High-resolution subtropical summer precipitation derived from dynamical downscaling of the NCEP/DOE reanalysis: how much small-scale information is added by a regional model?

Young-Kwon Lim · Lydia B. Stefanova · Steven C. Chan · Siegfried D. Schubert · James J. O’Brien

Received: 1 April 2010 / Accepted: 29 July 2010 / Published online: 10 August 2010
© Springer-Verlag 2010

Abstract This study assesses the regional-scale summer precipitation produced by the dynamical downscaling of analyzed large-scale fields. The main goal of this study is to investigate how much the regional model adds smaller scale precipitation information that the large-scale fields do not resolve. The modeling region for this study covers the southeastern United States (Florida, Georgia, Alabama, South Carolina, and North Carolina) where the summer climate is subtropical in nature, with a heavy influence of regional-scale convection. The coarse resolution (2.5° latitude/longitude) large-scale atmospheric variables from the National Center for Environmental Prediction (NCEP)/DOE reanalysis (R2) are downscaled using the NCEP/Environmental Climate Prediction Center regional spectral model (RSM) to produce precipitation at 20 km resolution for 16 summer seasons (1990–2005). The RSM produces realistic details in the regional summer precipitation at 20 km resolution. Compared to R2, the RSM-produced monthly precipitation shows better agreement with observations. There is a reduced wet bias and a more realistic spatial pattern of the precipitation climatology compared with the interpolated R2 values. The root mean square errors of the monthly R2 precipitation are reduced over 93% (1,697) of all the grid points in the five states (1,821). The temporal correlation also improves over 92% (1,675) of all grid points such that the domain-averaged correlation increases from 0.38 (R2) to 0.55 (RSM). The RSM accurately reproduces the first two observed eigenmodes, compared with the R2 product for which the second mode is not properly reproduced. The spatial patterns for wet versus dry summer years are also successfully simulated in RSM. For shorter time scales, the RSM resolves heavy rainfall events and their frequency better than R2. Correlation and categorical classification (above/near/below average) for the monthly frequency of heavy precipitation days is also significantly improved by the RSM.

1 Introduction

Regional climate models (RCMs) have been actively studied for many years as tools for producing fine scale climate information over particular regions (e.g., Dickinson et al. 1989; Giorgi 1990; Giorgi et al. 1992; Liu et al. 1994; Takle et al. 1999; Fennessy and Shukla 2000; Anderson et al. 2003; Roads et al. 2003; Liang et al. 2004; Fu et al. 2005; Castro et al. 2007; Seo et al. 2007; Mearns et al. 2009). RCMs, however, are still limited in their ability to accurately reproduce the important small-scale atmospheric features from the larger-scale lateral boundary conditions (Castro et al. 2005; Rockel et al. 2008). In particular, as discussed in Jenkins (1997) and Kunkel et al. (2002), accurate simulation of the seasonal summer precipitation, particularly for subtropical and tropical areas, remains a major challenge due to the frequent local-scale convective activity that occurs throughout the season. The present study takes on the challenge of improving upon the
regional to local spatial scale subtropical summer precipitation estimates obtained from large-scale reanalyses.

The southeastern United States (Florida, Georgia, Alabama, South Carolina, and North Carolina), which has a subtropical summer climate, is a particularly challenging region to obtain measurably skillful summer precipitation forecasts (Kunkel et al. 2002; Lim et al. 2009). Roads et al. (2003) found that regional climate simulations and forecasts depict the precipitation intensity better for the western USA. Since a large proportion of the summer rainfall in the southeastern USA is convective in nature, the use of very fine-scale spatial resolution is essential for rainfall simulation in this region (Mullen and Buizza 2001). Therefore, this study utilizes a high resolution (20 km) RCM to simulate subtropical summer precipitation for this region given realistic large-scale fields of temperature, humidity, winds, and surface pressure. This region frequently faces small, localized areas of severe thunderstorms associated with extremely heavy rainfall during summer. It is also a region noted for some of the largest agricultural areas in the USA, producing a variety of crops and fruits. Improvements in the simulation of summer precipitation, including such characteristics as monthly to seasonal means, interannual variation, and frequency of heavy rainfall events, could potentially have a wide range of applications in agriculture (Robertson et al. 2007; Baigorria et al. 2008), water resource management, and decision making.

The regional model used in this study is the National Center for Environmental Prediction/Environmental Climate Prediction Center (NCEP/ECPC) regional spectral model (RSM) (Juang and Kanamitsu 1994). It is used to downscale large-scale analysis fields with the goal of producing improved simulations of summer precipitation on subseasonal to interannual time scales. The RSM has been well tested by the atmospheric science community and is widely used for regional climate simulation studies (Mearns et al. 2009). It has the very attractive feature of a scale selective bias correction that allows a downscaling ratio >1:10 (Juang and Hong 2001; Kanamitsu and Kanamaru 2007). Kanamitsu and Kanamaru (2007) have downscaled the NCEP/NCAR Reanalysis I (Kalnay et al. 1996) to a target resolution of 10 km over California with great success. In this study, the large-scale lateral forcing to the RSM is provided by the NCEP/DOE reanalysis II (R2) (Kanamitsu et al. 2002). R2 is dynamically downscaled with RSM to a resolution of 20 km, which corresponds to at least one or two grid points per county. For the assessment of the RSM performance, we focus on the model’s ability to add smaller scale information that is not resolved by the global reanalysis. We will use the observed high-resolution precipitation obtained from the National Weather Service (NWS) Cooperative Observer Program (COOP) for validation of the small-scale precipitation structure produced by the RSM. This downscaling of global reanalysis also can be understood as regional data assimilation without using observation, as discussed in von Storch et al. (2000). We note that the horizontal resolution in our study is finer than that of the North American Regional Reanalysis (NARR) (Mesinger et al. 2006).

