006). Bukovsky and Karoly (2007) conclude that although NARR is imperfect, it is superior to other reanalysis products for precipitation and it adequately captures extreme events, even over the topography of the western U.S. Becker et al. (2009), however, note that NARR has a systematic bias toward more frequent, lighter precipitation and extremes are underestimated in the eastern United States. In accordance with Mesinger et al. (2006), our precipitation comparisons are restricted to land and over the CONUS because NARR is less reliable where limited and coarser-scale data were assimilated. Since the NARR precipitation fields represent the CONUS well, we use NARR instead of CFSR precipitation fields which have not been adjusted by observational assimilation. Our analysis indicates that CFSR is much wetter than NARR (not shown), which is corroborated by Higgins et al. (2010) and Mo et al. (2011) who showed systematic overprediction of precipitation by CFSR throughout the CONUS.

a. Mean 2-m temperature and precipitation

Although the focus of this work is on simulating extreme events, we first evaluate the mean values of 2-m temperature and total precipitation in the WRF simulations over different temporal scales because changes in the means will affect the extreme values. The 20-yr mean 2-m temperature is computed for each of the three WRF simulations and compared against CFSR (Fig. 2). All three WRF simulations show a slight warm bias (>0.5 K) in the Plains (see Fig. 1) and along the southeastern Atlantic coast compared to CFSR. The differences from CFSR are more pronounced in NN, where the warm bias exceeds 1.5 K in the southern Plains and a large area of cool bias of more than 0.5 K extends throughout southeastern Canada. As in the comparison of NARR with CFSR, all three WRF simulations have large differences from CFSR in complex terrain, and the patterns, signs, and magnitudes of the differences in complex terrain are consistent when compared to the difference between NARR and CFSR (Fig. 2). Differences between NARR and the WRF simulations are not as pronounced as in the comparisons with CFSR especially in complex terrain (not shown), which suggests that the NARR topography may be more consistent with WRF than the topography used in the global CFSR.

The precipitation predicted by WRF is too high compared to NARR throughout much of the domain (Fig. 3). Average annual precipitation in WRF is particularly exaggerated in complex terrain and east of the Rocky Mountains. Although the average annual precipitation in WRF is too high regardless of whether nudging is used, the WRF simulations all correctly predict that the highest precipitation amounts occur along the northwestern coast and in the eastern United States.

The evaluation of the extremes in this paper focuses on the Midwest region (Fig. 1), which has only gradual changes in topography; the other regions are presented in less detail to permit a broader analysis. The differences between NARR and CFSR in 20-yr mean 2-m temperature are typically within ±0.5 K throughout the Midwest (Fig. 2), and those differences are overall the smallest of the regions in Fig. 1. In the Midwest, NN has little bias compared to CFSR (Fig. 2), except for a slight cool bias between -1.5 K and -0.5 K around the northern, eastern, and southern peripheries of that region. AN has a slight warm bias (0.5–1.5 K) in the Midwest, and SN is the least biased compared to CFSR for the 20-yr mean 2-m temperature (Fig. 2).

Figure 4 shows a time series of the monthly area-average 2-m temperature difference from CFSR for the Midwest region for each of the three WRF simulations. Although the 20-yr mean 2-m temperature from NN compares well to CFSR and arguably may be as good as or better than AN and SN, examining only the mean 2-m temperature over the 20-yr period can be misleading (cf. Fig. 2 and Fig. 4). The monthly area-average 2-m temperature over the 20-yr period shows deviations greater than 4 K in NN (Fig. 4). These month-to-month differences in NN indicate the RCM's inability to correctly simulate weather conditions that are consistent with the large-scale driving fields and show that the modest mean annual bias (Fig. 2) results from averaging large monthly biases that have the opposite sign (Fig. 4). Both AN and SN reduce the monthly deviations from CFSR to less than ±2 K (Fig. 4). Each year, the most pronounced monthly cold bias in the Midwest in NN is typically in July or August (Fig. 4), and that cold bias is mitigated by both forms of nudging, slightly more strongly by AN than SN. AN is slightly warmer than SN for most months throughout the 20-yr period, which is consistent with the relative comparisons of AN and SN to CFSR (Fig. 2). AN and SN improve the average monthly predictions of 2-m temperature throughout the domain compared to NN (Fig. 5). In NN, there is a pronounced cold bias (approaching 3 K) in the eastern U.S. in the summer, which is mitigated by either form of nudging.

The three WRF simulations generally overpredict precipitation by 10–50 mm per month compared to NARR (Figs. 4 and 6), which is consistent with the overpredictions in Fig. 3. The largest monthly differences in the Midwest (Fig. 4) are typically in NN, and the differences are progressively reduced in SN and AN. Some months in the 20-yr period also have noticeable underpredictions of area-average precipitation of more than 25 mm, particularly in NN. In addition, the phase of the errors in NN is often not aligned with the errors in AN and SN, which suggests that the individual weather events in NN may be misrepresented. Such large differences in area-average precipitation in NN over a one-month period (both overprediction and underprediction) indicate the RCM's inability to accurately characterize prolonged periods of heavy rain and dry spells that could contribute to flooding and drought, and the resulting errors in the surface heat fluxes would affect the ability of the RCM to predict extreme temperatures (e.g., Lynn et al. 2007). Overall, using either form of interior nudging improves the regional prediction of monthly precipitation by WRF, and AN gives better predictions than SN for 5 of the 6 regions (Fig. 6).

b. Spectra of downscaled fields

Since variability can influence extreme events (Katz and Brown 1992; Meehl et al. 2000a), spectra are examined to determine the effects of nudging on variability at different spatial scales. Spectra represent the contribution of each wavenumber to the total variance and can indicate how well the large-scale fields from R-2 are captured and reproduced by WRF. In addition, comparing the WRF spectra to NARR shows if WRF is producing variability at the smaller scales where value should be added from the downscaling process.

One-dimensional spatial spectra are computed along rows of the 36-km domain (grid-relative west-east) for R-2, NARR, and the three WRF simulations. The spectra are computed every 6 h, and all data for each month are averaged over the 20-yr period. The data are detrended by fitting the fields along each model row to a quadratic least-squares regression, then using the regression to remove linear and parabolic trends. After subtracting the row mean, a Hamming window (Kaimal and Kristensen 1991) is used to taper the rows to force periodicity for the spectral computations. Following Kaimal and Kristensen, the final spectra are multiplied by 2.52 to compensate for the reduction of variance from the Hamming window.

In January, the variability in the long waves (longer than $4\Delta x$ for R-2) in 500-hPa temperature over the 20-yr period is consistent with R-2 in all three WRF simulations at 36-km (Fig. 7). WRF retains much of the large-scale variability from R-2 via the lateral boundaries during January when there is strong synoptic forcing, though there is a slight reduction in variability in NN at long wavelengths compared to the other spectral representations of January. In the mesoscale wavelengths (between $4\Delta x$ for R-2 and $4\Delta x$ for WRF), both NN and SN add variability at a magnitude that is consistent with NARR, while AN has reduced variability compared with NARR. Even by weakening the nudging on the 36-km domain compared to model defaults, the analysis nudging technique may be nudging too strongly toward the R-2 fields and, as a result, unrealistically suppressing variability in the wintertime 500-hPa temperature. Thus, the nudging coefficients used for AN should be further revised for regional climate simulations to achieve the optimal balance between mesoscale variability and fidelity to the driving fields. Approaching 4Δx in WRF, all three WRF runs have higher variability than NARR, suggesting the downscaled runs have too much variance at those scales.

In July, the long waves in 500-hPa temperature are consistent between R-2 and the nudged WRF simulations. However, there is much greater and unrealistic variability in NN (note the logarithmic ordinate axis in Fig. 7). This suggests that without interior nudging, weak synoptic forcing through the lateral boundaries allows WRF too much freedom to generate variability. Simply comparing the three WRF simulations could lead to the conclusion that using either interior nudging technique in WRF adversely impacts the variability in the multi-decadal regional climate prediction. However, the variability in NN is neither present in the large-scale driving fields (R-2), nor is it corroborated by the NARR. At the mesoscale wavelengths, AN has reduced the variance compared to SN and NN during July. SN is seemingly effective for producing large-scale variability that is consistent with NARR while also allowing the RCM to develop smaller-scale variability.

Examining 700-hPa water vapor mixing ratio for January and July (Fig. 8) suggests the large-scale moisture fields from R-2 are generally retained, but there is too much variability in all three WRF simulations regardless of whether interior nudging is used. The increased humidity variance in WRF is consistent with the overprediction of precipitation in all WRF simulations. Unlike for 500-hPa temperature (and momentum fields, not shown), the variance of 700-hPa water vapor mixing ratio with AN is not unrealistically suppressed. This suggests that analysis nudging may be adjusting the variance in the moisture fields toward the observed state, which is also consistent with the better predictions of precipitation by AN than SN (Fig. 4), or that the humidity is strongly controlled by fields in the PBL that are not nudged. Recall that the analysis nudging technique in WRF can adjust the water vapor mixing ratio field, while spectral nudging cannot.

To focus on the long waves where the RCM should be consistent with the large-scale driving fields, energy spectra are shown in Fig. 9 with a linear ordinate axis. At 250 hPa, the energy in the January meridional wind is reduced for all three WRF simulations compared to the representations in R-2 and NARR. NN has notably lower energy than both AN and SN, where energy in the long waves is increased to approach the reference fields. In July, the 250-hPa meridional wind spectra are qualitatively similar to January, but the magnitudes are smaller because the synoptic transport has a smaller meridional component in July in this domain. At 500 hPa for January, the distinctions between the WRF runs and the reference fields are small, although NN still has slightly lower energy compared to the other runs. However, at 500 hPa in July, NN has greater energy than the other WRF runs and the reference fields (consistent with Fig. 7). In addition, compared to July at 250 hPa, the 500-hPa spectral energy of the meridional wind has the opposite sign of the error, so the distribution of energy in NN in the column is in error, and interior nudging notably acts to mitigate that error under weak synoptic forcing. The analogous zonal wind spectra (not shown) are qualitatively similar to Fig. 9.

As shown in Figs. 7–9, a larger total variance in the RCM simulations is not an indication of added value. Comparing the total variance of RCM simulations only to each other is not enough to determine the best representation of regional climate. The added or reduced variance at the large scales in NN (Figs. 7–9) represents an undesired deviation from the driving fields, and those errors in variance at larger scales may cascade down and contaminate the smaller scales. The spectra suggest that using interior nudging (AN or SN) produces larger-scale features that are more consistent with the driving fields. The adverse impacts of AN at smaller scales may be mitigated by further decreasing the nudging strength (Bowden et al. 2012a).

c. Annual totals of daily exceedances of extreme thresholds

To evaluate extremes, we first examine exceedances of 2-m temperature and precipitation thresholds from the RCM compared to those computed from CFSR (temperature) and NARR (precipitation). For the RCM simulations and the high-resolution reanalyses, the number of days in each year that the threshold was exceeded at each grid cell was tallied. Those annual tallies for each threshold were then area-averaged within each region (see Fig. 1). The thresholds are based on the Annual Climatological Summary maintained by the NOAA National Climatic Data Center. The thresholds also align well with a subset of the 27 extreme indices suggested by the World Climate Research Programme Climate Variability and Predictability (CLIVAR) Expert Team on Climate Change Detection and Indices (e.g., Karl et al. 1999). Hot and cold thresholds for daily temperature and high daily precipitation thresholds are examined. The analysis for R-2 is not shown because the temperature data are too temporally coarse (6-h) to capture threshold values, and the precipitation estimates from R-2 are biased high (e.g., Guirguis and Avissar 2008; Wang et al. 2011).

Figure 10 shows the area-averaged number of days with 2-m temperature >90°F (32.2°C), or "summer days", based on hourly data. None of the RCM simulations predicts as many area-average exceedances of the 90°F threshold as the CFSR for the Midwest region in any of the 20 years simulated. Compared to CFSR, NN underestimates the annual number of summer days by as many as 40 days across the Midwest region. Both forms of interior nudging improve the simulation of summer days compared to NN, although AN and SN still typically underestimate the number of summer days by 10–20 days compared with CFSR. For the summer day threshold in the Midwest over this period, AN performs best. The underprediction of summer days in all WRF simulations (Fig. 10) is consistent with a persistent overprediction of precipitation in the region (Figs. 3 and 4), where the surface energy balance is likely tilted more toward latent heating because of the moist ground. In addition, the underprediction of temperatures at the "summer day" threshold is consistent with Fig. 4, which shows the largest underprediction of temperature typically occurs in July and is most pronounced in NN.

Figure 11 shows a comparison of the WRF simulations to CFSR over the Midwest region for three cold thresholds: number of days with temperature <32°F (0°C, frost days), number of days with maximum temperature <32°F (0°C, freeze days), and number of days with temperature <0°F (-17.8°C). For the first decade of the 20-yr simulation, all three WRF simulations tended to underpredict the number of frost days, but the number of area-average frost days for the Midwest was typically within five days of CFSR for all three WRF runs during the second decade. NN often had the largest differences from CFSR. Both AN and SN predicted similar numbers of frost days for most years and represented an improvement over NN throughout the 20-yr period.

