Diagnosis and Attribution of a Seasonal Precipitation Deficit in a U.S. Regional Climate Simulation

WILLIAM J. GUTOWSKI JR.

Department of Geological and Atmospheric Sciences, and Department of Agronomy, Iowa State University, Ames, Iowa

FRANCIS O. OTIENO

Department of Geological and Atmospheric Sciences, Iowa State University, Ames, Iowa

RAYMOND W. ARRITT

Department of Agronomy, Iowa State University, Ames, Iowa

EUGENE S. TAKLE

Department of Geological and Atmospheric Sciences, and Department of Agronomy, Iowa State University, Ames, Iowa

ZAITAO PAN

Department of Agronomy, Iowa State University, Ames, Iowa

(Manuscript received 2 October 2002, in final form 8 September 2003)

ABSTRACT

Precipitation from a 10-yr regional climate simulation is evaluated using three complementary analyses: self-organizing maps, bias scores, and arithmetic bias. Collectively, the three reveal a precipitation deficit in the south-central United States that emerges in September and lingers through February. Deficient precipitation for this region and time of year is also evident in other simulations, indicating a generic problem in climate simulation.

Analysis of terrestrial and atmospheric water balances shows that the 10-yr average precipitation error for the region results primarily from a deficit in horizontal water vapor convergence. However, the 10-yr average for fall only suggests that the primary contributor is a deficit in evapotranspiration. Evaluation of simulated temperature and soil moisture suggests the model has insufficient terrestrial water for evaporation during fall. Results for winter are mixed; errors in both evapotranspiration and lateral moisture convergence may contribute substantially to the precipitation deficit. The model reproduces well both the time-average and time-filtered large-scale circulation, implying that the moisture convergence error arises from an error in simulating mesoscale circulation.

1. Introduction

Regional climate models (RCMs) are among the tools developed to address concerns of climate change impact at regional scales. However, before RCMs can be used for these purposes, better understanding of their behavior in simulating current climate is needed to characterize their strengths and weakness and, ideally, improve their simulation. Such understanding is especially important for correct interpretation of results from climate projections under different scenarios and for estimating confidence bounds of climate projections. This task requires a variety of model assessments that can provide broad understanding of model behavior.

We use such a variety of assessments here to evaluate precipitation produced by a 10-yr RCM simulation of the United States. This simulation is one part of a sequence of contemporary and future climate simulations that have been cross compared to assess confidence in their climate changes (Pan et al. 2001b). Evaluation of this simulation on its own is also important, as it provides a basis for diagnosing specific areas of concern and directions for improvement.

Evaluations of extratropical simulation tend to focus on annual averages and extreme seasons (e.g., MacAvaney et al. 2001; Giorgi et al. 2001), implicitly assuming that these cover the major challenges for good simulation. However, a model must also simulate tran-

Corresponding author address: William J. Gutowski Jr., 3021 Agronomy Hall, Iowa State University, Ames, IA 50011.
E-mail: gutowski@iastate.edu

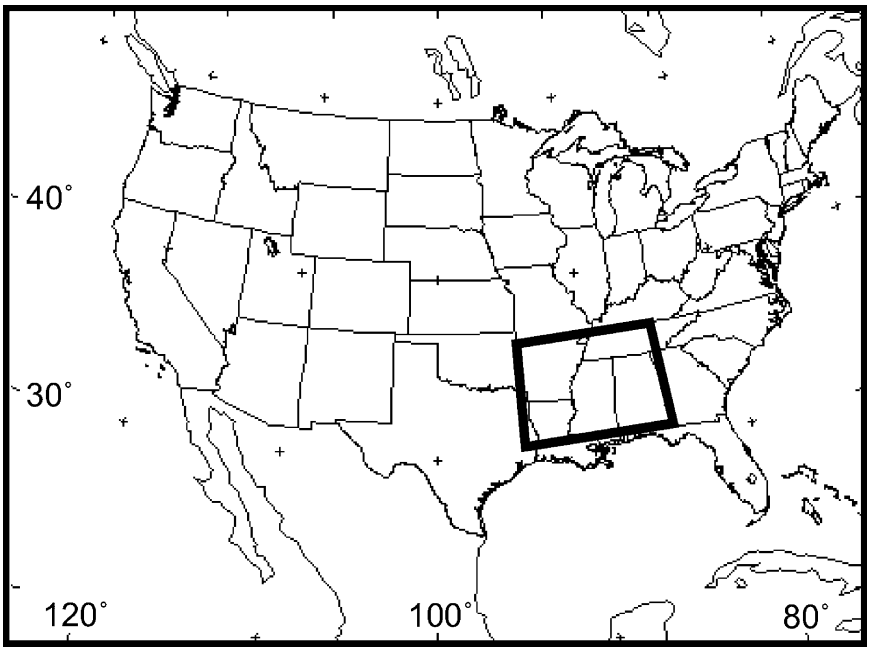


FIG. 1. Model domain. Heavy interior lines mark the MD box.

sition seasons well, especially for climate change impacts assessments. For example, transition seasons govern growing-season length and may have strong influence on flooding and water resources as periods of snowmelt. The analyses here will highlight a southern U.S. precipitation deficit that emerges during September–October–November (SON) and lingers through December–January–February (DJF). We will also present evidence of similar deficits in climate simulations by other models, so these results may apply to U.S. precipitation simulation in general.

Section 2 gives a description of the model and the 10-yr simulation it performed, along with the observational datasets used to evaluate the simulation. Section 3 presents the precipitation analyses, and section 4 diagnoses possible sources for the most prominent precipitation deficit emerging from the analyses. The paper concludes in section 5 with a discussion of the results and suggested causes for the deficit.

2. Model output and observations

a. Simulation

We analyze a 1979–88 simulation performed by the Second-Generation Regional Climate Model (RegCM2; Giorgi et al. 1993a,b). The model computed precipitation using a simplified version (Giorgi and Shields 1999) of the Hsie et al. (1984) explicit moisture scheme and the Grell (1993) convection parameterization. The model also used the Biosphere–Atmosphere Transfer Scheme (BATS) version 1e (Dickinson et al. 1993) land surface model and the Holtslag et al. (1990) nonlocal boundary layer turbulence parameterization. Radiative transfer used the Community Climate Model version 2 (CCM2) radiation package (Briegleb 1992). The model did not include a gravity wave drag parameterization. The simulation domain (Fig. 1) was the North American grid of the Project to Intercompare Regional Climate Simulations (Takle et al. 1999), with 101 × 75 grid points spaced 52 km apart in a Lambert conformal projection centered at 37.5°N, 100°W. The vertical grid used 14 sigma layers with finer resolution near the surface. Pan et al. (2001b) give further details of the simulation and discuss general features of the precipitation output and its change under greenhouse warming.