The main goal of this study is to quantify the improvements in the spatial and temporal structure of subtropical summer precipitation obtained by using a 20 km resolution RCM (RSM) to downscale the large-scale fields obtained from the NCEP/DOE reanalysis II. The results should also provide useful information about the potential for improved predictive skill for the southeastern USA region, where the seasonal climate is closely linked with many natural systems and the human environment, including agriculture, forestry, water management, vegetation, tourism, and urban development.

The paper is organized as follows. The RSM, the experimental design with model configuration, and the observational data used for validation are described in Sect. 2. Section 3 describes the dynamically downscaled summer precipitation and its validation. This is followed in Sect. 4 by the discussion and concluding remarks.

2 Regional climate model and data

2.1 RSM

The RSM used in this study was first developed at the NCEP (Juang and Kanamitsu 1994), and subsequently maintained and further improved at the Scripps Institution of Oceanography/Experimental Climate Prediction Center (Kanamitsu et al. 2005). It is a spectral model with a physics package similar to the NCEP Global Forecast System (GFS) (Han and Pan 2006), including a suite of optional mesoscale cloud water schemes (e.g., Slingo 1987; Tiedtke 1993; Zhao and Carr 1997; Hong et al. 2004). The model can easily be nested inside a global model for daily weather and seasonal climate forecasts. Since the RSM is a spectral model, spectral decomposition is applied to the difference between the full field and the time-evolving background global analysis field (Juang and Kanamitsu 1994; Juang et al. 1997). The model is based on the primitive equation system under a hydrostatic approximation. We attempt the downscaling to 20 km spatial scale in order to allow for reasonable recognition of small-scale rainfall. We expect, based on the work of Juang and Hong (2001), that the large downscaling ratio (2.5° lon.–lat. → 20 km) will work reliably with spectral nesting.

The main characteristics of the model used in this study are summarized in Table 1. We tested the sensitivity of the
model responses to changes in model physics and found that the physics listed in Table 1 provided a reasonable seasonal pattern of summer precipitation (e.g., bias and general spatial patterns) for the southeastern USA.

### 2.2 Model experimental design

The NCEP/DOE Reanalysis (R2) (Kanamitsu et al. 2002) is used for the initial and boundary forcing to the RSM. The R2 dataset is six hourly, with a horizontal resolution of 2.5\(^\circ\) latitude and longitude and it provides the RSM with atmospheric forcing at 28 vertical pressure-sigma levels. The global R2 fields are downscaled with the RSM for a period of 16 years, from 1990 to 2005. RSM is initialized on 25 May of each year and integrated through the end of August. The oceanic SST boundary conditions are taken from the European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis (Fiorino 2004) at a resolution of 1\(^\circ\) of latitude and longitude. The SST data were interpolated to daily values using a mean conserving interpolation scheme (Taylor et al. 2000; Kanamitsu and Kanamaru 2007).

The geographical domain for the reanalysis downscaling covers the states of Florida, Georgia, Alabama, South Carolina, and North Carolina, as shown in Fig. 1. The model domain for simulation is larger than the domain shown in Fig. 1 by approximately 20 grid points along each longitudinal and latitudinal direction to accommodate the lateral boundary nudging zone (Kanamitsu and Kanamaru 2007). The large scale R2 field is downscaled by the RSM to a 20-km-resolution regional grid.

### 2.3 Interpolation of the R2 data via objective analysis

The Interpolation of the large-scale R2 precipitation fields is done to allow for comparison with the downscaled precipitation derived from the RSM. For the objective analysis (OA) of the R2 precipitation, we applied the Cressman OA scheme (Cressman 1959), which makes successive corrections to an initial guess, with increasingly smaller radii of influence. At each step, a correction factor based on a distance-weighted formula was applied to errors in order to minimize them. The error is defined as the difference between the value at the R2 grid point and the interpolated value at the fine-scale grid point. The final product is the objectively analyzed R2 precipitation on the 20 km RSM grid.

### 2.4 Observations

The source for the observed precipitation used for validation is the NWS COOP. The COOP has more than 100 years of observational data that help define the climate and long-term climate change over the USA. COOP weather stations provide a densely distributed record of daily weather observations over the entire USA (http://www.nws.noaa.gov/climate/). The gridded dataset for the southeastern USA region was provided by the Florida Climate Center (http://www.coaps.fsu.edu/climate_center/). The data period in this study covers the period of 1990–2005 with a daily time interval.
A map with the locations of all currently active COOP stations in the continental USA is available at ftp://ftp.ncdc.noaa.gov/pub/data/inventories/COOP-MAP.GIF. These densely distributed station data are converted to 20 km × 20 km grids using the Cressman OA scheme (Cressman 1959). As a result, each county is on average represented by at least one grid point in Georgia and North and South Carolina, where the counties’ areas are relatively small, compared to those in other states. Florida and Alabama have one or two grid points in most counties. The resulting 20 km gridded COOP dataset has 1,821 grid points, covering Florida, Georgia, Alabama, South Carolina, and North Carolina. The gridded values were compared with those from the station data to ensure the consistency of the fields.