For some years during the period, NN approximately predicted the area-average number of freeze days in the Midwest compared to CFSR (Fig. 11), but other years underpredicted the number of freeze days by more than 10. However, AN and SN consistently predicted the number of area-average annual freeze days within five days of CFSR. All three WRF simulations were consistent with CFSR in characterizing the number of very cold days (temperature <0°F) throughout the 20-yr period, though the most notable differences from CFSR occurred in NN.

Across all regions, the distributions of the 20-yr annual exceedances of the hot (90°F) and cold (32°F) thresholds are shown in Fig. 12. In NN, there is reduced interannual variability and too few exceedances of the hot threshold in the Midwest, Northeast, and Southeast, which is consistent with the strong summer cold biases shown in Fig. 5. In all of those regions, both AN and SN increase the interannual variability and the number of exceedances to be more consistent with CFSR. In the Northwest and Southwest, NN overpredicts the exceedances of the hot threshold, and this overprediction is mitigated with nudging. For the cold threshold, NN tends to artificially increase the interquartile range in the northern regions, where >100 cold days occur annually. For the nudged runs, the interquartile ranges are closer to CFSR than NN is in those regions. Nudging does not suppress the prediction of cold days relative to NN or to CFSR in most regions, although there is a slight reduction in the number of cold days predicted in the Plains in all WRF runs.

To understand the ability of WRF to simulate heavy precipitation events, comparisons are made to NARR estimates of numbers of days with precipitation exceeding thresholds of 0.5 in and 1.0 in (similar to CLIVAR indices of ≥10 mm and ≥20 mm). Figure 13 shows that for both precipitation thresholds, all three WRF simulations overpredict the annual area-average number of days that each threshold was surpassed in the Midwest compared to NARR. The overprediction of precipitation at the high thresholds by WRF occurs for each year of the 20-yr simulation period (Figs. 13 and 14), and it is consistent with the general overprediction of precipitation shown in Figs. 3, 4, and 6. In general, the overpredictions occur most frequently in NN, which suggests that without interior nudging, the configuration of WRF used here has a tendency to generate more heavy precipitation events than are observed. In general, NN predicts about ten more days ≥0.5 in and about five more days ≥1.0 in per year than were observed in the Midwest (using NARR as the benchmark). At the 0.5 in threshold, the SN simulation tends to overpredict the number of days as often as NN (Figs. 13 and 14). The precipitation event totals at both thresholds are best matched with NARR in AN in five of the six regions, possibly because AN is the only simulation that constrains moisture on the interior of the domain. Radu et al. (2008) showed that spectral nudging exaggerated the intensity of wintertime precipitation events unless a constraint toward specific humidity was introduced. Thus more heavy precipitation events are erroneously predicted without using interior nudging, and AN appropriately suppresses the number of events toward the observed state.

d. Monthly extremes and interannual variability

Here, extremes are assessed relative to the 20-yr climatology by examining monthly-averaged daily maximum and minimum 2-m temperature, monthly-averaged diurnal temperature range, and total monthly precipitation. As in the previous subsection, values are tabulated at each grid cell and aggregated to form an area average. For each of the 12 months, the means and standard deviations are computed relative to each model run's distribution to account for the bias in the RCM predictions (e.g., Figs. 2–4) and to track the annual cycle in the Midwest region. This subsection not only addresses extremes, but also the effects of nudging on the mean, variability, and timing of events in the RCM. To examine the effects of the variability on the extremes, two standard deviations from the mean ($\pm 2\sigma$) are considered outlier months. Assuming the data are normally distributed, approximately 1 in 22 values falls outside $\pm 2\sigma$, so those events occurring less than 5% of the time could be considered rare or extreme. Although this criterion is objective and practical, it is limited for precipitation which does not have a normal distribution, and its lower bound is 0.

Using the $\pm 2\sigma$ criterion, the CFSR identifies three exceptionally hot months and four exceptionally cold months in the Midwest region using the monthly area-averaged daily maximum 2-m temperature (Fig. 15). Four of those months (January 2006, September 1993, December 1989, and December 2000) were correctly characterized as exceptional in all three WRF runs, regardless of whether interior nudging was used. The exceptionally cold August 1992 was also identified as the coldest August in all three WRF runs, despite falling short of the -2σ criterion. (August 1992 is obscured for AN and SN in Fig. 15 because August 2004 has a similar value.) This shows that WRF can create credible predictions (e.g., from persistent and strong synoptic forcing through the lateral boundaries) and does not rely on nudging to compensate for shortcomings in physics. However, March 2000 and June 1988 were merely cast as unusually warm in NN, but correctly characterized as extreme by AN and SN. In fact, the summer of 1988 had the hottest June, July, and August of the 20-yr period, a prolonged period of drought in the Midwest. Without interior nudging, NN consistently underpredicted 2-m temperature during the summer months (consistent with Fig. 4), and did not identify 1988 as having a remarkably hot summer. In NN, July 1988 was 0.5 K cooler for the region than July 2006, its hottest July (a false alarm), which was only unusually warm (+1σ) in CFSR, AN, and SN. In addition, April 2006 was the hottest April of the 20-yr period in CFSR, AN, and SN, but without interior nudging, NN classified that month as near normal. Without interior nudging, WRF captured some of the extreme months during the 20-yr period, but had several misses and false alarms. Although imperfect, using interior nudging in WRF improves the representation of the extreme months, eliminates the misses and false alarms, and greatly improves the accuracy in characterizing the relative severity of the events.

As with daily maximum temperature, several months that had exceptionally hot or cold monthly area-averaged daily 2-m temperature minima (June 1992, August 1992, December 1989, December 2000) were correctly characterized in all three WRF runs, regardless of whether nudging was used (Fig. 16). However, without nudging, NN misclassified the severity of some months (October 1988 and 2007, which were the coldest and hottest Octobers at ±1σ rather than ±2σ, which suggests reduced interannual variability for October), missed extreme months altogether (June 2003, which was the 2nd

coldest in CFSR, AN, and SN, but average in NN), or simulated extreme conditions when they did not occur (November 2003, which was the 3rd hottest and +1σ in NN, but average in CFSR, AN, and, SN).

The diurnal range of the 2-m temperature can illustrate the effects of precipitation on temperature. As demonstrated with the maxima and minima of the daily 2-m temperature, WRF without nudging can sometimes accurately predict extreme events. February and March 1998 and November 1999 were correctly classified with exceptionally small diurnal range by all three WRF runs (Fig. 17), and June 1988 was exceptionally large in all three WRF runs. In other cases, nudging was necessary to intensify (May 1988, July 1988, November 1992, July–September 1993) or mitigate (November 1997) the magnitude of the diurnal range. Interior nudging was necessary to capture the magnitude of the expanded diurnal range during the extreme hot and dry summer of 1988. In addition, nudging correctly reduced the diurnal range during July–September 1993, following the record-breaking flooding events. The annual variability in the diurnal range in NN is erroneously largest in winter months (and enhanced compared to CFSR, AN, and SN), and smallest in summer months (and suppressed compared to CFSR, AN, and SN). This shows that interior nudging is needed to correctly simulate the intraannual and interannual variability in diurnal range.

Month-by-month area-average precipitation totals for the 20-yr period are shown in Fig. 18. Evaluating monthly precipitation totals over a region allows us to remove acute events (which are also important, but discussed as part of Figs. 13 and 14) and assess prolonged synoptic patterns that either increase or decrease widespread precipitation at some point in the year. Based on NARR for the 20-yr period, there were nine individual months with >+2σ area-average precipitation (exceptionally wet) in the Midwest, and one month with <–2σ area-average precipitation (exceptionally dry) in the Midwest (Fig. 18). Without interior nudging in WRF, NN predicted ten exceptionally wet months and no exceptionally dry months. However, of the ten exceptionally wet months identified by NN during the 20-yr period, only four of them actually verified as exceptionally wet; the other six months predicted as exceptionally wet by NN were usually only slightly wetter than average according to NARR. In addition, the exceptionally dry month (June 1988) was predicted to be only abnormally dry (<–1σ) by NN, and it was not even the driest June of the 20-yr period in NN. By contrast, the exceptionally dry year in June 1988 was correctly predicted by both AN and SN at <–2σ. June 1988 had <50% of the area-average monthly precipitation of the next driest June of the 20-yr period in both AN and SN, as in NARR.

AN identified eight exceptionally wet months, and SN identified ten exceptionally wet months. The months identified by AN and SN as exceptionally wet often matched those identified from NARR as exceptionally wet (see Fig. 18). In cases where there was disagreement on the extremity of the precipitation during the month, often the month was the wettest year for that month during the period in NARR and the WRF nudging cases, so the 2σ threshold may have been too strict. By contrast, in cases where NN was inconsistent with NARR, the errors in classifying the extremity of the monthly precipitation were much larger. For example, March 1998 was exceptionally wet (>+2σ) as classified by NARR and as predicted by AN and SN, but it was predicted as slightly wetter than average (between ±1σ) by NN. March 2002 was predicted as exceptionally wet by NN, but verified as slightly wetter than average in NARR and was correctly classified by AN and SN. Problems in NN also persisted in summer months, where August 2007 was an exceptionally wet month in NARR and was correctly predicted by AN and SN as the wettest August of the 20-yr period (Fig. 18); NN, however, classified August 2007 as abnormally dry ($<-1\sigma$). Lastly, the three wettest months during the 20-yr period in NARR were May 2004, June 1998, and July 1992 (Fig. 18). All three of those months were correctly predicted as the top three wet months by AN and SN, while NN did not identify any of those months among the three wettest. Overall, while imperfect and subject to refinement, applying interior nudging toward the coarse-resolution R-2 fields through AN and SN enabled WRF to identify extreme months in the Midwest region that were better matched to NARR than NN. Without interior nudging NN identified the approximate number of extreme wet months, and NN correctly identified four of the ten extreme months during the 20-yr simulation period. However, there were six misses and six false alarms for NN predictions of exceptionally wet months during the 20-yr period (and one egregious miss of the exceptionally dry month), which is unreliable for predicting extreme precipitation.

4. Summary

In this paper, the impacts of interior nudging on the prediction of extremes in regional climate modeling were explored. Using the WRF model as the RCM, three continuous simulations covering 1988–2007 were evaluated where the constraint toward the large-scale driving conditions was exercised either only at the lateral boundaries or via one of the two interior nudging techniques in WRF. The simulations were initialized with reanalysis fields from R-2 as a proxy for a coarse-resolution global climate model.

Comparisons of the spectra from WRF output fields were made against R-2 to determine if the WRF simulations were consistent with the driving model at large scales. Finer-scale comparisons of the WRF simulations were drawn against comparable-resolution reanalyses from the NARR and CFSR products.

We showed that nudging improves the prediction of monthly means over a multi-decadal period, which is consistent with other studies using shorter (1-yr or less) simulations (e.g., Miguez-Macho et al. 2004; Castro et al. 2005; Lo et al. 2008, Rockel et al. 2008; Alexandru et al. 2009; Bowden et al. 2012a). By constraining only at the lateral boundaries, WRF often but not always captures the interannual variability, which is also noted in Bowden et al. (2012b), and some of the extremes. However, interior nudging improves the simulation of the mean 2-m temperature and both the hot and cold extreme thresholds, so nudging improves the distribution and does not simply shift a model bias. Using interior nudging is clearly an advantage for simulating extreme wet and dry precipitation periods during the multi-decadal period. All WRF runs overpredicted precipitation totals through the multi-decadal period (as in Caldwell et al. 2009) regardless of whether nudging was used. Yet, both forms of interior nudging reproduced extreme events with greater accuracy and did not produce the false alarms and misclassifications of events when nudging was not used. Overall, interior nudging preserved the variability in the large scales from the driving fields and adjusted the smaller-scale variability toward the high-resolution reanalyses.

These results should not be used to compare the interior nudging techniques directly because of differences in their fundamental approaches and the variables that are nudged. However, the application of nudging in WRF for regional climate modeling stands to be improved to capitalize on the strengths of both methods. Although analysis nudging is not theoretically applicable for regional climate modeling, using it is preferable to not using interior nudging. Here, the analysis nudging simulation is heuristic because its precipitation means and extremes are consistently more accurate than the other two runs in five of the six regions in our domain, so it is plausible that spectral nudging in WRF can be improved.