The simulation used output from the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis (Kalnay et al. 1996) for initial and lateral boundary conditions. The simulation started on 1 October 1978, in part to allow the model to build its own snow cover, but we discarded the first three simulated months to reduce the potential influence of spinup on the model's climatology. The reanalysis also provided evolving sea surface temperature (SST). However, in the poorly resolved Gulf of California and North American Great Lakes, the model supplemented the reanalysis data with Advanced Very High Resolution Radiometer (AVHRR) retrievals (Brown et al. 1993) and buoy observations from the Great Lakes Environmental Research Laboratory, respectively. The evolving water vapor field used for lateral boundary conditions has lower accuracy than other upper-air fields (Kalnay et al. 1996), which is potentially an important factor in our simulation. Accuracy of the

initial soil moisture also may be questionable because of climatological constraints placed on reanalysis soil moisture (Roads et al. 1999; Roads and Betts 2000), but simulated soil moisture (Pan et al. 2001a) shows no evidence of slow, multiyear spinup from a poor initial condition. We otherwise assume that the driving data replicate the actual atmosphere, so that all precipitation error is attributable to the model itself unless clearly linked to lateral boundary conditions or soil moisture.

b. Observations

1) VEMAP PHASE II PRECIPITATION

The Vegetation/Ecosystem Modeling and Analysis Project (VEMAP) has produced an 1895–1993 time series of monthly average precipitation on a 0.5° grid for the contiguous United States (Kittel et al. 1995, 1997, 2000; Schimel et al. 2000). Data processing used a local, moving-window kriging model (Haas 1990) to produce serially complete station records using monthly observations from the National Climatic Data Center's Historical Climatalogy Network (~1200 stations from 1895 to present), the Cooperative Observing Network (additional 6000–8000 stations from 1951–90), and more recent snowpack telemetry (SNOTEL) observations. These records were then merged using the Parameter-Elevation Regressions on Independent Slopes Model (PRISM; e.g., Daly et al. 1994) for spatial interpolation with topographic adjustment.

2) HIGGINS DAILY PRECIPITATION

Higgins et al. (1996) have produced gridded hourly precipitation using observations spanning 1963–93 from approximately 2500 stations, which they interpolated to a 2.0° (latitude) by 2.5° (longitude) grid using a modified Cressman (1959) scheme. Like the VEMAP dataset, it covers only the contiguous United States.

3) NCEP–NCAR REANALYSIS

The NCEP–NCAR reanalysis (NNR) is a product of data assimilation. We assume that, for the observation-rich United States, the NNR gives the best available rendition for observational, gridded upper-air fields. Our analyses use geopotential heights, which Kalnay et al. (1996) describe as strongly influenced by observations and therefore among the most reliable NNR fields. We also use NNR precipitable water, although with caution, as the NNR forecast model likely has strong influence on this field.

4) UNH/GRDC GRIDDED RUNOFF CLIMATOLOGY

The University of New Hampshire (UNH) and the Global Runoff Data Centre (GRDC) have produced a monthly, annual-cycle climatology of runoff on a global 0.5° grid (Fekete et al. 1999). The method combines observed river discharge with a climatology-driven water-balance model to compute composite runoff fields that are consistent with observed discharge volumes and the water-balance model's spatial distributions. Periods of record for the 15 contributing discharge sites in our southern U.S. focus region range from 12 to 92 yr and end in the mid-1980s. Median record length is 49 yr.

3. Precipitation error analysis

a. Methods

To gain a broad characterization of precipitation errors in the simulation, we employed three complementary diagnostic methods: self-organizing maps, bias scores, and arithmetic bias. We applied the methods to the contiguous United States and, for the bias scores and arithmetic bias, to a Mississippi Delta (MD) box (31°–37°N, 85°–95°W; Fig. 1) that emerged as a special focus in our analyses because of its error characteristics. To compare observed and simulated precipitation, we interpolated the model's monthly average precipitation to the VEMAP grid using a Cressman (1959) scheme.

Self-organizing maps (SOMs; Kohonen 1995) are two-dimensional arrays of maps that display characteristic behavior patterns of a field (e.g., Main 1997; Cavazos 1999, 2000; Michaelides et al. 2001; Hewitson and Crane 2002). The SOM array is a discrete representation of a continuous pattern space occupied by the field examined. Individual maps in the array represent nodes in a projection of this space onto a two-dimensional surface, with the two dimensions showing the two primary pattern transitions for the field examined. The input maps themselves determine the degree and types of pattern transition, hence the ''self-organizing'' nature of the resulting array. The SOM array is trained on a sequence of input maps through an artificial neural net technique. In contrast to eigenvector techniques such as empirical orthogonal functions, maps in the SOM array do not necessarily favor the largest scales in the input data, but rather the scales most relevant to the field for the domain and resolution examined.

The array size (number of nodes) is not predetermined, but the SOM's discretization of the pattern space provides guidance in this choice. Small arrays of maps (e.g., 2×4) may allow fairly rapid SOM generation but poor discretization of the pattern space spanned by the input field. Large arrays (e.g., 10×12) may offer finer but perhaps unnecessary discretization at the expense of more cumbersome computation. For the monthly precipitation examined here, a 4×6 array of maps appears to discriminate well the primary patterns of observed and simulated precipitation during our 10-yr study period.

SOM generation uses an iterative procedure that compares each member of a sequence of maps, such as 10 years of monthly precipitation, to an existing set and

nudges the closest map in the set toward the input map. The initial maps in the set consist of random numbers. A well-constructed SOM set is independent of initial iteration conditions, which we determined was the case for our results through tests using alternative initial conditions and sequences of map ingestion. Our metric for determining the closest, or best-fit, map is the Euclidian norm,

$$D^{ij} = \sqrt{\frac{\sum_{k=1}^{n} (P_k^i - P_k^j)^2}{n}}, \quad (1)$$

where P is precipitation at grid point k of n total points, i refers to one of the input precipitation maps, and j refers to one of the SOM maps. To ensure relatively smooth transitions between SOM maps, the procedure also nudges toward the input map all SOM maps in a neighborhood about the best-fit map, with the degree of nudging and neighborhood size shrinking with iteration. A well-constructed SOM set minimizes

$$D = \sqrt{\sum_{i=1}^{I} (D^{iJ_i})^2}, \quad (2)$$

where I is the total number of input maps, and J_i is the closest member of the SOM array to input map i. Our application combines 10 years of monthly precipitation maps from observations and model output to yield $I = 10 \times 12 \times 2 = 240$. We performed 20 000 iterations with this set of input maps to arrive at our final SOM array, a procedure that took roughly 30 CPU min on a DEC 3000 workstation.

With a trained SOM array, one can track the evolution of a field through pattern space by the sequence of SOM maps chosen to be closest to each member i of the input sequence. We constructed a climatological annual cycle of monthly precipitation for observations and for model output by extracting the frequency distribution of the closest, or best-fit, maps for each calendar month's 10 renditions and then computing the centroid of this distribution in the pattern space spanned by the SOM array. Connecting centroids for observations and model output separately then gave their respective annual cycles. Comparison of annual cycles reveals the degree of agreement between simulated and observed precipitation patterns.

The bias score (Wilks 1995) is the ratio

$$B = \frac{N_M}{N_O}, \quad (3)$$

where N_M is the number of event occurrences in the model, and N_O is the number of event occurrences in observations. Here, an event is monthly precipitation at a grid point exceeding a specified threshold. If $B = 1$, then the model exceeds the threshold exactly the same number of times as the observations, with larger (smaller) score meaning that the model produces more (fewer) events. Monthly precipitation for most of the United States averages 1–4 mm day^{-1}. Thresholds well above this range will have $N_O = 0$, and thus B will be indeterminate. Except for deserts, thresholds approaching zero for monthly precipitation will yield $B \to 1$, in essence simply confirming that both observed and simulated precipitation occurred some time during the month. Thus, bias scores will be most meaningful for thresholds between these extremes. We used monthly average precipitation rates of 1, 2, and 4 mm day^{-1} as thresholds.

Finally, the arithmetic bias (or simply, bias) is the simulation − observation difference. We present biases here for seasonal maps of the United States and for monthly precipitation differences averaged over the MD box.

b. Results

Prominent features in the SOM array (Fig. 2) are high precipitation over the northwestern and southeastern United States and low precipitation in the Southwest. Moving from left to right among the maps shows increasing southeastern U.S. precipitation, while moving up and down the array shows variation in the northwestern precipitation. The upper-row maps show typical summer climatological patterns: little precipitation in the Southwest, a strong west–east gradient in the middle of the country, and largest amounts along the southeastern coast. In contrast, the maps in the lower row show typical winter climatological patterns: large precipitation amounts in the Northwest and generally less precipitation compared to summer in the eastern two-thirds of the country, except in the south-central United States. The maps thus display a range of monthly precipitation patterns for the United States that correspond in part to the annual cycle.