3 Results

3.1 Monthly mean precipitation and bias

The daily precipitation derived from RSM is summed for each month (June, July, and August) to compare with observations. The monthly rainfall is then averaged over 16 years (1990–2005). The rainfall patterns produced by the RSM (middle row) and the observations (bottom row) are shown in Fig. 2, along with the coarse-scale precipitation obtained from R2 (top row). Spatial patterns from the left column to the right represent the accumulated monthly rainfall for June, July, and August. Overall, the results indicate that both R2 and RSM tend to overestimate the observed rainfall. This is particularly true for R2, which also exhibits a very limited ability to produce any detailed regional-scale rainfall features. In addition, the rainfall maximum produced by R2 is displaced from the observed maximum over central and southern Florida. R2 produces more precipitation over Georgia than over Florida in June (Fig. 2a) and July (Fig. 2b). We note that the precipitation obtained directly from the NCEP reanalysis and associated hindcasts both tend to overestimate the summer precipitation over the southeastern USA and exhibit a rainfall maximum over Georgia (Lim et al. 2009). Liang et al. (2004) also noted that R2 produces substantially heavier summer rainfall over the southeastern USA and Mexico (by >100 mm/month).

We see that the downscaling appears to produce realistic regional scale precipitation and improves upon the unrealistic spatial pattern seen from R2. For instance, it is clear from a comparison between RSM and R2 that the rainfall patterns derived from RSM (Fig. 2d–f) are closer to the observation (Fig. 2g–i), although they overestimate the precipitation over Georgia and South Carolina and underestimate the precipitation over western Alabama (Fig. 2d–f) (Fig. 3). The observed rainfall maximum over central and southern Florida is successfully simulated, and the increased precipitation along the coastal areas (Fig. 2g–i) relative to inland areas is also reasonably well reproduced by the RSM (Fig. 2d–f).

Previous studies suggest that many of the uncertainties/errors in regional model downscaled results are linked to the biases in the large-scale forcing data (Pan et al. 2001; Wu et al. 2005). Such bias in the large-scale atmospheric fields could effectively be transferred to the regional model and thereby lead to similar bias in the downscaled precipitation. Another possible reason for the bias in regional model is the formulation deficiency of the high-resolution rainfall derivation. This problem could be solved by applying more appropriate physical parameterizations for the southeastern USA in the regional model. In this study, the downscaled precipitation produced by RSM reveals substantially reduced wet biases over Alabama, Georgia, South Carolina and North Carolina. Figure 3 shows the reduction in the wet bias in the downscaled RSM precipitation (bottom row) compared with that from the reanalysis (top row). Note that the reanalysis precipitation in Fig. 3 is the R2 precipitation objectively analyzed to the 20 km grid. Critical discussion of the reasons for reduction in bias is not the main scope of this study. However, we suggest, based on the partial agreement with Roads et al. (2003) producing the reliable annual mean precipitation using the RSM, that the combination of relaxed Arakawa-Schubert (RAS) convective scheme with the RSM in the present study may play a positive role in producing the reasonable precipitation patterns from the large-scale fields for the southeastern USA. We also found that the ECMWF SST is more reliable for reducing large-scale wet bias than Atmospheric Model Intercomparison Project (AMIP) II SST as an oceanic boundary condition. The monthly mean AMIP II SST is generally higher than ECMWF SST, resulting in larger downscaled precipitation amount when AMIP II SST is applied in the RSM (figure not shown).

Figure 4 compares the RSM (red), the R2 (blue) and the observed (black) monthly rainfall variation for the summer months of the entire 16-year period. Here, the accumulated monthly rainfall has been area-averaged over the study domain (FL, GA, AL, SC, and NC). It is clear that R2 overestimates the observed rainfall, confirming the wet biases described previously. The interannual rainfall variation also appears to be poorly reproduced in R2 (blue) relative to RSM (red). The precipitation simulated by RSM, on the other hand, exhibits good agreement and reduced bias with the observed rainfall variation throughout the period. The mean bias of the RSM and R2 precipitation is, respectively, +26 mm (RSM) and +102 mm (R2). The root mean square (RMS) of the random part of the error is also calculated (Table 2). It shows that the
monthly precipitation error in R2 is substantially reduced in RSM (RMS (RSM) = 25 mm, RMS (R2) = 49 mm). RMS of the random error indicates the average skill in depicting the overall variability (Roads et al. 2003). From these two different error estimates, we find that both bias and random error are significantly reduced by the RSM. In addition, the standard deviation of the R2 monthly precipitation is one and a half times as large as the observed standard deviation ratio ((RSM/Obs.) = 1.37, (R2/Obs.) = 1.54). Correlations of RSM and R2 monthly precipitation with observation are, respectively, 0.89 and 0.52, indicating that the downscaling by the RSM better...
captures the observed interannual variation. Interannual variation will be discussed in more detail for individual grid points in Sect. 3.3.

3.2 Seasonal anomaly distributions during dry/wet summer years

3.2.1 Wet summer years

In order to investigate the regional model’s ability for properly simulating wet/dry summers (Jenkins 1997), we investigate the anomalous rainfall patterns associated with the four wettest and the four driest years. From the results shown in Fig. 4 we find that the wettest summers in the time period are 1992, 1994, 2003, and 2005, while the driest summers are 1990, 1993, 1998, and 2000.