Our results clearly indicate that using interior nudging for regional climate modeling *with reasonable settings* will not inappropriately squelch temperature and precipitation extremes over prolonged periods in mid-latitudes. In some cases, increased spatial variability and larger extremes were predicted without using interior nudging, but those predictions were inaccurate. Using an interior constraint toward the large-scale fields is absolutely necessary to consistently predict extreme events that are faithful to the large-scale atmospheric circulation and approach the verified values. Because there is no consensus on whether nudging is appropriate for regional climate modeling (e.g., Rummukainen 2010), this research adds confidence to use nudging for dynamical downscaling particularly when there is an interest in extreme events. Nudging techniques must be used appropriately (i.e., nudging toward synoptic-scale waves for spectral nudging, and using relaxation timescales that are sufficiently long for analysis nudging) to maximize the benefit from them. However, we did not explore whether model biases could be masked and/or exacerbated by nudging. If the downscaling techniques are extended to global climate fields (i.e., Type 3 or Type 4 rather than Type 2, following Castro et al. 2005), then the resultant regional climate projections may include the effects of biases in the global climate fields that will not be overcome by nudging. Our results reflect one configuration of WRF, and the generality of our conclusions should be evaluated for other configurations of WRF and other RCMs. Using historical data, WRF provides realistic regional climatology and captures some interannual variability without interior nudging. However, accurately capturing changes in the interannual variability of critical thresholds of 2-m temperature and precipitation are important to generate credible, problem-focused climate assessments (e.g., Tryhorn and DeGaetano 2011), and that can best be achieved today by using interior nudging techniques in the RCM.

Acknowledgments.

Lara Reynolds and Chris Misenis (CSC) provided technical support to generate some of the simulations shown in this paper. Kiran Alapaty and S.T. Rao (U.S. EPA) provided technical feedback on this paper. The critique of three anonymous reviewers served to strengthen the manuscript. The United States Environmental Protection Agency through its Office of Research and Development funded and managed the research described here. It has been subjected to the Agency's administrative review and approved for publication.

References

Alexandru, A., R. de Elia, R. Laprise, L. Separovic, and S. Biner, 2009: Sensitivity study of regional climate model simulations to large-scale nudging parameters. *Mon. Wea. Rev.*, **137**, 1666-1685.

Arritt, R., and M. Rummukainen, 2011: Challenges in regional-scale climate modeling. *Bull. Amer. Meteor. Soc*., doi:10.1175/2010BAMS2971.1.

Becker, E. J., E. H. Berbery, and R. W. Higgins, 2009: Understanding the characteristics of daily precipitation over the United States using the North American Regional Reanalysis. *J. Climate*, **22**, 6268-6286.

Bowden, J. H., T. L. Otte, C. G. Nolte, and M. J. Otte, 2012a: Examining interior grid nudging techniques using two-way nesting in the WRF model for regional climate modeling. *J. Climate*, **25**, 2805-2823.

Bowden, J. H., C. G. Nolte, and T. L. Otte, 2012b: Using continuous multi-decadal regional climate simulations to examine the impact of the large-scale circulation on the regional climatology. *Clim. Dyn*., submitted.

Bukovsky, M. S., and D. J. Karoly, 2007: A brief evaluation of precipitation from the North American Regional Reanalysis. *J. Hydrometeor*., **8**, 837-846.

Caldwell, P., H.-N. S. Chin, D. C. Bader, and G. Bala, 2009: Evaluation of a WRF dynamical downscaling simulation over California. *Climatic Change*, **95**, 499–521.

Castro, C. L., R. A. Pielke, Sr., and G. Leoncini, 2005: Dynamical downscaling: Assessment of value retained and added using the Regional Atmospheric

Modeling System (RAMS). *J. Geophys. Res*., **110**, D05108, doi:10.1029/2004JD004721.

Cha, D.-H., C.-S. Jin, D.-K. Lee, and Y.-H. Kuo, 2011: Impact of intermittent spectral nudging on regional climate simulation using Weather Research and Forecasting model. *J. Geophys. Res*., **116**, D10103, doi:10.1029/2010JD015069.

Chen, F., and J. Dudhia, 2001: Coupling and advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part I: model implementation and sensitivity. *Mon. Wea. Rev*., **129**, 569–585.

Christensen, J. H., and Coauthors, 2007: Regional climate projections. 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 [Solomon, S., et al., Eds.], Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Colin, J., M. Déqué, R. Radu, and S. Somot, 2010: Sensitivity study of heavy precipitation in Limited Area Model climate simulations: influence of the size of the domain and the use of the spectral nudging technique. *Tellus*, **62A**, 591–604.

Davies, H. C., and R. E. Turner, 1977: Updating prediction models by dynamical relaxation: An examination technique. *Quart. J. Roy. Meteor. Soc*., **103**, 225–245.

Deng, A., D. R. Stauffer, J. Dudhia, T. L. Otte, and G. K. Hunter, 2007: Update on analysis nudging FDDA in WRF-ARW. Proceedings, *8th WRF Users' Workshop*, Boulder, CO, National Center for Atmospheric Research, 4.8. [Available online at http://www.mmm.ucar.edu/wrf/users/workshops/WS2007/abstracts/4-8_Deng.pdf.]

Dulière, V., Y. Zhang, and E. P. Salathé, Jr., 2011: Extreme precipitation and temperature over the U.S. Pacific Northwest: A comparison between observations, reanalysis data, and regional models. *J. Climate*, **24**, 1950–1964.

Feser, F., B. Rockel, H. von Storch, J. Winterfeldt, and M. Zahn, 2011: Regional climate models add value to global model data: A review and selected examples. *Bull. Amer. Meteor. Soc.*, **92**, 1181–1192.

Garrett, C., and P. Müller, 2008: Extreme events. *Bull. Amer. Meteor. Soc.*, **89**, 1733.

Giorgi, F., 1990: Simulation of regional climate using a limited area model nested in a general circulation model. *J. Climate*, **3**, 941–963.

Grell, G. A., and D. Dévényi, 2002: A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. *Geophys. Res. Lett.*, **29**, 1963

Guirguis, K. J., and R. Avissar, 2008: An analysis of precipitation variability, persistence, and observational data uncertainty in the western United States. *J. Hydrometeor.*, **9**, 843–865.

Higgins, R. W., V. E. Kousky, V. B. S. Silva, E. Becker, and P. Xie, 2010: Intercomparison of daily precipitation statistics of the United States in observations and in NCEP reanalysis products. *J. Climate*, **23**, 4637–4650.

Hong, S.-Y., and J.-O. J. Lim, 2006: The WRF single-moment 6-class microphysics scheme (WSM6). *J. Korean Meteor. Soc.*, **42**, 2, 129–151.

Hong, S.-Y., Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev*., **134**, 2318–2341.

Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins, 2008: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *J. Geophys. Res*., **113**, D13103, doi:10.1029/2008JD009944.

Kaimal, J. C., and L. Kristensen, 1991: Time series tapering for short data samples. *Bound.-Layer Meteor*, **57**, 187–194.

Kanamaru, H., and M. Kanamitsu, 2007: Scale-selective bias correction in a downscaling of global analysis using a regional model. *Mon. Wea. Rev*., **135**, 334–350.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter, 2002: NCEP–DOE AMIP-II Reanalysis (R-2). *Bull. Amer. Meteor. Soc.*, **83**, 1631–1643.

Karl, T. R., N. Nicholls, and A. Ghazi, 1999: CLIVAR/GCOS/WMO workshop on indices and indicators for climate extremes: Workshop summary. *Climatic Change*, **42**, 3–7.

Katz, R. W., and B. G. Brown, 1992: Extreme events in a changing climate: Variability is more important than averages. *Climatic Change*, **21**, 289–302.

Lo, J. C.-F., Z.-L. Yang, and R. A. Pielke, Sr., 2008: Assessment of three dynamical downscaling methods using the Weather Research and Forecasting (WRF) model. *J. Geophys. Res*., **113**, D09112, doi:10.1029/2007JD009216.

Lynn, B. H., R. Healy, and L. M. Druyan, 2007: An analysis of the potential for extreme temperature change based on observations and model simulations. *J. Climate*, **20**, 1539–1554.

Meehl, G. A., and Coauthors, 2000a: An introduction to trends in extreme weather and climate events: Observations, socioeconomic impacts, terrestrial ecological impacts, and model projections. *Bull. Amer. Meteor. Soc.*, **81**, 413–416.

Meehl, G. A., F. Zwiers, J. Evans, T. Knutson, L. Mearns, and P. Whetton, 2000b: Trents in extreme weather and climate events: Issues related to modeling extremes in projections of future climate change. *Bull. Amer. Meteor. Soc.*, **81**, 427–436.

Mesinger, F., and Coauthors, 2006: North American Regional Reanalysis. *Bull. Amer. Meteor. Soc.*, **87**, 343–360.

Miguez-Macho, G., G. L. Stenchikov, and A. Robock, 2004: Spectral nudging to eliminate the effects of domain position and geometry in regional climate model simulations. *J. Geophys. Res.*, **109**, D13104, doi:10.1029/2003JD004495.

Mladjic, B., L. Sushama, M. N. Khaliq, R. Laprise, D. Caya, and R. Roy, 2011: Canadian RCM projected changes to extreme precipitation characteristics over Canada. *J. Climate*, **24**, 1565–2584.

Mo, K. C., L. N. Long, Y. Xia, S. K. Yang, J. E. Schemm, and M. Ek, 2001: Drought indices based on the Climate Forecast System Reanalysis and ensemble NLDAS. *J. Hydrometeor.*, **12**, 181–205.

Nigam, S., and A. Ruiz-Barradas, 2006: Seasonal hydroclimate variability over North America in global and regional reanalyses and AMIP simulations: varied representation. *J. Climate*, **19**, 815–837.

Otte, T. L., 2008: The impact of nudging in the meteorological model for retrospective air quality simulations. Part I: Evaluation against national observation networks. *J. Appl. Meteor. Climatol.*, **47**, 1853–1867.

Pielke, R. A., 1984: Mesoscale Meteorological Modeling. Academic Press, 612 pp.

Pielke, Sr., R. A., R. Wilby, D. Niyogi, F. Hossain, K. Dairuku, J. Adogoke, G. Kallos, T. Seastedt, and K. Suding, 2012: Dealing with complexity and extreme events using a bottom-up, resource-based vulnerability perspective. *AGU Monograph on Complexity and Extreme Events in Geosciences*, Amer. Geophys. Union, in press.

Radu, R., M. Déqué, and S. Somot, 2008: Spectral nudging in a spectral regional climate model. *Tellus*, **60A**, 898–910.

Ruiz-Barradas, A., and S. Nigam, 2006: IPCC's twentieth-century climate simulations: Varied representations of North American hydroclimate variability. *J. Climate*, **19**, 4041–4058.

Rummukainen, M., 2010: State-of-the-art with regional climate models. *WIREs Climate Change*, **1**, 82–96.

Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, doi:10.1175/2010BAMS3001.1.

Stauffer, D. R., and N. L. Seaman, 1990: Use of four-dimensional data assimilation in a limited-area model. Part I: Experiments with synoptic-scale data. *Mon. Wea. Rev.*, **118**, 1250–1277.

_____, and _____, 1994: Multiscale four-dimensional data assimilation. *J. Appl. Meteor*., 33, 416–434.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, M. G. Duda, X.-Y. Huang, W. Wang, and J. G. Powers, 2008: A description of the Advanced Research WRF Version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp.

Trenberth, K. E., and Coauthors, 2007: *Observations: Surface and Atmospheric Climate Change*. In: *Climate Change 2007. The Physical Science Basis*. Contribution of WG 1 to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. S. Solomon, D. Qin, M. Manning, Z. Chen, M. C. Marquis, K. B. Averyt, M. Tignor and H. L. Miller (Eds.) Cambridge University Press. Cambridge, U. K., and New York, NY, USA, 235–336, plus annex online.

Tryhorn, L., and A. DeGaetano, 2011: "2100? It doesn't keep me up at night!" Lessons for the next generation of climate assessments. *Bull. Amer. Meteor. Soc*., **92**, 1137–1148.

von Storch, H. H. Langenberg, and F. Feser, 2000: A spectral nudging technique for dynamical downscaling purposes. *Mon. Wea. Rev*., **128**, 3664–3673.

Wang, W., P. Xie, S.-H. Yoo, Y. Xue, A. Kumar, and X. Wu, 2011: An assessment of the surface climate in the NCEP climate forecast system reanalysis. *Clim. Dyn*., **37**, 1601–1620.

Werth, D., and A. Garrett, 2011: Patterns of land surface errors and biases in the Global Forecast System. *Mon. Wea. Rev*., **139**, 1569–1582.

List of Figures

FIG. 1. WRF 108-km and 36-km domains, and the regions used for model evaluation: Northwest (NW), Southwest (SW), Plains (PL), Midwest (MW), Southeast (SE), and Northeast (NE). From Bowden et al. (2012a).

FIG. 2. Mean 2-m temperature difference (K) from CFSR for 1988–2007 from NARR and from WRF simulations NN, AN, and SN.

FIG. 3. Mean annual precipitation (mm) for 1988–2007 from NARR and from WRF simulations NN, AN, and SN.

FIG. 4. Monthly area-averaged (a) 2-m temperature difference from CFSR (K) and (b) precipitation difference from NARR (mm) for the Midwest region (refer to Fig. 1) for three WRF runs: NN (green), AN (blue), and SN (red).