The maps in Fig. 2 appear in a rectangular array for plotting convenience. However, this array does not show their true separation as measured by (1). A higher dimensional space is required, but an approximate projection onto a two-dimensional space known as a Sammon map (Sammon 1969; Kohonen 1995) is sufficient for many purposes, including ours. The dots in Fig. 3 show the Sammon map of the Fig. 2 array, where the dots are separated approximately by the differences between neighboring maps. In the Sammon map, the array retains an approximately rectangular shape. The largest separation of points occurs between the middle rows, indicating distinct warm–cold season differences in precipitation patterns. Point separation is also larger along the upper (summer) rows than along the lower (winter) rows, indicating greater spread of precipitation patterns in summer than in winter.

Figure 3 also shows the centroids of each month's best-fit precipitation maps in this array, computed using the method described in the previous section and projected to their Sammon-map positions. The solid (ob-

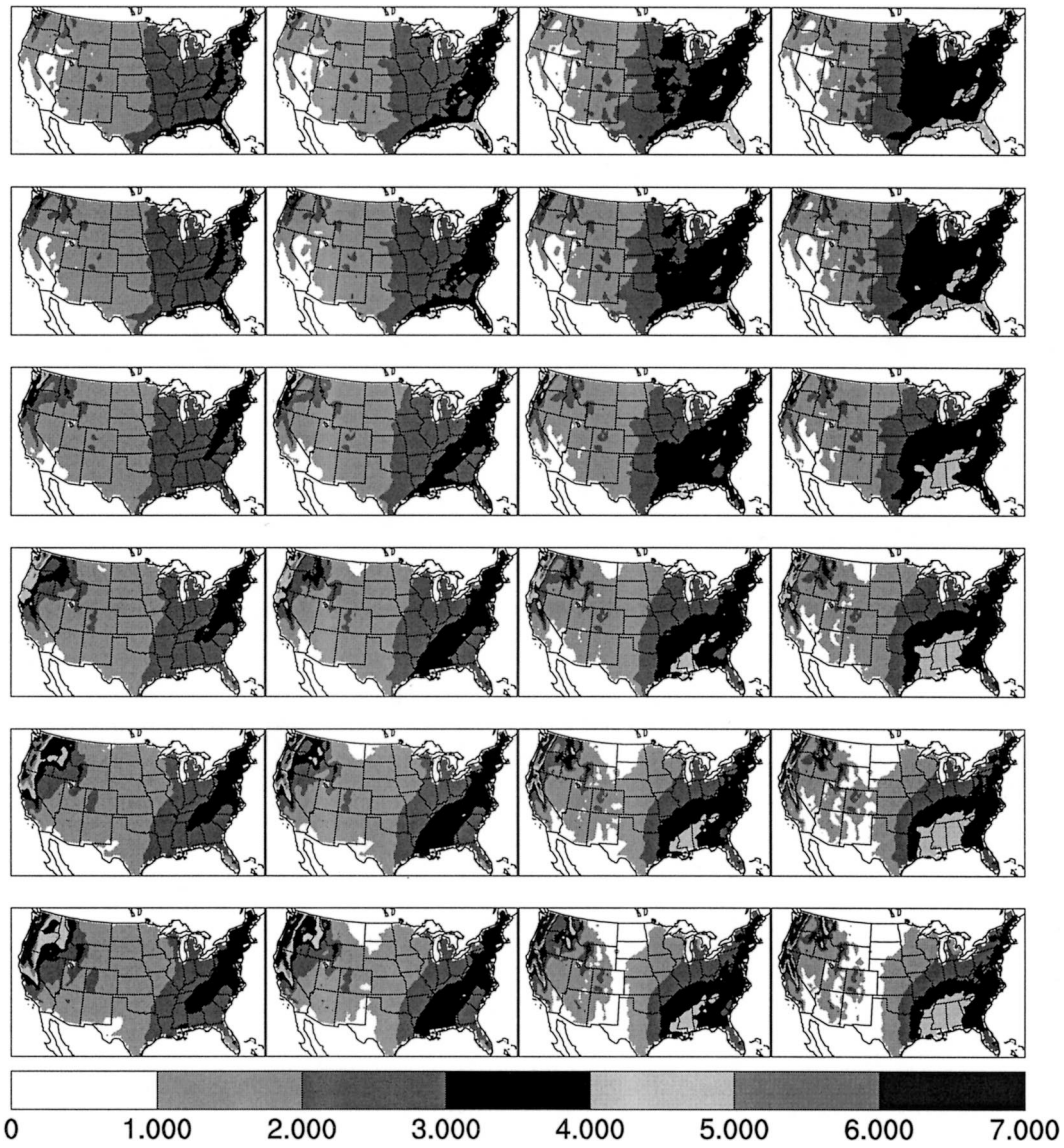


FIG. 2. SOM array of monthly precipitation. The two axes of the array give the two principal variations of precipitation in the pattern space it spans. See text for details. (Units are mm day^{-1}.)

servations) and dashed (simulation) lines connecting centroids show annual-cycle trajectories through the SOM pattern space. A distinct feature of the trajectories is the separation between VEMAP and RegCM2 trajectories from September through March, with the VEMAP trajectory residing on the right-hand side of the Sammon map and the model trajectory residing on the left-hand side. The VEMAP trajectory also shows greater movement through the Sammon map during November–January than does the RegCM2 trajectory. The model thus displays less pattern variability than the observations during the cold months. Examination of the three lowest rows in the SOM array (Fig. 2) reveals the differences in simulated (left-hand side) and observed (right-hand side) precipitation patterns. The simulated and observed patterns are similar in the far western United States. The most prominent difference involves the large south-central U.S. precipitation amounts in the observed patterns that the model fails to replicate.

The simulation trajectory is relatively close to the observations trajectory during April–August. This agreement could be a result of fortuitous averaging when computing centroids. However, the frequency distribution for each month's best-fit maps (Otieno 2001) is roughly the same in the simulation and the observations during June–July–August (JJA). The model simulates the spatial pattern of JJA precipitation much better than SON or DJF precipitation.

We obtained model-observation separation of centroids by computing for each calendar month the fre-

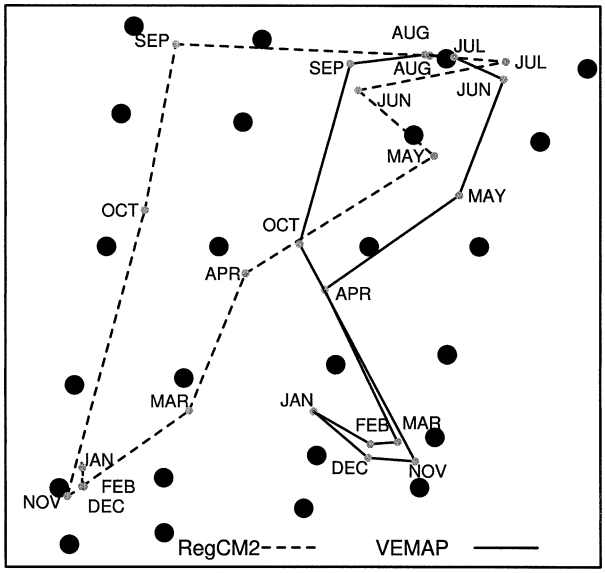


FIG. 3. Sammon map of the SOM array and trajectory of monthly precipitation centroids for VEMAP and RegCM2.