The seasonal anomaly distributions for the wet summers (Fig. 5) indicate that, overall, both R2 and RSM reproduce reasonably well the observed above-average precipitation for all four wet summers (Fig. 5a–h). A detailed comparison between R2 and RSM indicates, however, that the RSM produces more accurate anomalous precipitation distributions at regional scales. The distributions are in good agreement with those of the observations: this is quantified by spatial correlations for the 4 years (Table 3). The underestimated positive R2 anomalies spread across most of the regions are improved in the RSM results. In particular, the large magnitudes of the observed positive anomalies over eastern Alabama, Georgia, and northern South Carolina in 1994 (Fig. 5j), central Alabama and Georgia in 2003 (Fig. 5k), and southern Alabama and the border of Georgia and Alabama in 2005 (Fig. 5l) are better simulated in RSM (Fig. 5f–h). However, the RSM does not always produce the appropriate sign of very localized observed precipitation anomalies. For instance, only positive anomalies are found over Alabama, Georgia, and Florida Panhandle in 1992 and 1994 in both R2 and RSM, while in the observations there are pockets of negative anomalies present within the predominantly positive-anomaly field (Fig. 5a, b, e, f, i, j). This is consistent with

---

Fig. 3 Biases of the climatological monthly rainfall for June (left column), July (middle column), and August (right column). The top panel represents the biases of the R2 precipitation whereas the bottom panel the biases of the RSM. For calculation of the biases of the R2 precipitation, large-scale R2 precipitation has been objectively analyzed to the spatial scale of 20 km. The scale is denoted by color bar attached on the right side.
previous studies of the limitations of regional downscaling of large-scale reanalysis (Castro et al. 2005; Diaconescu et al. 2007).

Table 3 represents the spatial correlations of the precipitation anomaly patterns shown in Fig. 5. The R2 precipitation values with 20 km resolution were constructed before calculating the correlations. They were obtained by assuming that the R2 values are the same at all 20-km-scale grid points within the R2 precipitation grid. Please note that R2 precipitation here was not interpolated to the RSM grid in order to avoid the smoothing effect that might result in artificial increase in spatial correlation (Rossa et al. 2008).

The correlation values of the RSM and R2 with the observation are >0.5 in every case. Particularly, the correlations in three cases (1992, 1994, and 2003) demonstrate the ability of the RSM to simulate the detailed wet summer precipitation patterns from the large-scale boundary conditions. The spatial correlations between the RSM results and observations are greater than those obtained with R2, by ranging from 0.12 (2003) to 0.18 (1992). However, correlation of the RSM for the case of 2005 is lower than that of the R2. It appears that the cause for this is the misrepresentation of the smaller scale precipitation over South and North Carolina, and western Alabama by the RSM.

3.2.2 Dry summer years

We next examine the dry summers. As seen in Fig. 6, the RSM reasonably simulates the negative rainfall anomalies for these four dry summers. Note that the southeastern USA region experienced a dry summer in 1993 while the Midwest region had a severe flood (Fennessy and Shukla 2000). An examination of the detailed features in Fig. 6 shows that the RSM successfully simulates the dry years. Based on the monthly rainfall amount shown in Sect. 3.1 and the anomalous patterns shown in this section, it is clear that R2 precipitation overestimates the rainfall amount (consistent with previous studies, e.g., Liang et al. 2004) and at the same time, in many cases, underestimates the magnitude of anomalies. Those problematic features in R2 are improved by downscaling with the RSM. For example, the underestimation of the negative precipitation anomalies in R2 over Georgia, eastern Alabama, and South Carolina in 1990 (Fig. 6a) and over Georgia and North and South Carolina in 1993 (Fig. 6b) is ameliorated by the RSM downscaling (Fig. 6e, f). Observed positive anomalies scattered over Florida peninsula in 1990 are successfully produced at local spatial scales in RSM (Fig. 6e, i), while R2 shows a negative anomaly over the region (Fig. 6a). Relatively weaker negative anomalies or even slightly positive anomalies over the western Alabama and negative anomalies over Florida Panhandle in 1993 (Fig. 6j) are better simulated by RSM (Fig. 6f). The incorrect positive R2 anomaly over the southern tip of Florida in 2000 is instead a more correct negative anomaly in the downscaled results except for the Ft. Myers area (Fig. 6d, h, l).

Spatial correlations are calculated for these dry summer cases, as described in the previous section for the wet summer cases. Table 3 shows that both RSM and R2 exhibit reasonable correlation values with the exception of the low correlation of R2 in 1993. Correlations are remarkably increased by the RSM in 1993 (0.43 → 0.71) and in 1990 (0.76 → 0.88). However, correlation values of the RSM and R2 with the observation are nearly identical in 1998 and 2000, indicating that the fine-scale anomalous features resolved at the exact locations are not overwhelming against the falsely resolved fine-scale features for these 2-year cases.

3.3 Seasonal to interannual variation

The seasonal to interannual variation of precipitation is investigated in more detail by looking at the precipitation
at individual local grid points for every summer month during the 16 years. Twelve grid points are selected representing four cities in Florida, and two cities in each of the remaining four states. The grid points are selected to provide an approximately evenly distributed representation of the southeastern USA region (Fig. 1). We plot the accumulated monthly precipitation for these grid points for the summers of the entire 16-year period (Fig. 7). The comparison between RSM and R2 results reveals that the observed precipitation (black) is better reproduced by the RSM (red) than by R2 (blue). While R2 substantially overestimates the observed precipitation, the RSM monthly precipitation values are closer in amplitude to the observations. RMS errors and correlations for the time series in Fig. 7 are provided in Table 4, quantifying the improvement obtained by downscaling with the RSM.

RMS and correlation maps for the monthly summer precipitation are shown in Fig. 8. We again remind the

![Figure 5](image1)  
**Fig. 5** Monthly rainfall (mm/month) anomaly distribution for the selected wet summer (JJA) years. Wet summer years picked from Fig. 4 are 1992, 1994, 2003, and 2005. Figures from the top row represent the rainfall distribution obtained from the R2 (top), the RSM (middle), and observation (bottom). The scale is denoted by color bar attached on the right side.