FIG. 5. 20-yr-average of monthly area-averaged 2-m temperature difference from CFSR (K) for 6 regions (refer to Fig. 1) for three WRF runs: NN (“N”), AN (“A”), and SN (“S”).

FIG. 6. 20-yr-average of monthly area-averaged precipitation difference from NARR (mm) for 6 regions (refer to Fig. 1) for three WRF runs: NN (“N”), AN (“A”), and SN (“S”).

FIG. 7. Spectral variance from 500-hPa temperature for R-2, NARR, and WRF simulations NN, AN, and SN averaged for (a) January, and (b) July.

FIG. 8. Same as Fig. 7, but for 700-hPa water vapor mixing ratio.

FIG. 9. Low-frequency kinetic energy spectra for R-2, NARR, and WRF simulations NN, AN, and SN averaged for (a) January 250-hPa meridional wind, (b) July 250-hPa meridional wind, (c) January 500-hPa meridional wind, and (d) July 500-hPa meridional wind.

FIG. 10. Annual area-averaged number of days with 2-m temperature above 90°F for the Midwest region. Data are shown from CFSR ("O") and WRF runs NN ("N"), AN ("A"), and SN ("S").

FIG. 11. Annual area-averaged number of days with (a) 2-m temperature below 32°F, (b) maximum 2-m temperature below 32°F, and (c) 2-m temperature below 0°F for the Midwest region. Data are shown from CFSR ("O") and WRF runs NN ("N"), AN ("A"), and SN ("S").

FIG. 12. 20 years of annual area-averaged number of days with 2-m temperature greater than 90°F (DT90, gray) and less than 32°F (DT32, white) for the 6 regions in Fig. 1. Data are shown from CFSR and WRF runs NN, AN, and SN. Boxes are drawn from 25th

to 75^{th} percentiles with 50^{th} percentile shown in center of each box, and whiskers at minimum and maximum values.

FIG. 13. Annual area-averaged number of days with (a) precipitation greater than 0.5 in and (b) precipitation greater than 1.0 in for the Midwest region. Data are shown from NARR ("O") and WRF runs NN ("N"), AN ("A"), and SN ("S").

FIG. 14. 20 years of annual area-averaged number of days with precipitation greater than 0.5 in (DP05, gray) and precipitation greater than 1.0 in (DP10, white) for the 6 regions in Fig. 1. Data are shown from NARR and WRF runs NN, AN, and SN. Boxes are drawn from 25^{th} to 75^{th} percentiles with 50^{th} percentile shown in center of each box, and whiskers at minimum and maximum values.

FIG. 15. Monthly area-averaged daily maximum 2-m temperature (K) for the Midwest region for 1988-2007. Data are shown from CFSR (upper-left) and WRF runs NN (upper-right), AN (lower-left), and SN (lower-right). The solid black line indicates the 20-yr, monthly mean of the daily maximum 2-m temperature, the dashed black lines indicate ±1 standard deviation from the mean, and the gray shading indicates ±2 standard deviations from the mean. The data are color-coded by year, with the earliest years in blues progressing to reds in the later years.

FIG. 16. Same as Fig. 15, but for daily minimum 2-m temperature (K).

FIG. 17. Same as Fig. 15, but for daily diurnal range (K).

FIG. 18. Same as Fig. 15, but for precipitation (mm).

TABLE 1. Nudging coefficients (s^{-1}) and domain-relative wave numbers used for analysis and spectral nudging simulations. Time scales (h) that correspond to the nudging coefficients and length scales (km) that correspond to the wave numbers are in parentheses. Fields that are not applicable are indicated by –.

	Wind	Potential Temp.	Water Vapor Mixing Ratio	Geo-potential	West-east wave number	South-north wave number
Analysis Nudging (108-km)	3.0×10^{-4} (0.9)	3.0×10^{-4} (0.9)	4.5×10^{-5} (6.2)	–	–	–
Analysis Nudging (36-km)	1.0×10^{-4} (2.8)	1.0×10^{-4} (2.8)	1.0×10^{-5} (27.8)	–	–	–
Spectral Nudging (108-km)	3.0×10^{-4} (0.9)	3.0×10^{-4} (0.9)	–	3.0×10^{-4} (0.9)	5 (1728)	3 (1800)
Spectral Nudging (36-km)	3.0×10^{-4} (0.9)	3.0×10^{-4} (0.9)	–	3.0×10^{-4} (0.9)	4 (1674)	2 (1512)

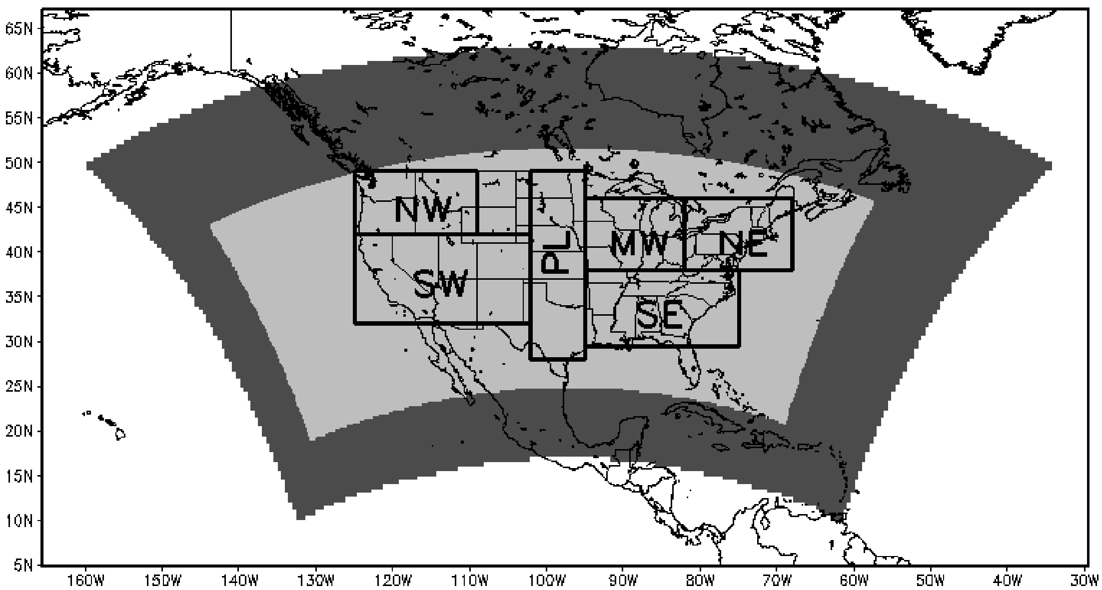

FIG. 1. WRF 108-km and 36-km domains, and the regions used for model evaluation: Northwest (NW), Southwest (SW), Plains (PL), Midwest (MW), Southeast (SE), and Northeast (NE). From Bowden et al. (2012a).

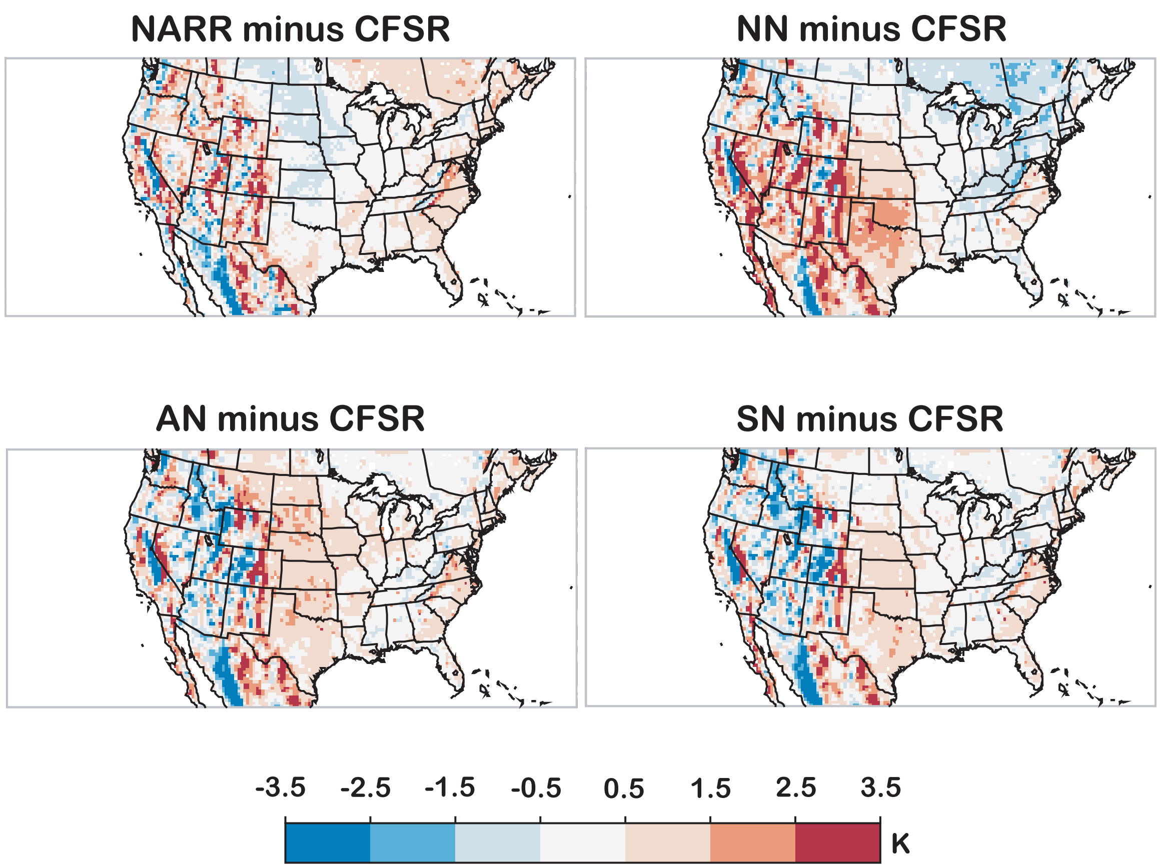

FIG. 2. Mean 2-m temperature difference (K) from CFSR for 1988–2007 from NARR and from WRF simulations NN, AN, and SN.

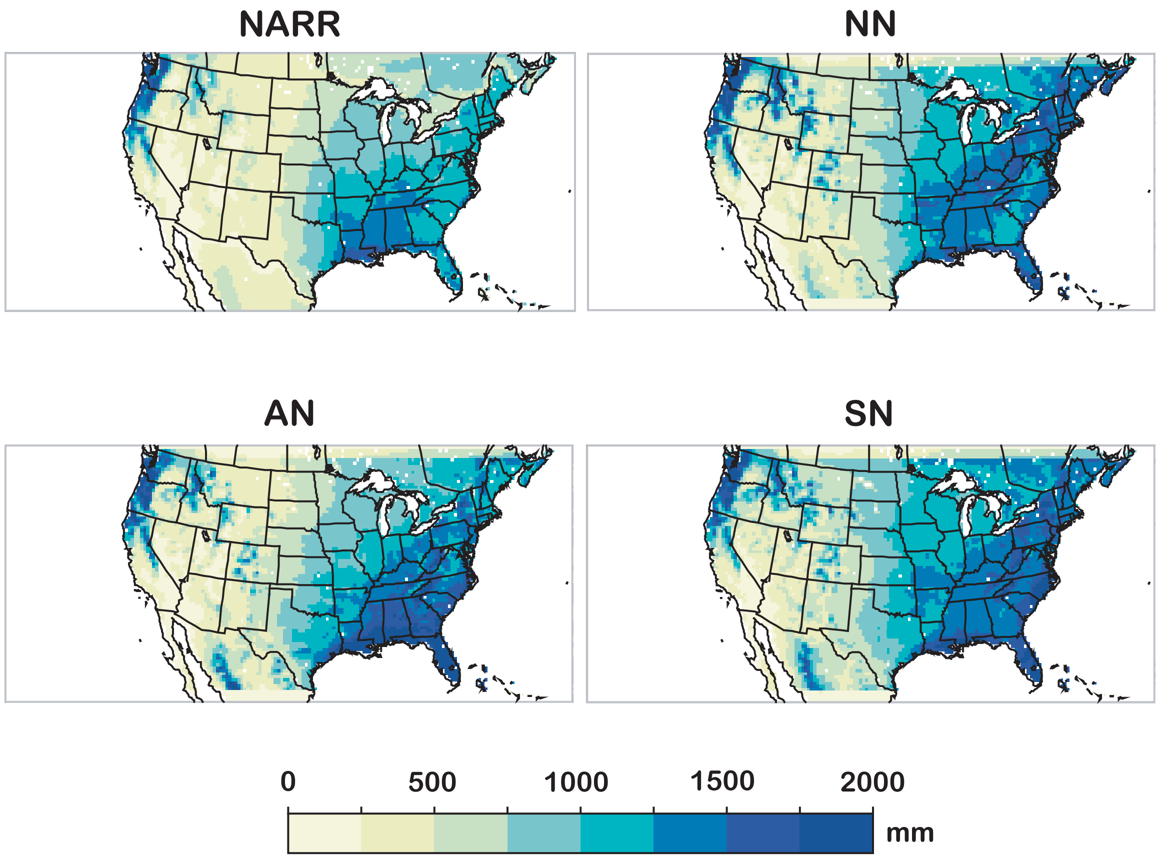

FIG. 3. Mean annual precipitation (mm) for 1988–2007 from NARR and from WRF simulations NN, AN, and SN.