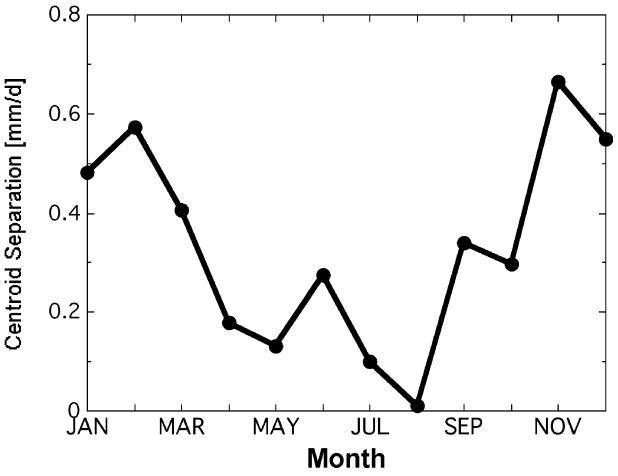


FIG. 4. Annual cycle of the separation between simulated and observed centroids of each month's 10-yr frequency distribution in the SOM array.

quency-weighted average of the maps used to determine the centroid location and then computing the model-observation separation of these averages using (1). Model-observation differences (Fig. 4) are small in July and August, increase to maximum difference in November, and remain large through DJF. The greatest contribution to the separation in SON and DJF comes from the region of large observed precipitation in the Mississippi Delta (MD box).

Centroid paths might differ simply because the simulation output samples only a portion of the SOM space sampled by the observations. For each month, we used the 10-yr frequency distribution of best-fit maps to count the number of distinct maps chosen and to compute monthly coefficients of kurtosis. The map count is a simple measure of the range of patterns in the simulation and the observations. The coefficient of kurtosis is a parametric statistic that measures the breadth of a frequency distribution relative to the normal distribution. Positive kurtosis signifies large tails in a distribution relative to the normal distribution. Both map count and kurtosis (Table 1) show that the simulation output samples as wide a variety of SOM maps as the observations. Thus, the simulation's spread in interannual pattern variability is as large as the observations', albeit shifted with respect to observations.

The bias and bias score also show relatively large simulation error in the south-central United States (Otieno 2001). Seasonal average bias for the 10-yr period (Fig. 5) shows a prominent simulation deficit in this region during SON. There are also narrow bands of positive and negative bias along the U.S. West Coast that appear to result from the model's relatively smooth topography compared to observed topography (Pan et al. 2001b). Bias is generally smaller across the United States during JJA. This is in part due to the annual minimum in precipitation during JJA along the West Coast, but JJA is also the season of annual maximum precipitation for much of the central United States. Simulation − observation bias for the MD box (Fig. 6) shows persistent, relatively large negative values in SON and DJF for each year. In contrast, simulated JJA precipitation tends to have small, positive bias. March–April–May (MAM) has strongly negative bias in the first 2 yr of simulation, but later years show less negative and even slightly positive bias, so its errors are not persistently large as in DJF and SON.

The bias scores (Fig. 7) are mainly less than 1 during SON and DJF, with the score decreasing as threshold increases. The model thus has more difficulty simulating large monthly average precipitation, with fewer events than observed. However, during July and August, the opposite behavior occurs: bias score exceeds 1 and becomes larger with increasing threshold. Even though the model replicates the observed summer precipitation pattern well, it has too many months in the MD box with larger-than-observed monthly average precipitation, and the problem is worse for larger precipitation thresholds. Except for this problem, the SOMs, the bias, and the bias score all point to the same location and time of year as the most prominent precipitation error in the simulation.

This precipitation error appears in other simulations. Giorgi and Shields (1999) used a nearly identical version of RegCM2 to simulate a similar domain, but for

TABLE 1. Annual average of monthly coefficient of kurtosis and count of best-fit maps.

Source	No. of maps	Kurtosis
RegCM2	5.3	9.2
VEMAP	5.5	8.4

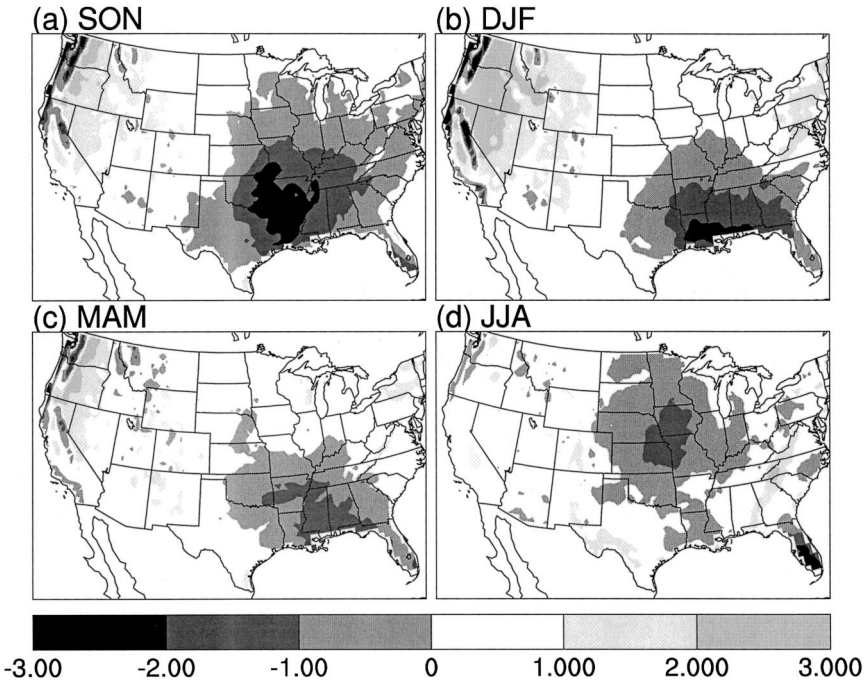


FIG. 5. Seasonal precipitation bias averaged over the 10-yr simulation period for (a) SON, (b) DJF, (c) MAM, and (d) JJA. (Units are mm day^{-1}.)

March 1993–February 1996, with boundary conditions from the European Centre for Medium-Range Weather Forecasts (ECMWF). Their lower Mississippi basin region has large negative bias in SON and DJF. The High Resolution Limited Area Model with Hamburg Physics (HIRHAM; Christensen et al. 1996; Christensen et al. 1998) simulated the same domain and time period as our simulation, with the same boundary conditions (Pan et al. 2001b). HIRHAM has similar negative precipitation bias in the south-central United States (not shown) that emerges in SON and fades during DJF and MAM.

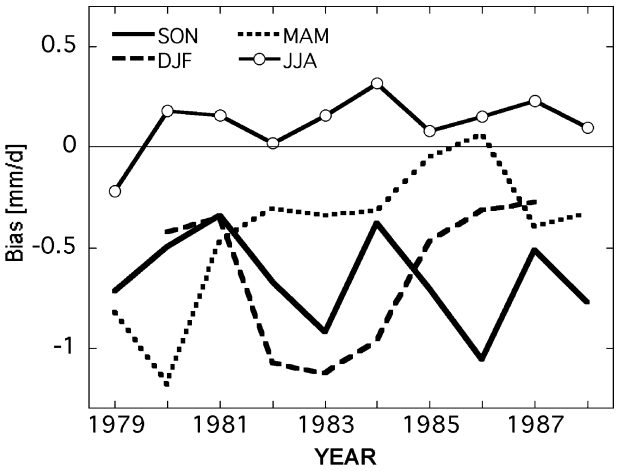


FIG. 6. Interannual variability of seasonal precipitation bias averaged over the MD box. DJF points are plotted for the calendar year of January.