![Table 3](image2)  
**Table 3** Spatial correlations of R2 and RSM seasonal precipitation anomalies with the observations for individual wet summer years (92, 94, 03, and 05) and dry summer years (90, 93, 98, and 00)

<table>
<thead>
<tr>
<th>Wet years</th>
<th>1992</th>
<th>1994</th>
<th>2003</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSM versus observation</td>
<td>0.68</td>
<td>0.75</td>
<td>0.78</td>
<td>0.64</td>
</tr>
<tr>
<td>R2 versus observation</td>
<td>0.50</td>
<td>0.61</td>
<td>0.66</td>
<td>0.74</td>
</tr>
<tr>
<td>Dry years</td>
<td>1990</td>
<td>1993</td>
<td>1998</td>
<td>2000</td>
</tr>
<tr>
<td>RSM versus observation</td>
<td>0.88</td>
<td>0.71</td>
<td>0.78</td>
<td>0.80</td>
</tr>
<tr>
<td>R2 versus observation</td>
<td>0.76</td>
<td>0.43</td>
<td>0.79</td>
<td>0.78</td>
</tr>
</tbody>
</table>

Note that the large-scale R2 precipitation anomalies were objectively analyzed to 20 km grids before calculating correlations.
reader that the R2 precipitation has been objectively analyzed to a spatial scale of 20 km (Cressman 1959) before computing RMS and correlation. Figure 8 clearly demonstrates the improvement in the simulation of monthly summer precipitation by the downscaling. A substantial reduction in RMS is achieved from the RSM simulation (Fig. 8a–c). The RMS for the random error of the RSM precipitation ranges from $\sim 1.5$ to $\sim 3.0$ mm/day, while most of the R2 precipitation shows a range from $\sim 2$ to $\sim 3.5$ mm/day. RMS values are up to 4 mm/day in southwestern Georgia and Florida Panhandle. The comparison of the RMS between RSM and R2 reveals that the RMS is remarkably reduced in the RSM simulation in over 93% (1,697/1,821) of the total grid points (blue colored grid points in Fig. 8c). A few exceptions (yellow colored areas) are found primarily in North Carolina and southern tip of Florida, where the RMS of RSM precipitation is comparable to or greater than that of R2 precipitation (Fig. 8c).

The correlations with monthly precipitation observations are also improved for the RSM simulation (compared with R2) at almost all grid points (Fig. 8d–f). This improvement is encouraging because previous regional climate simulations forced by large-scale model or reanalysis had difficulty in achieving increased correlation for the southeastern USA. Liang et al. (2004) documented that a change in convective schemes (Grell and Kain-Fritsch scheme) can result in switching the sign of the model precipitation bias, suggesting the extreme sensitivity of the precipitation to convective scheme. The correlation values of the RSM monthly summer precipitation in this study exceed the statistically significant threshold value.
(0.28) at 95% confidence level over almost all grid points (99%) (1,811/1,821), except a few grid points [e.g., southwest of Jacksonville (FL)] (Fig. 8e). In R2, on the other hand, more than half of all grid points in the central part of the model domain show correlations below 0.4 (Fig. 8d). The correlation increase in RSM relative to R2 is more pronounced in the central part of the domain (Fig. 8f). An increase in correlation is achieved by the RSM simulation at 92% (1,675/1,821) of all grid points (Fig. 8d–f). The area-averaged correlation of the RSM and R2 monthly summer precipitation with observations over the domain is, respectively, 0.55 and 0.38.

The dominant principal modes are extracted from the monthly summer RSM precipitation and the interpolated monthly summer R2 precipitation to identify the spatial features of the dominant modes and their seasonal to
interannual variation using an empirical orthogonal function (EOF) decomposition (following Zhu and Liang 2007). The spatial patterns of each mode and the corresponding PC time series are compared with observations. Figure 9 shows the first two EOFs from R2 (top), RSM (second row), and observations (third row). The corresponding PC time series are compared with observations. The first mode extracted from R2, RSM, and observations (bottom panel). The first mode extracted from R2, RSM, and observations represents an almost identical signal as the three PC time series (Fig. 9d). The eigenvector distribution from the observations depicts a larger variance over central Georgia, southern Alabama, and South Carolina than over other regions (Fig. 9c).

The second mode (Fig. 9e–h) is characterized by an east–west dipole in the observations (Fig. 9g). This pattern appears to explain the difference in precipitation characteristics between inland and coastal regions that presumably is associated with regional scale features including land sea breeze effects. Mearns et al. (2003) found that the eastward gradient of summer precipitation [i.e., more (less) over the coastal (continental) regions] is observed on regional scales over the southeastern USA (Fig. 3 in Mearns et al. 2003). This feature is not reflected well in the coarse-scale fields (Mearns et al. 2003; Castro et al. 2007). The RSM does a good job in capturing this observed feature (Fig. 9f, g). R2, on the other hand, produces a rather different pattern that shows a northwest to southeast gradient (Fig. 9e). This pattern looks similar to the precipitation zone influenced by the large-scale low-level flow along the western edge of the Bermuda high (Fig. 2 in Higgins et al. 1997). We speculate that the downscaled precipitation over the eastern domain may be reflecting the effect of the interannual variability in the westward extent of the Bermuda high on the coastal sea breeze and inland convection. Further analysis will be necessary for drawing firmer conclusions associated with this spatial feature.