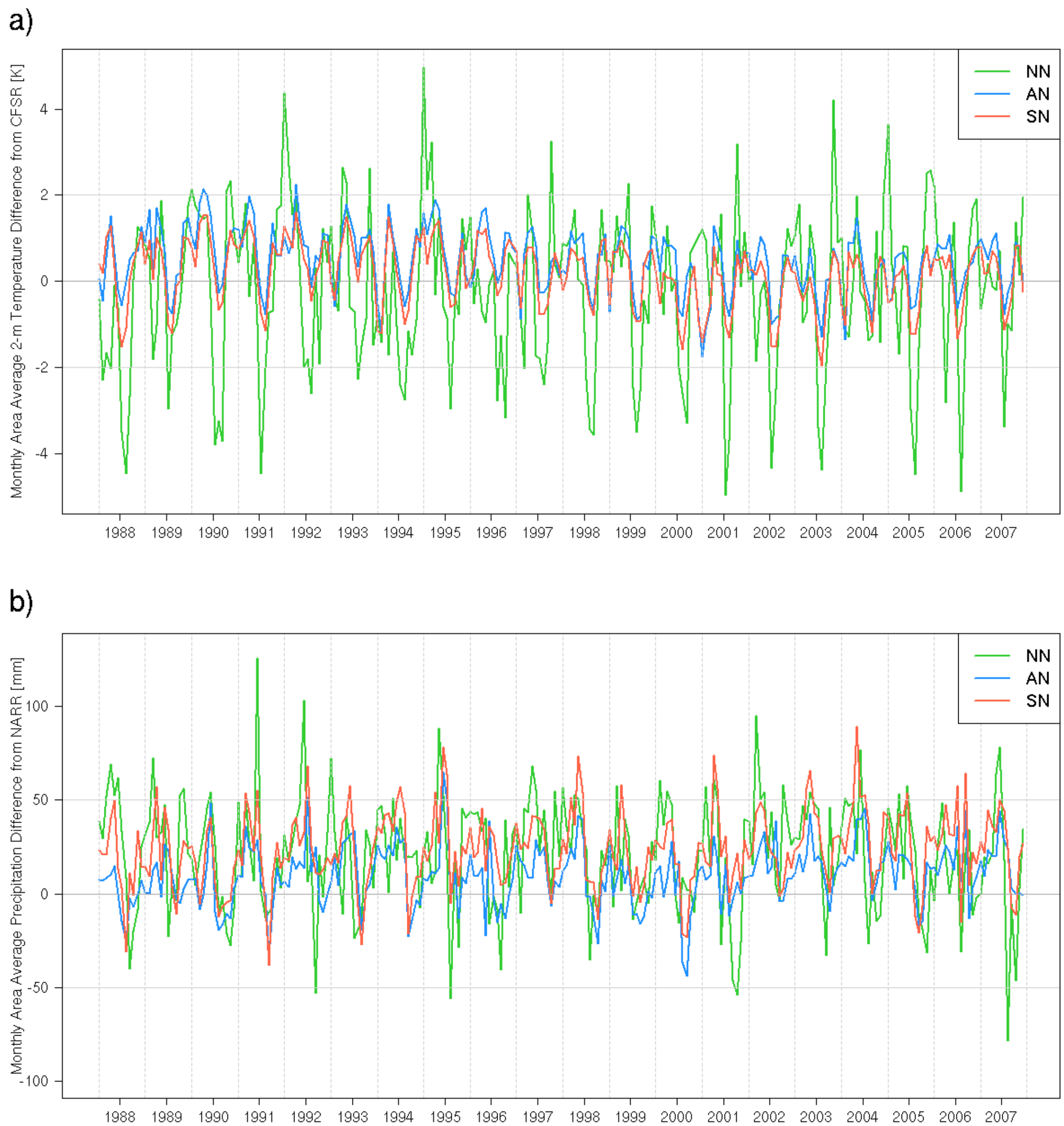

FIG. 4. Monthly area-averaged (a) 2-m temperature difference from CFSR (K) and (b) precipitation difference from NARR (mm) for the Midwest region (refer to Fig. 1) for three WRF runs: NN (green), AN (blue), and SN (red).

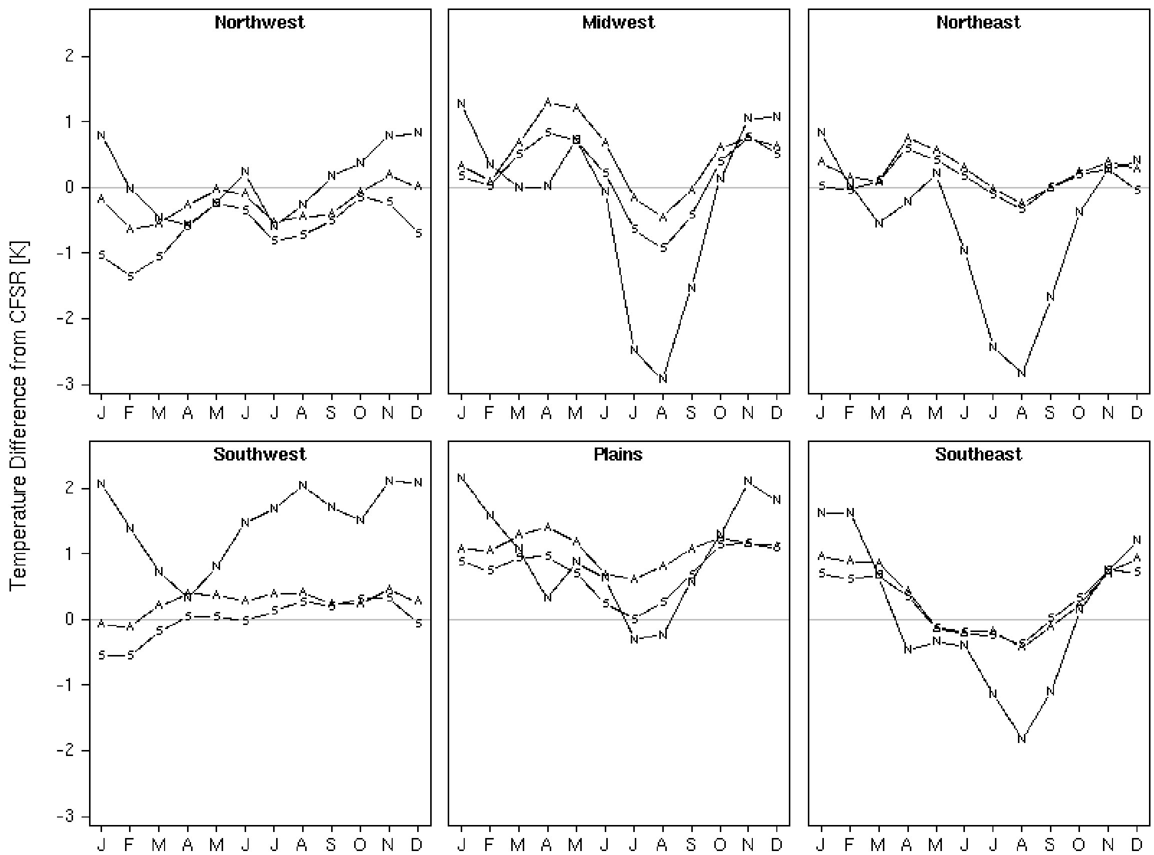

FIG. 5. 20-yr-average of monthly area-averaged 2-m temperature difference from CFSR (K) for 6 regions (refer to Fig. 1) for three WRF runs: NN ("N"), AN ("A"), and SN ("S").

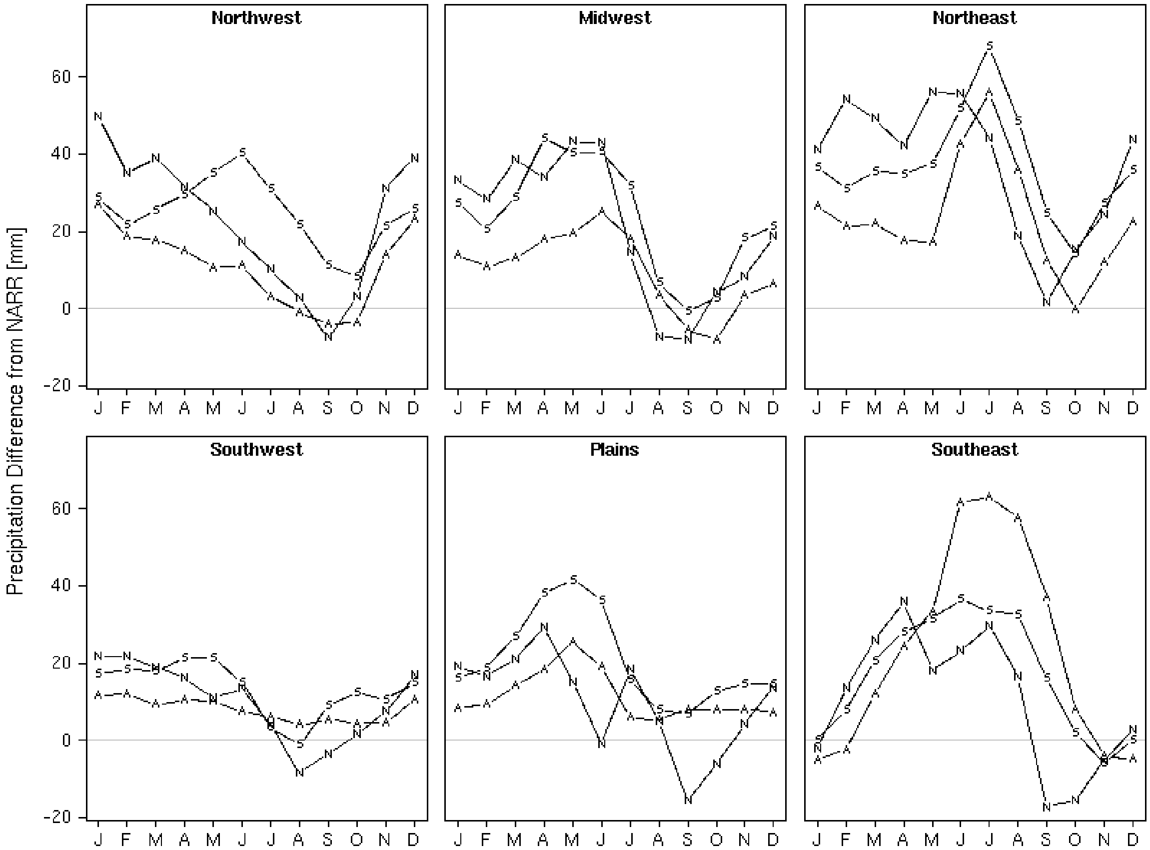

FIG. 6. 20-yr-average of monthly area-averaged precipitation difference from NARR (mm) for 6 regions (refer to Fig. 1) for three WRF runs: NN ("N"), AN ("A"), and SN ("S").