Duffy et al. (2003) have used the NCAR Community Climate Model Version 3 (CCM3) and the climatological annual cycle of SSTs for 10-yr global simulations at resolutions of T42 (corresponding grid spacing ~300 km), T170 (~75 km) and T239 (~50 km). DJF precipitation in all three simulations has negative bias in the south-central United States, with bias decreasing as resolution increases. The bias magnitude in the Mississippi Delta region at T239 is similar to that in our RegCM2 simulation. Finally, J. Boyle (2002, personal communication) has examined monthly precipitation error for the United States in 18 atmospheric global climate models (AGCMs) that simulated 1979–95 for the Atmospheric Model Intercomparison Project (AMIP; Gleckler 2001). A common principal component anal-

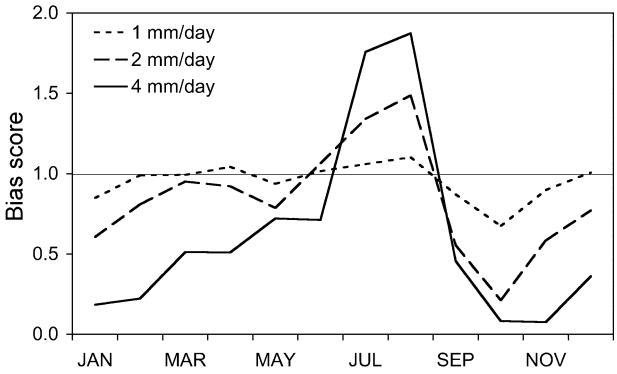


FIG. 7. Annual cycle of bias score for the MD box using three different thresholds.

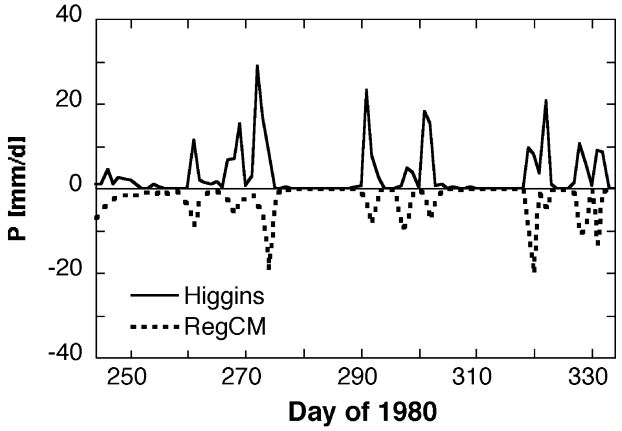


FIG. 8. Time series of daily precipitation in the MD box for SON. Note that RegCM2 output has been multiplied by −1 for clarity.

ysis (Flury 1988) gives a leading principal vector for error in 11 of the models that has its extremum in western Tennessee and Kentucky (i.e., along the northern edge of our MD box), with maximum negative bias in September and October. This feature may also correspond to the bias seen in our simulation.

RegCM2 ingests reanalysis output at its lateral boundaries and thus should replicate observed large-scale features that affect precipitation. Time series of daily observed and simulated SON precipitation for the MD box (e.g., Fig. 8) show that the model reproduces fairly well the observed precipitation episodes as well as intervening dry periods. Similar behavior occurred in all other SON periods, including relatively wet (1984) and dry (1987) SON (Otieno 2001). Because of the match between observed and simulated precipitation episodes, we used the Higgins time series for 1979–88 SON to identify precipitation events, defined as one or more consecutive days when observed, daily precipitation averaged over the MD box exceeded 1 mm. This definition produced 131 precipitation events, averaging 2.8 days. Over 95% of the events were 7 days or shorter, with 33% lasting only 1 day. The simulation had dry bias in 93% of the events, with an average bias per event of -5.4 mm day^{-1}. For the remaining events, the simulation had an average wet bias of $+1.6$ mm day^{-1}. There was no clear year-to-year trend in the bias, but bias magnitude increased through SON, with November having roughly twice the bias per event (-6.2 mm day^{-1}) as September (-3.5 mm day^{-1}). Fractional bias showed no clear dependence on either length or intensity of precipitation episode. The model thus tended to reproduce observed precipitation episodes, but with less precipitation than observed.

4. Sources of precipitation error

In this section we investigate potential sources of the deficit in simulated precipitation in the south-central United States, focusing on the seasons SON and DJF, when the simulation error emerges and maintains large magnitude (e.g., Figs. 4 and 7).

a. Circulation

The model's time-average 500-hPa geopotential heights for SON and DJF are nearly identical to corresponding NNR heights (Otieno 2001). Differences across almost all of the contiguous United States are less than 5 m, which is within the measurement error of observed 500-hPa heights. This behavior is consistent with many regional climate models (Takle et al. 1999; Gutowski et al. 2000).

Applying Blackmon (1976) bandpass and high-pass filters to the geopotential height time series allows comparison of storm track behavior in observations and the simulation. For these filters, bandpass variance is mainly in the range 2.5–6.0 days, whereas high-pass variance is mainly in the range 1–2 days. The 500-hPa bandpass variance (Fig. 9) in both the NNR and the simulation shows two regions of large variance in the northwestern and northeastern United States that correspond to storm tracks. For both seasons, the model tends to have weaker variance in the core of the storm tracks but somewhat larger variance in the central United States. Differences between simulated and NNR bandpass variances are less than 15% generally and even smaller over the south-central United States. The 500-hPa high-pass-filtered variances (Otieno 2001) show slightly weaker simulated versus observed variance in the northeastern storm track and very small differences in the south-central United States. Differences between observed and simulated circulation thus are relatively small for synoptic variability, especially in the south-central United States. This result is consistent with Fig. 8, suggesting that the model produces acceptable large-scale dynamics for generating precipitation events.

Circulation error in shallow weather systems would not be detected in a 500-hPa analysis. However, the dry bias in the overwhelming majority (93%) of our precipitation episodes indicates that nearly all weather systems produced too little precipitation and that a subset of events that were shallow was not the culprit. Indeed, the episodes with largest dry biases tended to be the strongest precipitation producers (not shown), suggesting that relatively deep weather systems were important contributors to the precipitation deficit.

b. Water balance

The atmospheric and terrestrial water balances provide an alternative method for assessing possible contributions to precipitation error, since all errors must balance. This approach needs water balances in terms of quantities available from observations and model output and also needs to recognize potentially significant differences in how the water balance equations are applied to simulated and observed behavior.