3.4 Synoptic to daily scale extreme events

(heavy rainfall)

3.4.1 Spatial patterns for heavy rainfall cases

We now turn to the rainfall features present on sub-monthly time scales. Specifically, we examine the simulated spatial rainfall distributions in the case of heavy rainfall. The heavy rainfall cases are selected from the observed 16-year record of daily rainfall. We compare the observed spatial patterns of rainfall during these events with the spatial patterns derived from both R2 and RSM. We then investigate the number of heavy rainfall events every month for the entire 16-year record.

Figure 10 shows the area-averaged rainfall time series with a daily time interval. We again see the significant overestimation of precipitation by R2 (blue line). The rainfall variation derived from the RSM (red line), however, shows a good agreement with the observed precipitation (black line), in terms of both amplitude and timing. As shown in Table 4, the correlation of the area-averaged daily summer precipitation (10-day running averaged) from RSM and from R2 with observations is, respectively, 0.86 and 0.51. Calculations of the corresponding RMS error yield 2.07 and 3.75 mm/day, respectively.

Four observed heavy rainfall cases are picked from this time series (Fig. 10) and the corresponding spatial rainfall distributions are plotted in Fig. 11. Figure 11 shows the rainfall anomaly patterns derived from R2 (top), RSM (middle), and observations (bottom) for the selected heavy rainfall cases. The second (July 04–07, 1994) and the fourth (July 10–11, 2005) cases are related to tropical cyclones, Alberto and Dennis, respectively. Anomalies are defined as deviations from the 16-year total summer mean.

Although coarsely resolved, the distribution of positive/negative anomalies seen from R2 is not very different from the observed anomalies in terms of large-scale features. The positive anomalies, with the exception of western Alabama in 1992 (Fig. 11a, i), the positive anomaly over Georgia and southeastern Alabama with a negative anomaly over North Carolina and southern Florida in 1994 (Fig. 11b, j), and the positive anomaly over Alabama, Georgia, and Florida with a negative anomaly over South and North Carolina in 2005 (Fig. 11d, i) are all examples of R2 roughly matching the geographical extent of the observed rainfall anomaly (Fig. 11a–d, i–l). Due to its coarse resolution, however, R2 lacks the ability to produce detailed features on finer spatial scale. In addition, the

<table>
<thead>
<tr>
<th>City</th>
<th>RMS Error RSM</th>
<th>RMS Error R2</th>
<th>Correlation RSM</th>
<th>Correlation R2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tallahassee</td>
<td>81</td>
<td>94</td>
<td>0.50</td>
<td>0.34</td>
</tr>
<tr>
<td>Jacksonville</td>
<td>65</td>
<td>72</td>
<td>0.52</td>
<td>0.45</td>
</tr>
<tr>
<td>Orlando</td>
<td>65</td>
<td>91</td>
<td>0.64</td>
<td>0.48</td>
</tr>
<tr>
<td>Miami</td>
<td>70</td>
<td>88</td>
<td>0.57</td>
<td>0.43</td>
</tr>
<tr>
<td>Atlanta</td>
<td>73</td>
<td>84</td>
<td>0.49</td>
<td>0.33</td>
</tr>
<tr>
<td>Tifton</td>
<td>65</td>
<td>78</td>
<td>0.61</td>
<td>0.32</td>
</tr>
<tr>
<td>Birmingham</td>
<td>54</td>
<td>84</td>
<td>0.72</td>
<td>0.40</td>
</tr>
<tr>
<td>Montgomery</td>
<td>61</td>
<td>95</td>
<td>0.71</td>
<td>0.33</td>
</tr>
<tr>
<td>Columbia</td>
<td>56</td>
<td>69</td>
<td>0.64</td>
<td>0.41</td>
</tr>
<tr>
<td>Charleston</td>
<td>63</td>
<td>71</td>
<td>0.58</td>
<td>0.49</td>
</tr>
<tr>
<td>Charlotte</td>
<td>73</td>
<td>77</td>
<td>0.40</td>
<td>0.33</td>
</tr>
<tr>
<td>Raleigh</td>
<td>57</td>
<td>76</td>
<td>0.60</td>
<td>0.48</td>
</tr>
</tbody>
</table>
magnitude of the R2 precipitation anomalies tends to be smaller than that of the observed anomalies.

Using the large-scale information from R2, however, RSM is successful in producing a reasonable regional structure of precipitation for the heavy rainfall events (Fig. 11e–h). For example, the region of maximum rainfall along the Atlantic coast extending from the northeastern Florida, through the eastern part of Georgia, and up to South and North Carolina in 1992 (Fig. 11i), the wet anomaly band extending from the western Florida Panhandle to the eastern Georgia in 1994 (Fig. 11j), the precipitation maxima over the western edge of Florida Panhandle, central Georgia, South Carolina, and the western North Carolina in 2003 (Fig. 11k), and the precipitation maximum over the western Florida Panhandle, western Georgia and Alabama in 2005 (Fig. 11l) are reasonably well simulated by the RSM without losing the magnitude of the precipitation anomalies (Fig. 11e–h). The negative anomaly over southern Florida peninsula in 2003 and 2005 cases is also better resolved by the RSM (Fig. 11g, h, k, l).

However, the RSM does not show improved spatial patterns in all cases. The opposite sign of the anomaly in the RSM compared with the observations over the northwestern Alabama in 1994 is a good example of the RSM not

---

**Fig. 8** Left column: geographical distribution of the root mean square error (RMSE) of the interpolated R2 precipitation (mm/day) (top panel) and the downscaled RSM precipitation (middle panel). The difference subtracting RMSE of R2 from that of RSM is plotted on the bottom panel. The scale is denoted by color bar on the left side (mm/day). Right column: same as the left column but for correlation of monthly precipitation. The scale is denoted by color bar on the right side. Note that 0.28 is the statistically significant threshold value at 95% confidence.
being able to alter the positive sign of the R2 precipitation (Fig. 11b, f, j).