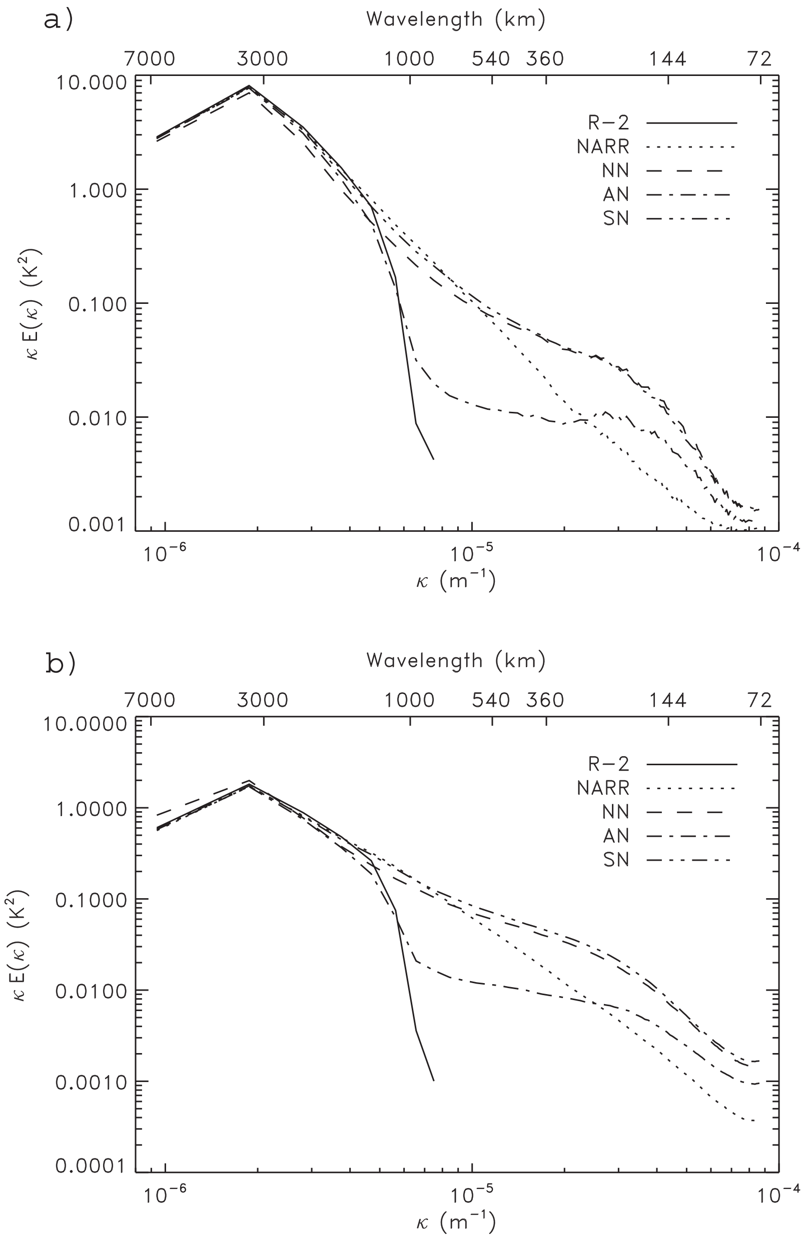

FIG. 7. Spectral variance from 500-hPa temperature for R-2, NARR, and WRF simulations NN, AN, and SN averaged for (a) January, and (b) July.

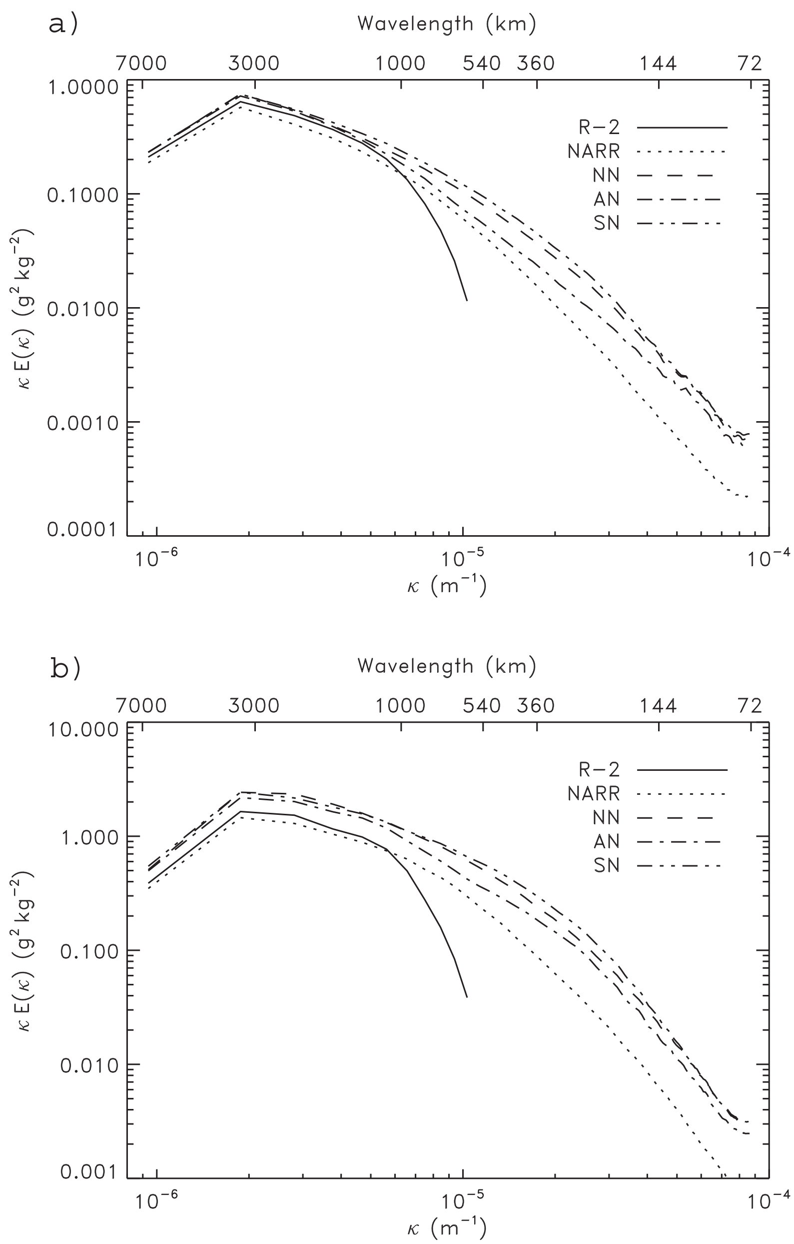

FIG. 8. Same as Fig. 7, but for 700-hPa water vapor mixing ratio.

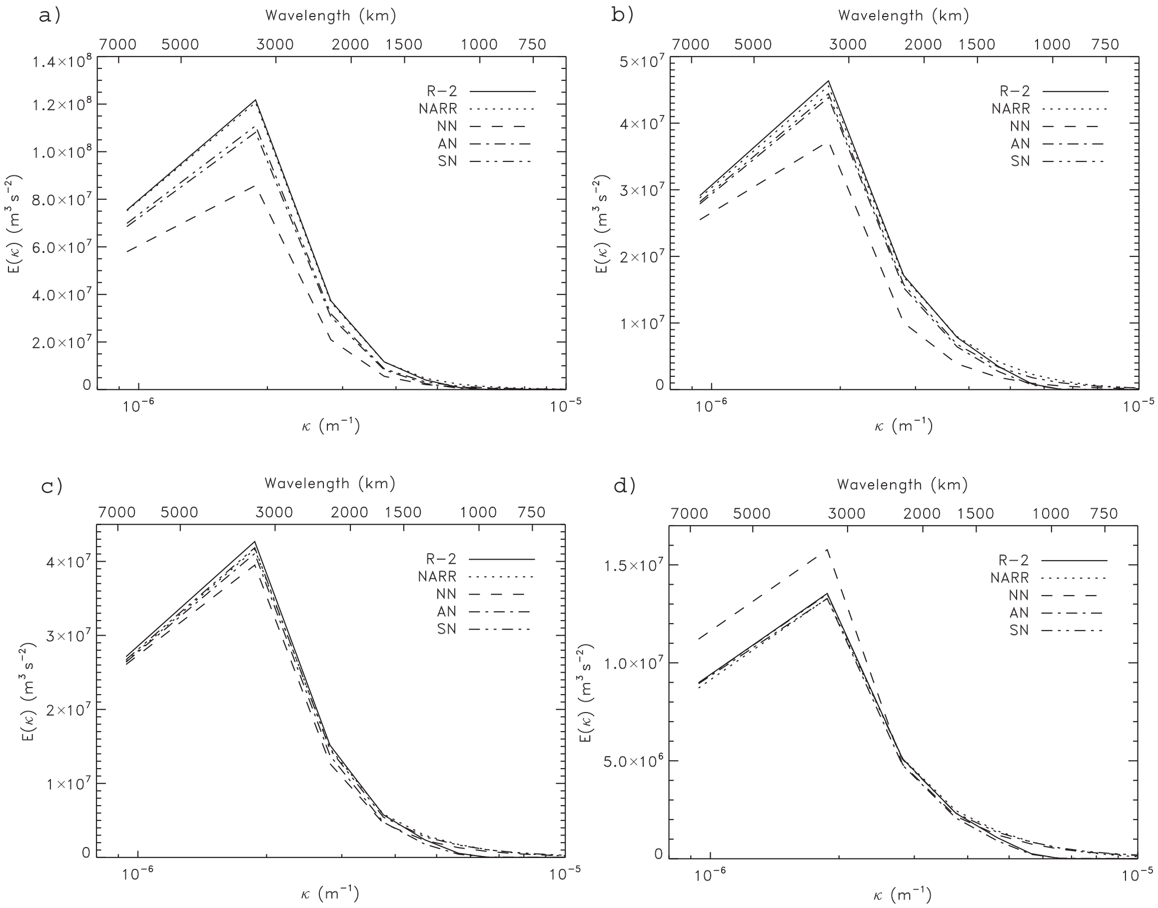

FIG. 9. Low-frequency kinetic energy spectra for R-2, NARR, and WRF simulations NN, AN, and SN averaged for (a) January 250-hPa meridional wind, (b) July 250-hPa meridional wind, (c) January 500-hPa meridional wind, and (d) July 500-hPa meridional wind.

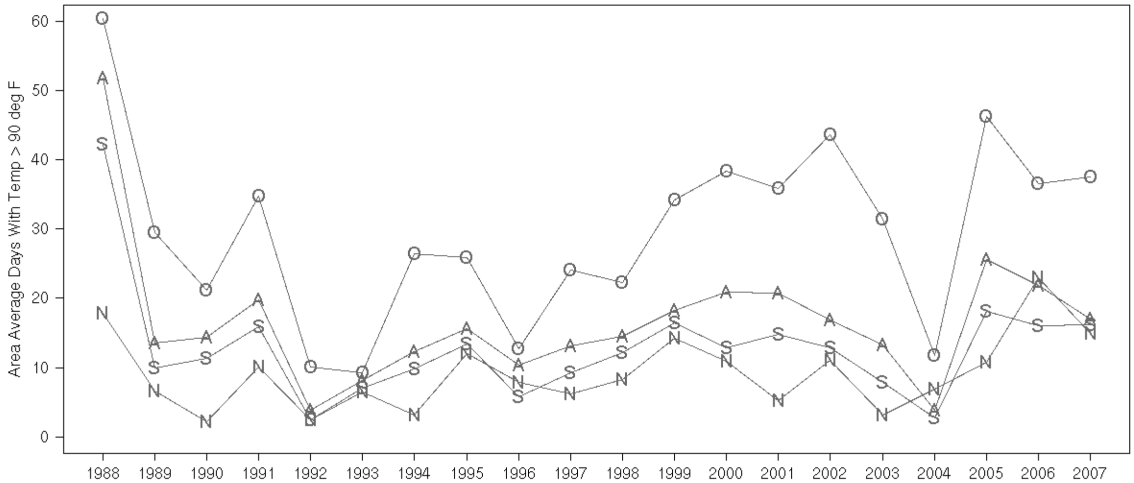

FIG. 10. Annual area-averaged number of days with 2-m temperature above 90°F for the Midwest region. Data are shown from CFSR (“O”) and WRF runs NN (“N”), AN (“A”), and SN (“S”).

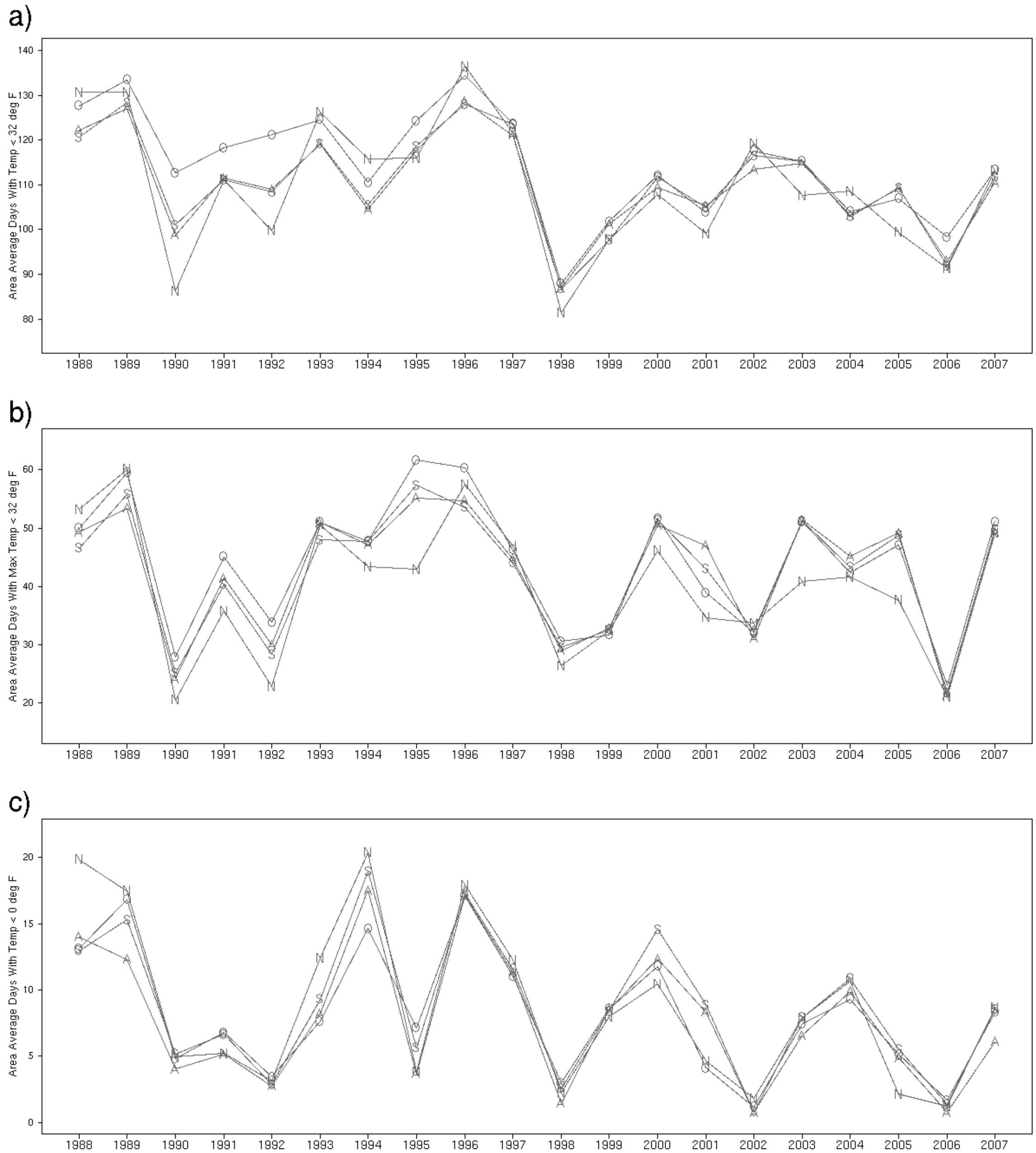

FIG. 11. Annual area-averaged number of days with (a) 2-m temperature below 32°F, (b) maximum 2-m temperature below 32°F, and (c) 2-m temperature below 0°F for the Midwest region. Data are shown from CFSR (“O”) and WRF runs NN (“N”), AN (“A”), and SN (“S”).