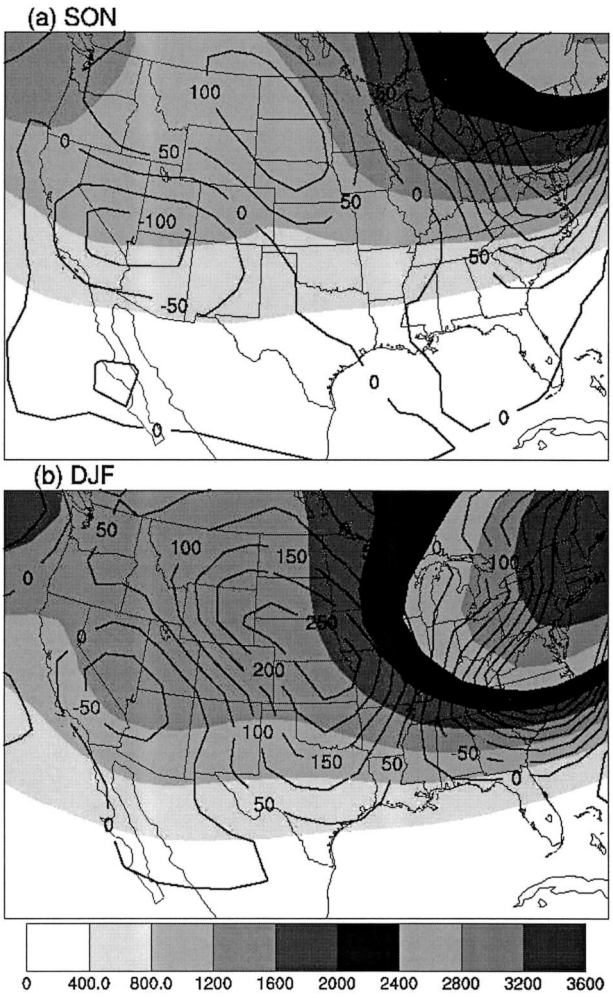


FIG. 9. Bandpass (2.5–6.0 days) variance of NNR 500-hPa geopotential height (shaded) and simulation minus NNR difference (contour lines) for (a) SON and (b) DJF. (Units are m^2.)

Consider the terrestrial water balance equation in the form

$$P - E = \frac{dW}{dt} + R, \tag{4}$$

where E, P, dW/dt, and R are the evapotranspiration, precipitation, change in subsurface storage, and runoff, respectively. Both the model and the real world satisfy (4). However, the terms on the right-hand side of (4) are not defined the same way in both realms. The subsurface storage, W, is the vertical integral over all water reservoirs below the surface. Simulated subsurface storage in RegCM2 is water in the root zone of BATS, while subsurface storage in the real world includes the root zone, vadose zone, and aquifers (e.g., Dingman 1994) and thus includes more water reservoirs than the BATS subsurface. Similarly, the model's runoff is water leaving the BATS land model via overland flow or subsurface drainage from its soil, whereas observed runoff is water entering the river network through overland flow, interflow, and base flow from aquifers (cf. Fekete et al. 1999). Thus, observed runoff is affected by groundwater retention not present in BATS. In a long-term average that eliminates storage changes, observed and simulated runoff should become equal. However, because W and R are not precisely the same in the model and the real world, their annual cycles are not necessarily the same. Thus, for long-term averages, we can evaluate differences in observed and simulated runoff, but for climatologies covering only part of the annual cycle, we have derived expressions for estimating errors in the model's water budget that avoid comparing observed and simulated values of R and dW/dt.

TABLE 2. Estimated water-budget errors in the MD box for a 10-yr average.

Term	Error (mm day^{-1})
ΔP	−1.02
ΔR	−0.71
ΔE	−0.31
ΔC	−0.71

Our analysis also uses the atmospheric water balance

$$\frac{dw_a}{dt} = E - P + C, \tag{5}$$

where dw_a/dt is the change in atmospheric water, and C is horizontal convergence of water in the atmosphere. Averaged over intervals of a month or longer, the atmospheric storage term tends to be very small compared to the others. Terms on the right-hand side of (5) are intended to represent the same processes in the model as in the real world.

1) TEN-YEAR AVERAGE

For long-term averages, the storage terms in both (4) and (5) become negligible. In this case, applying (4) and (5) to the model and the real world separately and taking differences yields simulation-observation errors

$$\overline{\Delta E} = \overline{\Delta P} - \overline{\Delta R}, \quad \text{and} \tag{6}$$

$$\overline{\Delta C} = \Delta\overline{R}, \tag{7}$$

where the overbar represents the long-term average.

We assume that the real world's subsurface storage changes are indeed negligible for our 10-yr period and that its 10-yr average annual cycle of runoff for 1979–88 is the same as UNH/GRDC runoff climatology. Under these assumptions, we use (6) and (7) to estimate time-average errors in atmospheric convergence and surface evapotranspiration for the MD box, thus inferring possible sources of precipitation error. Table 2 shows the 10-yr average ΔP, ΔR, and our estimated ΔE and ΔC. The model has less evapotranspiration than the real world, but this error is smaller than the precipitation error. For this region, the convergence error is roughly

twice as large as the evapotranspiration error, indicating that the primary reason for the model's precipitation deficit in the 10-yr average is a shortage of atmospheric water convergence into the region.

In the long-term average, a region's atmospheric convergence should equal its terrestrial runoff. Assuming this balance occurs for the 10-yr period and that the UNH/GRDC runoff climatology is an accurate estimate of runoff for this period, then the convergence error is 56% of annual average convergence. The proximity of the south-central United States to the model's southern boundary (Fig. 1) indicates that error in NNR moisture flux may be a contributor, since part of the atmospheric moisture in the southern United States evaporates from surfaces south of the simulation domain (Dirmeyer and Brubaker 1999) and must be ingested through lateral boundary conditions.

The accuracy of the model's large-scale circulation (e.g., Fig. 9) indicates that the 10-yr average error is due to error in the model's simulation of atmospheric water vapor. However, the difference between simulated and NNR precipitable water over the south-central United States and adjacent Gulf of Mexico is generally less than 10% for all seasons (Otieno 2001). The simulation appears to maintain a sufficiently full atmospheric water reservoir, but water does not converge and precipitate as intensely as in the real world.

2) SON AND DJF CLIMATOLOGY

The largest precipitation errors occur during fall and winter (Figs. 4–7), so we also analyze the climatological water budgets for SON and DJF. Changes in subsurface water storage may not be negligible over a season. Also as discussed earlier, observed and simulated annual cycles of R and dW/dt may not align even if the model simulates the atmospheric branch of the water cycle accurately because they do not represent precisely the same physical processes in the model and the real world. Thus, we seek error diagnostics that avoid comparing simulated and observed R and dW/dt and that depend as much as possible on quantities available in observations and model output. The second goal minimizes, but ultimately does not eliminate, the need to estimate magnitudes of poorly observed quantities, such as subsurface water storage.

We can write the simulation-observation error in $(P - E)$ as

$$\Delta(P - E) = (P - E)_S - (P - E)_o, \tag{8}$$

where the subscripts s and o refer to simulated and observed values, respectively. Applying (4) to observations gives

$$(P - E)_o = \left[\frac{dW}{dt} + R\right]_o. \tag{9}$$

Combining (8) and (9) gives an estimate for evapotranspiration error

$$\Delta E = \Delta P - [(P - E)_s - R_o] + \left.\frac{dW}{dt}\right|_o. \tag{10}$$

Using (5) to obtain the (observed − simulated) atmospheric water balance and substituting into (10) then yields

$$\Delta C = [(P - E)_s - R_o] - \left.\frac{dW}{dt}\right|_o. \tag{11}$$

Quantities in square brackets in (10) and (11) are known if we assume that the real world's annual cycle of runoff for 1979–88 is the same as UNH/GRDC runoff climatology. Again, we do not use simulated runoff because the model does not include a groundwater reservoir and its contribution to runoff by base flow, which is potentially significant (e.g., Dingman 1994) and which could alter the phase of the annual cycle of runoff.

The error diagnostics (10) and (11) avoid the unobserved E_o but do require estimates of $dW/dt|_o$. Simulated soil moisture change provides one estimate of its value, recognizing that simulated soil moisture does not include water in subterranean aquifers. Note also that if the model matched observations perfectly, then ΔE and ΔP would vanish, and $(P - E)$ and R would constrain $dW/dt|_o$. We assume that $dW/dt|_o$ is within the range of $(P - E)_s$ and R_o values, noting that positive $(P - E)$ or negative R increase subsurface water. We thus infer a range of likely values for $dW/dt|_o$ from $(P - E)_s$, R_o, and the simulated root-zone moisture change. Under these assumptions, (10) and (11) allow us to estimate evapotranspiration and atmospheric horizontal convergence errors using fields available from the model and observations for the MD box.