3.4.2 Correlation and categorical classification skill for the frequency of heavy rainfall events

The heavy rainfall days are counted each summer month over 16 years at individual 20-km-resolution grid points. A heavy rainfall day is defined as a rainy day with the rainfall amount exceeding a certain value [e.g., 1 standard deviation (Figs. 12, 13), or 2 standard deviations (Fig. 14) from the total JJA mean of the respective data]. In order to evaluate the ability of the RSM and R2 for producing these daily to synoptic scale features, we investigate the month-to-month variation of the frequency of heavy rainfall days over the 16 years (48 months). Figure 12 shows the month-to-month variation of the number of heavy rainfall days obtained from observations (black), the RSM (red), and interpolated R2 values (blue) for the selected grid points. The number of heavy rainfall events in any summer month
ranges between 0 and 12. At a glance, Fig. 12 indicates that both RSM and R2 reasonably reproduce the observed fluctuations. If the observed standard deviation and JJA mean are used for the threshold value for defining heavy rainfall events as opposed to the R2 mean and standard deviation, however, the R2 time series will significantly overestimate the observational number of heavy rainfall days because of the R2 precipitation’s wet biases discussed in Sect. 3.1. Table 6 shows the RMS error and correlation for the time series in Fig. 12, confirming the improved simulation of short-term heavy rainfall events by the RSM.

For a quantitative assessment of the skill in resolving these short-term rainfall events by R2 and RSM, two skill measures are calculated at individual grid points—correlation and Heidke skill score (HSS) (Heidke 1926; Jolliffe and Stephenson 2003). The HSS is a commonly used categorical verification score which measures categorical matches between model output and observations (Barnston 1992). A three-category (above/near/below average) classification is considered for this HSS calculation. The threshold values for the three category classification are ±1 standard deviation from the mean of the monthly frequency of heavy rainfall days for 48 months (3 months × 16 years). Based on the HSS formula (Heidke 1926), positive and negative HSS values indicate, respectively, skill above and below that of random chance.

The left column of Fig. 13 shows the correlation of RSM and R2 with observations for the heavy rainfall days per month. The threshold value used here for defining heavy rainfall events is one standard deviation plus the total JJA mean precipitation. The results show a noticeable increase in correlation of the monthly frequency of heavy rainfall events produced by the RSM relative to R2. While the correlation values obtained from R2 are below 0.4 over a majority of grid points (Fig. 13a), correlations calculated from the RSM precipitation show values exceeding 0.4 over many grid points (Fig. 13b). The improvement of the correlation values shows wide variation among local stations. Several grid points reach or exceed 0.6. As illustrated by the difference map in Fig. 13c, 93% (1,700) of all grid points show an increase in correlation values. The correlation value averaged over the domain is, respectively, 0.48 (RSM) and 0.28 (R2).

Finally, we assess the categorical classification skill for the monthly frequency of heavy rainfall days in terms of HSS (right column of Fig. 13). The HSS calculation is based on a three-category classification (above/near/below average), as described previously. The distribution of the HSS values in Fig. 13d, e reveals the improved categorical classification of the frequency of heavy rainfall days every month by RSM. HSS values are positive at 95% (1,724) of all grid points. The result from R2 also shows mainly positive HSS values over many grid points but concurrently, negative HSSs are observed over 23% (419) of all grid points (Fig. 13d). A comparison indicates that HSS values from RSM are, in general, greater than those from R2, as shown in the difference map in Fig. 13f. 73% (1,327) of all grid points show higher HSS values for RSM. The domain averaged HSS is, respectively, 0.20 (RSM) and 0.10 (R2).

In order to confirm the robustness of this conclusion, calculation of these skill measures is repeated for a different threshold value for defining heavy rainfall events. Similar relative differences between RSM and R2 are found again when for a threshold value of two standard deviations (Fig. 14). Overall features again show that the correlation and HSS calculated from the RSM are once again positive over a majority of grid points [89% (1,627)] (Fig. 14e). HSSs values for R2 are positive over only 69% (1,254) of total grid points (Fig. 14d). The area averaged HSS yields 0.16 for the RSM (Fig. 14e) and 0.07 for R2 (Fig. 14d).
A reasonable simulation of heavy rainfall frequency is valuable because seasonal rainfall has a strong contribution from extreme events (Higgins et al. 2007). The main conclusion that the monthly frequency of heavy rainfall events is better simulated by RSM than by R2 is robust with a change in the threshold value for defining heavy rainfall events, indicating that downscaling by the RSM is indeed very beneficial for better prediction of extreme events at fine spatial scales.

4 Concluding remarks

This study investigates the fidelity of the downscaled summer precipitation over the southeastern USA using a RCM. This is a particularly challenging region for which to produce realistic precipitation since a majority of rainfall events are caused predominantly by local- to regional-scale convection. The summer precipitation derived from the NCEP/ECPC RSM (Juang and Kanamitsu 1994; Kanamitsu et al. 2005; Kanamitsu and Kanamaru 2007), driven by the six-hourly large-scale atmospheric variables from the NCEP/DOE Reanalysis II (R2), has been validated for the southeastern USA region covering Florida, Georgia, Alabama, South Carolina, and North Carolina, for the 16 summers from 1990 to 2005. The oceanic boundary condition was obtained from the ECMWF reanalysis SST (Fiorino 2004) with a daily time interval (Taylor et al. 2000). The precipitation was down-scaled to 20 km so that the spatial grid distribution is denser than county-level. The spatial patterns and skill values of the
precipitation data derived from the RSM and R2 (interpolated by OA to a 20 km grid) were validated and compared in the investigation.