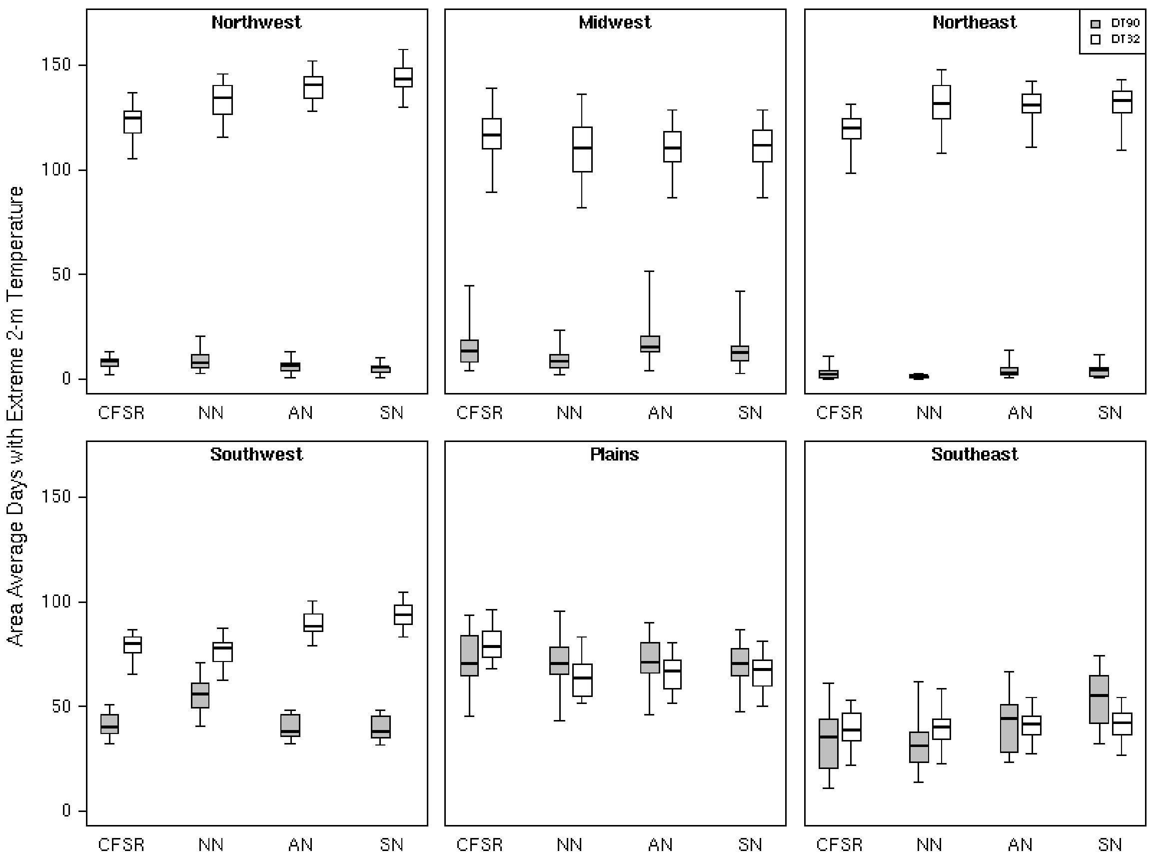

FIG. 12. 20 years of annual area-averaged number of days with 2-m temperature greater than 90°F (DT90, gray) and less than 32°F (DT32, white) for the 6 regions in Fig. 1. Data are shown from CFSR and WRF runs NN, AN, and SN. Boxes are drawn from 25th to 75th percentiles with 50th percentile shown in center of each box, and whiskers at minimum and maximum values.

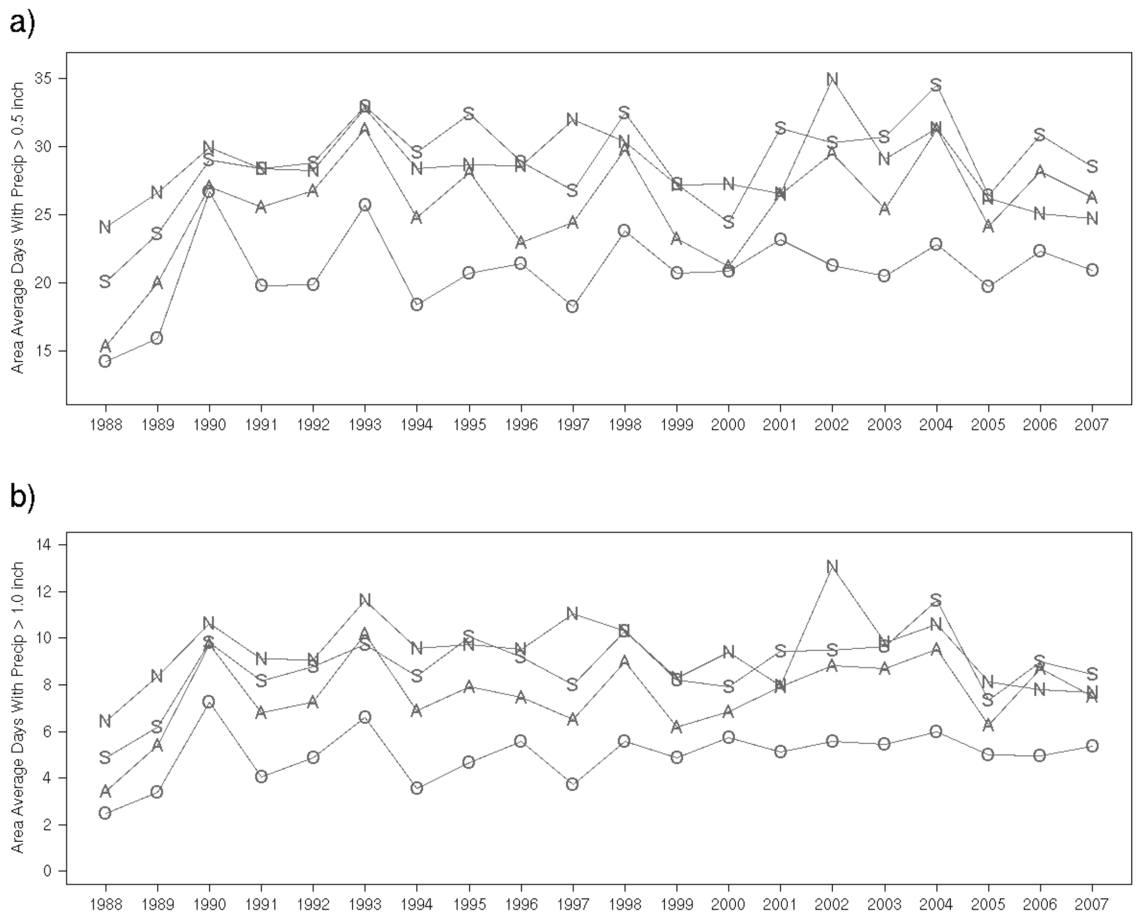

FIG. 13. Annual area-averaged number of days with (a) precipitation greater than 0.5 in and (b) precipitation greater than 1.0 in for the Midwest region. Data are shown from NARR (“O”) and WRF runs NN (“N”), AN (“A”), and SN (“S”).

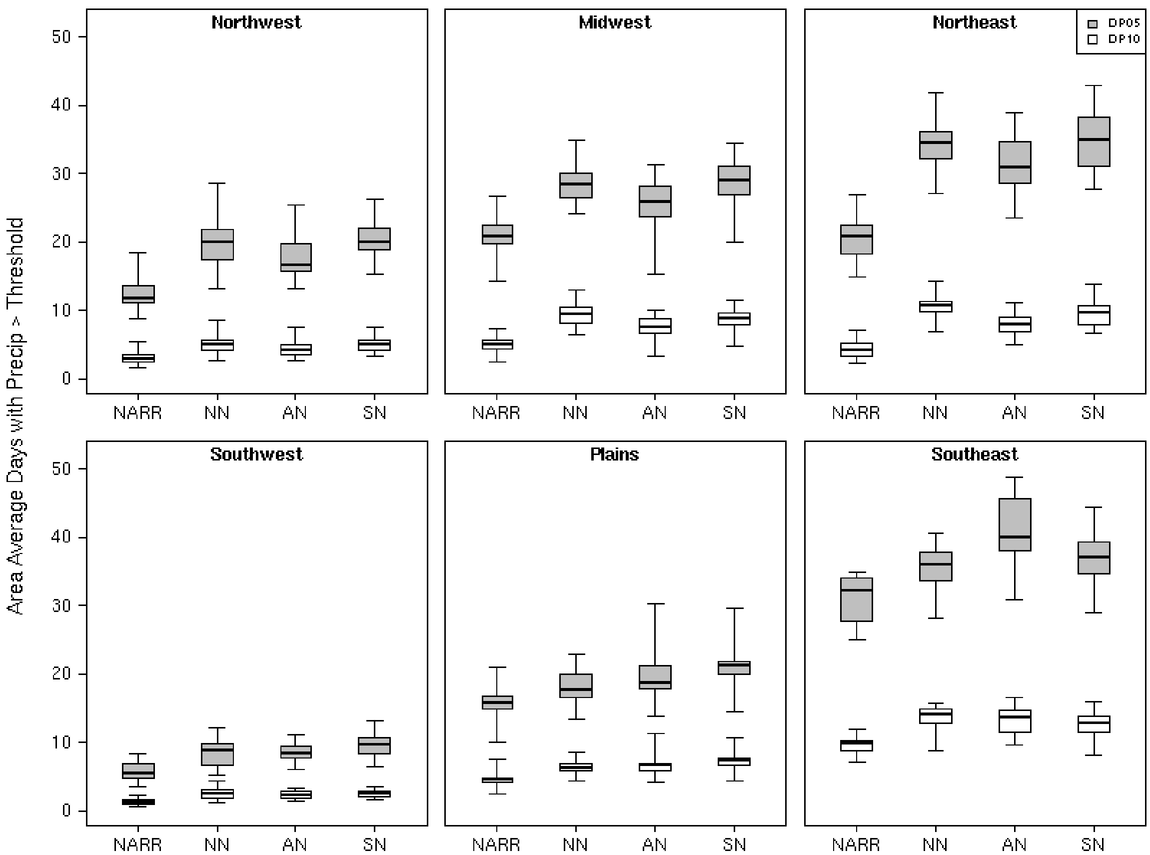

FIG. 14. 20 years of annual area-averaged number of days with precipitation greater than 0.5 in (DP05, gray) and precipitation greater than 1.0 in (DP10, white) for the 6 regions in Fig. 1. Data are shown from NARR and WRF runs NN, AN, and SN. Boxes are drawn from 25th to 75th percentiles with 50th percentile shown in center of each box, and whiskers at minimum and maximum values.