In the simulation, the annual cycle of soil moisture change tends to mirror the annual cycle of evapotranspiration, implying substantial control of model soil moisture by evapotranspiration (Fig. 10). During SON, when the precipitation deficit becomes large, monthly evapotranspiration wanes, and the change in simulated storage is small though still negative (-0.6 to -0.1 mm day^{-1}). Values of $(P - E)_s$ and $-R_o$ during SON range from -0.5 to $+1.1$ mm day^{-1}. These results suggest that a reasonable range of test values for $dW/dt|_o$ during SON is -0.6 to $+1.1$ mm day^{-1}. During DJF, simulated evapotranspiration becomes small and simulated precipitation increases, producing monthly soil moisture changes of 0.5 to 1.1 mm day^{-1}. The terms $(P - E)_s$ and $-R_o$ range from -2.7 to $+1.8$ mm day^{-1}. A reasonable range of test values for $dW/dt|_o$ during DJF thus appears to be -2.7 to $+1.8$ mm day^{-1}. The DJF estimate is rather liberal because it does not include cancellation of positive R by positive $(P - E)$. The monthly sum $\{(P - E)_s - R_o\}$ suggests that the true value of $dW/dt|_o$ more likely lies between -1.5 and $+0.6$ mm day^{-1}. Although we are not aware of any soil

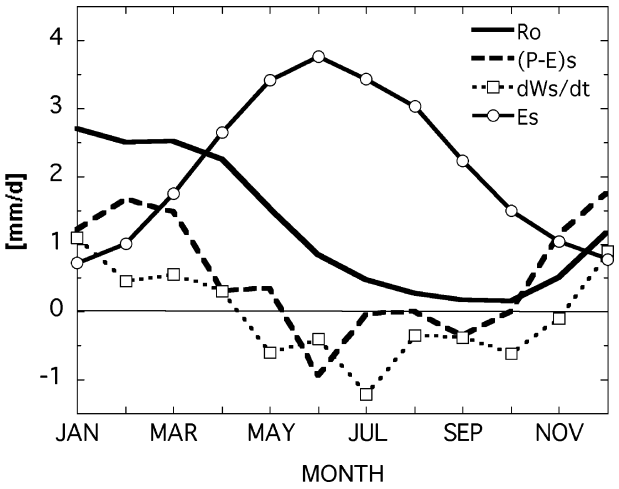
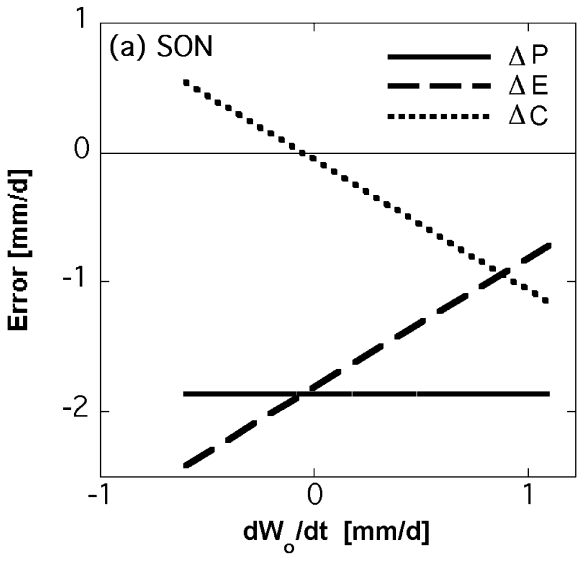


FIG. 10. Annual cycles of selected water balance terms in the MD box.



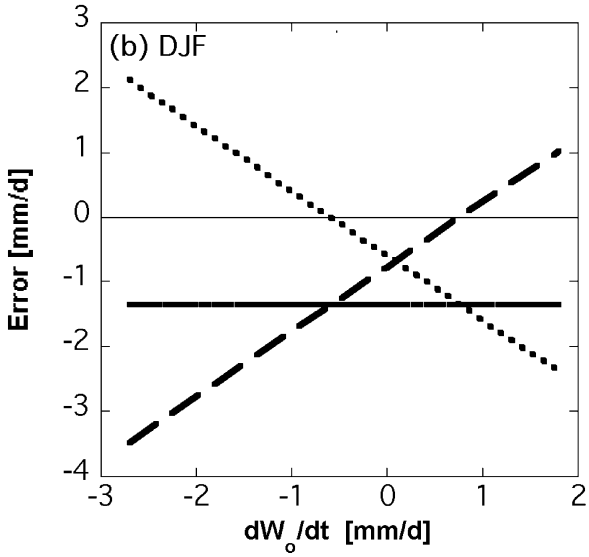


FIG. 11. Errors ΔP, ΔE, and ΔC vs dW_o/dt in the MD box for (a) SON and (b) DJF. Note different scales on (a) and (b); values are plotted for the test range of $dW/dt|_o$ in each season.

moisture measurements covering this domain for the period simulated, the order of magnitude of inferred $dW/dt|_o$ is consistent with moisture changes in the upper 1 m observed in neighboring Illinois for this period (Pan et al. 2001a).

For the MD box during SON (Fig. 11a), the estimated 10-yr average ΔE is more negative than ΔC for almost all test values of $dW/dt|_o$. In addition, for some of the test values, ΔC is positive, indicating that the simulation's circulation may even be importing too much water into the region during SON. The results suggest that the primary cause of precipitation deficit during SON for the south-central United States is a shortfall of surface moisture flux into the atmosphere. This may result from insufficient local moisture recycling or from missing reservoirs of terrestrial water.

During DJF (Fig. 11b), our estimated ''more likely'' range for $dW/dt|_o$, -1.5 to $+0.6$ mm day^{-1}, tends to have larger magnitudes for ΔE than ΔC. However, results for the entire range are mixed; insufficient vertical and horizontal water fluxes both appear to be important for the precipitation deficit. Insufficient moisture convergence during DJF may result from error in the NNR southern boundary condition for moisture and from differences between the observed and simulated DJF storm tracks.

5. Summary and discussion

We have evaluated precipitation from a 10-yr regional climate simulation using a variety of analyses (self-organizing maps, bias scores, and arithmetic bias) to determine the most prominent errors in the simulation. SOMs expose pattern differences, bias scores expose differences in temporal intensity distributions, and the arithmetic bias exposes excessive or deficient simulated precipitation. The three together reveal a precipitation deficit in the south-central United States that emerges in SON and lingers through DJF. The approach illustrates the advantages of using multiple perspectives to ensure that the most important errors emerge. Deficient precipitation for this region and time of year is also evident in other simulations, either by RegCM2 using a different source of boundary conditions or by different models, so the problem is not specific to this particular simulation, though the cause may of course differ between models.

Further analysis shows that large-scale atmospheric circulation is not a major contributor to the SON precipitation deficit in the south-central United States. Simulated precipitation events occur with approximately the same timing as observed, but they generally produce

less precipitation. Error diagnosis based on atmospheric and terrestrial water balances shows that the 10-yr average precipitation error results primarily from a deficit in horizontal water vapor convergence. However, for the 10-yr average over SON alone, the primary contributor appears to be a deficit in evapotranspiration, perhaps suggesting that the model does not produce as much local recycling of water as the real world or that a terrestrial water reservoir is missing in the model.