The results presented here offer promise for the use of RCMs for the study of high-resolution (20 km) precipitation variability in the southeastern USA. The regionally downscaled reanalysis with the RSM produces more realistic distributions of summer precipitation than the interpolated original reanalysis (R2). Wet biases of the reanalysis precipitation documented in the study domain are reduced by the downscaling. The regional downscaling dramatically reduces severe wet biases (greater than \( \sim 100 \text{ mm/month} \)) of R2 over northern Alabama, Georgia, central South Carolina and North Carolina.

An assessment of the seasonal to interannual variation in the precipitation also demonstrates an improved simulation...
of the high-resolution subtropical summer precipitation by the downscaling. Higher correlations and smaller RMS error are found in the RSM monthly precipitation. Correlations of the R2 monthly precipitation with observations range from 0.1 to 0.7. The corresponding values from the RSM-downscaled reanalysis range from 0.3 to 0.8, which is statistically significant at the 95 confidence level over nearly all of 1,821 local grid points (1,811/1,821). This reflects an increase in correlation values over 92% (1,675) of all grid points. The domain-averaged correlation values of the RSM and R2 precipitation are, respectively, 0.55 and 0.38. The calculated RMS error values also support the capability of RSM for simulating summer precipitation over the southeastern USA. The RMS of the monthly precipitation error for the RSM precipitation ranges from 1.5 to 3 mm/day, while the range for R2 is of 2–3.5 mm/day with 93% (1,697) of all grid points reflecting a reduction of RMS error by the RSM.

Calculation of the correlations of monthly frequency of heavy rainfall days for 16 years (48 summer months) demonstrates that the regional downscaling exhibits better agreement with observations. When heavy rainfall days are defined as those with precipitation anomaly exceeding one standard deviation, 93% (1,700) of all grid points show
improved correlations in the downscaled precipitation. The domain-averaged correlation values for heavy rainfall event frequency are, respectively, 0.48 (RSM) and 0.28 (R2). We also assess the categorical classification skill of the monthly frequency of heavy rainfall days in terms of HSS. Comparison of HSS values under this three-category classification (above/near/below average, with a threshold of one standard deviation from mean) reveals an improved categorical classification of the frequency of heavy rainfall events after downscaling. For RSM, HSS values are predominantly positive over 95% (1,724) of all grid points. The result from the R2 shows positive HSS values over 77% (1,402) of all grid points. 73% (1,327) of all grid points show higher HSS for RSM than for R2. The domain-averaged HSS is, respectively, 0.20 (RSM) and 0.10 (R2).

The present study finds that, given reasonable large-scale forcing (R2), the RSM is capable of adding small-scale information to the large-scale precipitation field. This study also confirms that a horizontal downscaling ratio larger than 10:1 can be used in RSM for the southeastern USA (Juang and Hong 2001). It seems that the combination of model physics including RAS convective scheme in the RSM make a positive contribution to the improved high-resolution summer precipitation simulation for the
Table 6 Root mean square errors (mm) and correlations of the monthly frequency of heavy rainfall days produced by R2 and RSM, respectively, with the observation (Fig. 12) for the selected city areas

<table>
<thead>
<tr>
<th>City</th>
<th>RMS error</th>
<th>R2</th>
<th>Corr.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tallahassee</td>
<td>2.2</td>
<td>3.0</td>
<td>0.59</td>
</tr>
<tr>
<td>Jacksonville</td>
<td>2.0</td>
<td>2.7</td>
<td>0.46</td>
</tr>
<tr>
<td>Orlando</td>
<td>2.2</td>
<td>2.9</td>
<td>0.47</td>
</tr>
<tr>
<td>Miami</td>
<td>1.6</td>
<td>2.5</td>
<td>0.66</td>
</tr>
<tr>
<td>Atlanta</td>
<td>2.5</td>
<td>2.8</td>
<td>0.34</td>
</tr>
<tr>
<td>Tifton</td>
<td>2.1</td>
<td>2.8</td>
<td>0.56</td>
</tr>
<tr>
<td>Birmingham</td>
<td>2.2</td>
<td>2.6</td>
<td>0.64</td>
</tr>
<tr>
<td>Montgomery</td>
<td>2.1</td>
<td>2.7</td>
<td>0.73</td>
</tr>
<tr>
<td>Columbia</td>
<td>2.2</td>
<td>2.5</td>
<td>0.52</td>
</tr>
<tr>
<td>Charleston</td>
<td>1.8</td>
<td>2.3</td>
<td>0.57</td>
</tr>
<tr>
<td>Charlotte</td>
<td>2.2</td>
<td>3.0</td>
<td>0.45</td>
</tr>
<tr>
<td>Raleigh</td>
<td>1.8</td>
<td>2.2</td>
<td>0.58</td>
</tr>
</tbody>
</table>

southeastern USA. However, critical investigation is necessary to more accurately assess the contribution of the individual physics and techniques to the bias reduction and improved simulation of precipitation variability. It should be also noted that no increased skill would be gained by dynamical downscaling if synoptic to larger scale variability is not resolved in the larger model or reanalysis (Rockel et al. 2008).

References

simulations of heavy precipitation events and seasonal anomalies: a comparison. J Hydrometeorol 3:322–334