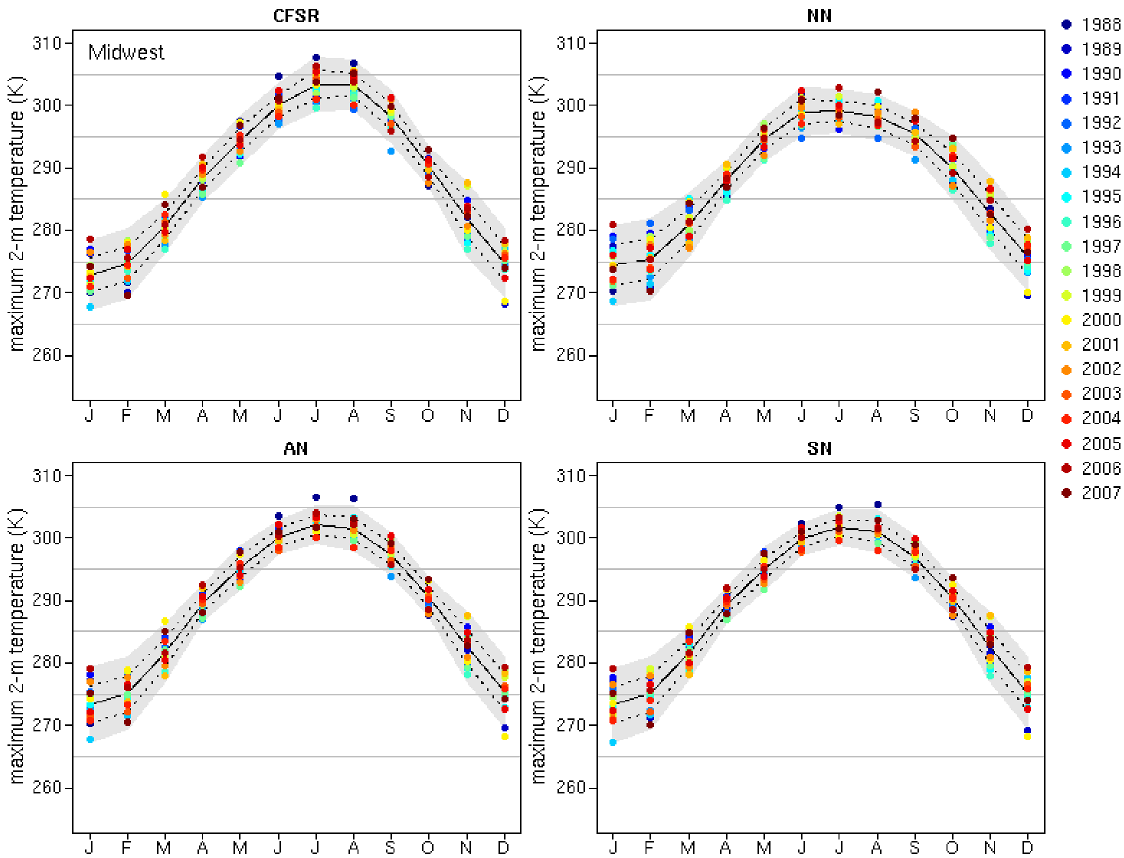

FIG. 15. Monthly area-averaged daily maximum 2-m temperature (K) for the Midwest region for 1988-2007. Data are shown from CFSR (upper-left) and WRF runs NN (upper-right), AN (lower-left), and SN (lower-right). The solid black line indicates the 20-yr, monthly mean of the daily maximum 2-m temperature, the dashed black lines indicate ±1 standard deviation from the mean, and the gray shading indicates ±2 standard deviations from the mean. The data are color-coded by year, with the earliest years in blues progressing to reds in the later years.

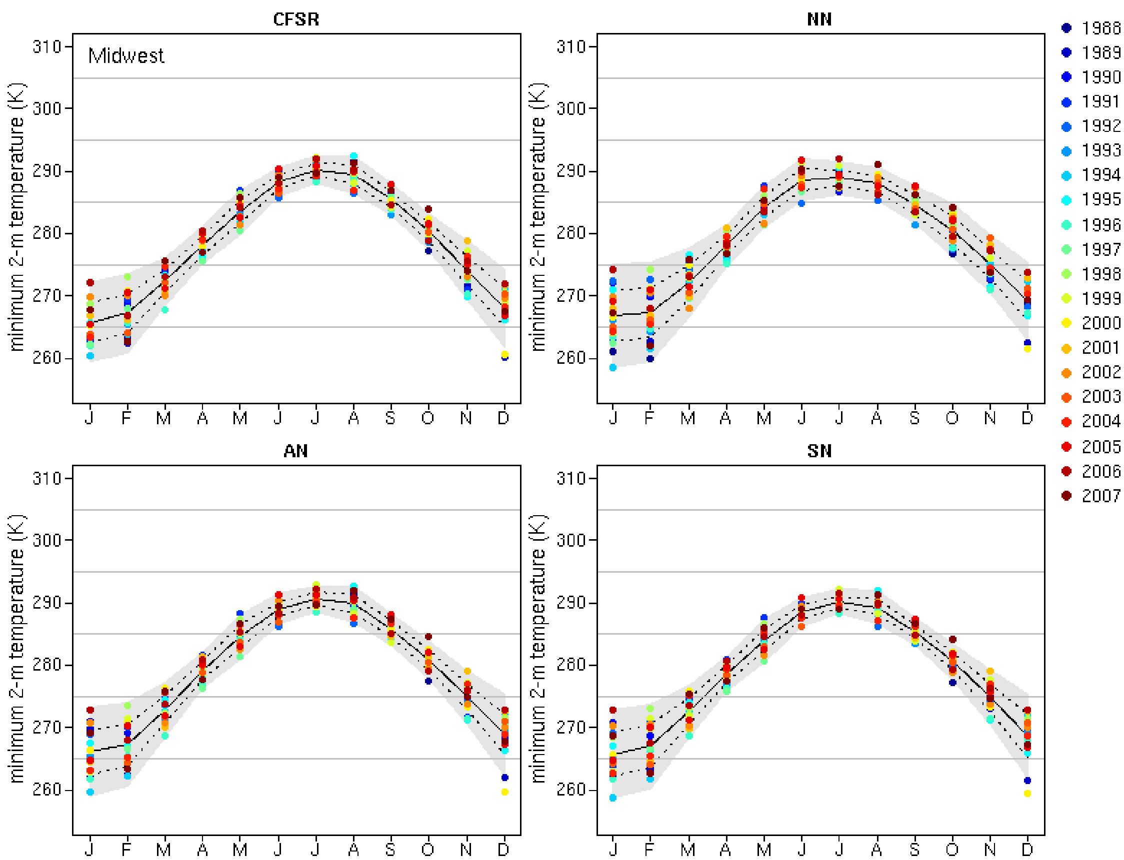

FIG. 16. Same as Fig. 15, but for daily minimum 2-m temperature (K).

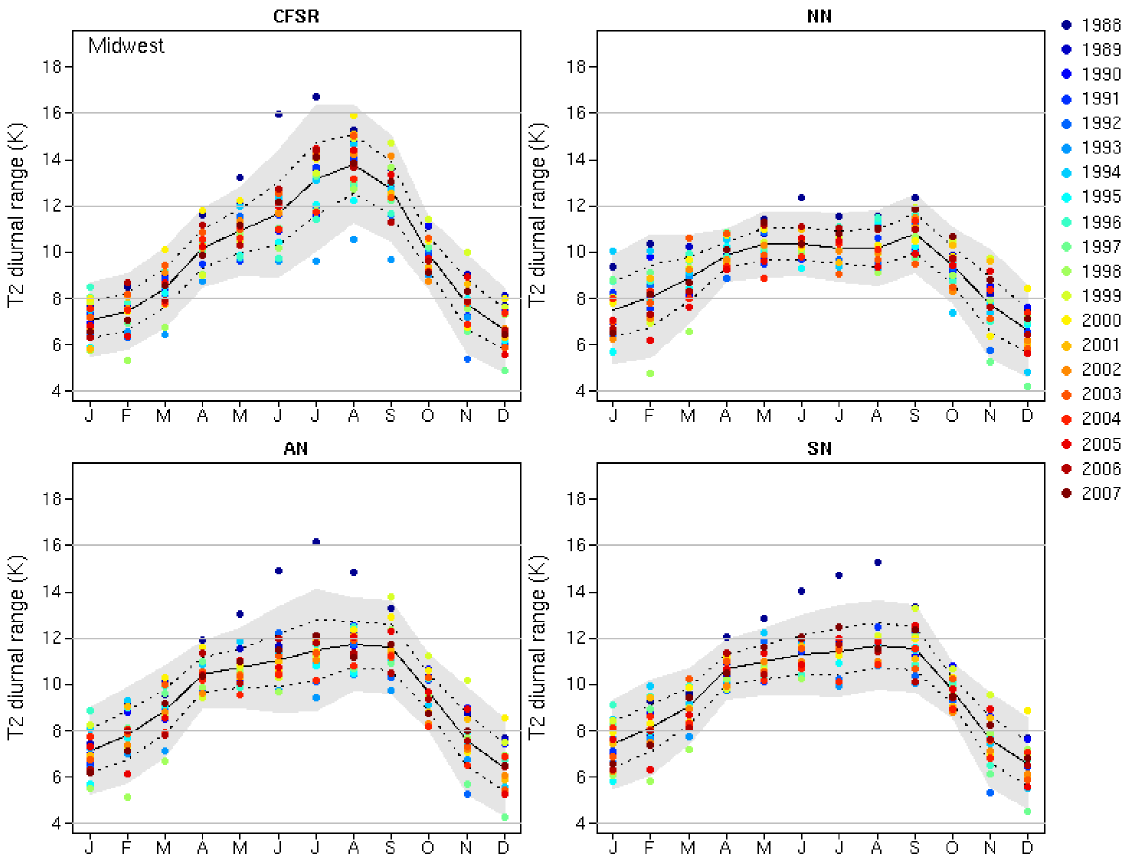

FIG. 17. Same as Fig. 15, but for daily diurnal range (K).

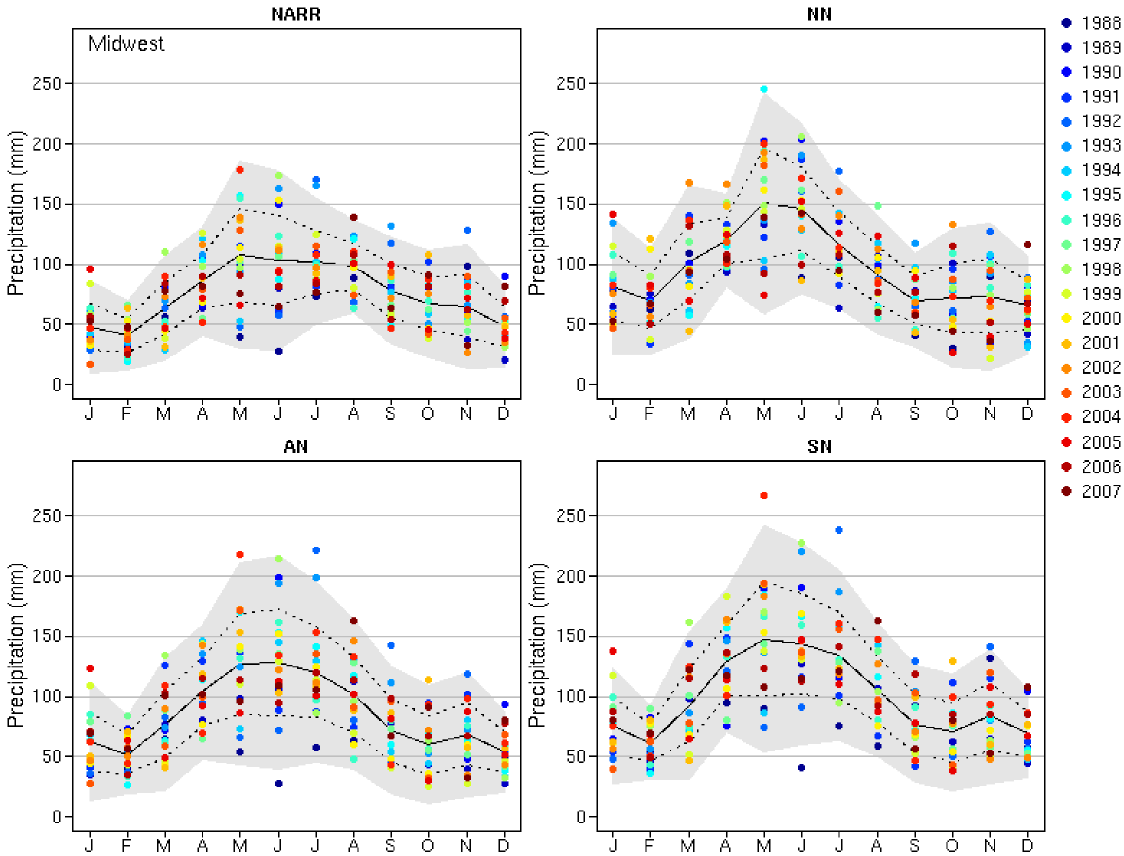

FIG. 18. Same as Fig. 15, but for precipitation (mm).