Several factors contribute to simulated evapotranspiration. Among the most important simulated fields for evapotranspiration in BATS (Dickinson et al. 1993) are soil temperature, soil moisture, and foliage temperature. Soil temperature determines the annually varying vegetation fraction and leaf area indices, whereas foliage temperature helps determine stomatal resistance. Soil moisture helps determine root resistance to transpiration. Cool temperature or dry soil moisture biases in the simulation can yield deficient evapotranspiration. Soil and foliage temperatures are not widely observed, so we use the simulation's surface air temperature bias as an indicator of soil and foliage temperature bias.

The simulated surface air temperature has relatively little bias in daily minimum temperature but does have a cool bias in daily maximum temperature (not shown). This could produce deficient evapotranspiration in BATS. However, there is no clear annual cycle in the temperature bias, and it tends to be less negative in the south-central United States than in adjacent regions. Temperature bias thus does not appear to cause deficient evapotranspiration leading to the south-central United States precipitation deficit. This suggests that there is insufficient water present for evapotranspiration, which may occur if the model's soil is too dry. Excessively dry soil during SON and DJF may occur for several reasons, such as excessive JJA evapotranspiration, insufficient water infiltration into the soil, or excessive drainage from the root zone.

Another factor may be missing surface water. Pitman (1991) and Bonan (1995) have argued that accurate simulation of surface evaporation in many regions may require a parameterization of subgrid-scale water bodies. Bonan's (1995) AGCM simulations showed statistically significant increases in surface evaporation for most, though not all, locations where parameterizations for subgrid-scale lake and swamp/marsh fractions were added, though his analysis focused only on July and January fluxes. The south-central United States contains many wetlands and other inland bodies of water (e.g., NRCS 2000) that are not included in the RegCM2 land use map for this region. Thus, a possible reason for the deficient evapotranspiration could be the lack of these water bodies in the model.

Deficiencies in NNR moisture flux along the model's southern boundary might also contribute to a precipitation deficit by causing insufficient moisture convergence, but the Giorgi and Shields (1999) simulation used the ECMWF analysis for boundary conditions, and the Duffy et al. (2003) simulations used a global model with no lateral boundary conditions. Thus, some factor other than lateral boundary conditions appears to be important. In the Duffy et al. (2003) simulations, the degree of bias diminishes with increasing resolution, indicating that better resolution of mesoscale circulation features such as fronts may be important for adequate simulation of south-central United States precipitation, especially in winter.

The appearance of the SON–DJF deficit in more than one model suggests, though does not guarantee, common causes. However, Bonan's (1995) inland water bodies dataset had negligible wet surface area in the south-central United States, in contrast to NRCS (2000). Better observational characterization of the land surface appears necessary to determine whether or not unresolved surface water could be an important factor in simulations that show an SON precipitation deficit.

Acknowledgments. We thank Chris Anderson for assistance with RCM output processing, Filippo Giorgi and Christine Shields for consultation on RegCM2, Phil Duffy and Jim Boyle for sharing unpublished research with us, Bruce Hewitson and Robert Crane for helpful discussions on SOMs, and the reviewers and editor for their helpful comments. The Climate System Modeling Program of the University Corporation for Atmospheric Research and the Ecosystem Dynamics and the Atmosphere Section, Climate and Global Dynamics Division of the National Center for Atmospheric Research gave access to VEMAP precipitation. NCAR archives also supplied the NCEP–NCAR reanalysis output. This research was supported by the National Oceanic and Atmospheric Administration (Grants NA86GP0572, NA16GP1582, and NA16GP1583), the National Science Foundation (Grant ATM-9911417), the National Institute for Global Environmental Change (Cooperative Agreement DE-FC03-90ER61010), and the Electric Power Research Institute, with additional computing support from NCAR.

REFERENCES

Blackmon, M., 1976: A climatological spectral study of the 500 mb geopotential height of the Northern Hemisphere. *J. Atmos. Sci.,* **33,** 1607–1623.

Bonan, G. B., 1995: Sensitivity of a GCM simulation to inclusion of inland water bodies. *J. Climate,* **8,** 2691–2704.

Briegleb, B. P., 1992: Delta-eddington approximation for solar radiation in the NCAR community climate model. *J. Geophys. Res.,* **97,** 7603–7612.

Brown, J. W., O. B. Brown, and R. H. Evans, 1993: Calibration of Advanced Very High Resolution Radiometer infrared channels: A new approach to nonlinear correction. *J. Geophys. Res.,* **98,** 18 257–18 268.

Cavazos, T., 1999: Large-scale circulation anomalies conducive to extreme precipitation events and derivation of daily rainfall in northeastern Mexico and southeastern Texas. *J. Climate,* **12,** 1506–1523.

——, 2000: Using self-organizing maps to investigate extreme cli-

mate events: An application to wintertime precipitation in the Balkans. *J. Climate,* **13,** 1718–1732.

Christensen, J. H., O. B. Christensen, P. Lopez, E. van Meijgaard, and M. Botzet, 1996: The HIRHAM4 Regional Atmospheric Climate Model. DMI Scientific Rep. 96-4, 51 pp. [Available from the Danish Meteorological Institute, Lyngbyvej 100, Copenhagen Ø, Denmark.]

Christensen, O. B., J. H. Christensen, B. Machenhauer, and M. Botzet, 1998: Very high-resolution regional climate simulations over Scandinavia—Present climate. *J. Climate,* **11,** 3204–3229.

Cressman, G., 1959: An operational objective analysis system. *Mon. Wea. Rev.,* **87,** 367–374.

Daly, C., R. P. Neilson, and D. L. Phillips, 1994: A statistical–topographic model for mapping climatological precipitation over mountainous terrain. *J. Appl. Meteor.,* **33,** 140–158.

Dickinson, R. E., A. Henderson-Sellers, and P. J. Kennedy, 1993: Biosphere–Atmosphere Transfer Scheme (BATS) version le as coupled to NCAR Community Climate Model. NCAR Tech. Note 387 + STR, 72 pp. [Available from the National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307-3000.]

Dingman, S. L., 1994: *Physical Hydrology.* Macmillan College, 575 pp.

Dirmeyer, P. A., and K. L. Brubaker, 1999: Contrasting evaporative moisture sources during the drought of 1988 and the flood of 1993. *J. Geophys. Res.,* **104,** 19 383–19 397.

Duffy, P. B., B. Govindasamy, K. Taylor, M. Wehner, A. Lamont, and S. Thompson, 2003: High resolution simulations of global climate. Part I: Present climate. *Climate Dyn.,* **21,** 371–390.

Fekete, B. M., C. J. Vorosmarty, and W. Grabs, 1999: Global, composite runoff fields based on observed river discharge and simulated water balances. WMO Global Runoff Data Center Rep. 22, 100 pp. [Available online at http://www.grdc.sr.unh.edu/html/paper/ReportUS.pdf.]

Flury, B., 1988: *Common Principal Components and Related Multivariate Models.* John Wiley and Sons, 258 pp.

Giorgi, F., and C. Shields, 1999: Tests of precipitation parameterizations available in the latest version of the NCAR regional climate model (RegCM) over the continental U.S. *J. Geophys. Res.,* **104,** 6353–6376.

——, M. R. Marinucci, G. T. Bates, and G. De Canio, 1993a: Development of a second-generation regional climate model (RegCM2). Part I: Boundary-layer and radiative transfer. *Mon. Wea. Rev.,* **121,** 2794–2813.

——, ——, ——, and ——, 1993b: Development of a second-generation regional climate model (RegCM2). Part II: Convective processes and assimilation of lateral boundary conditions. *Mon. Wea. Rev.,* **121,** 2814–2832.

——, and Coauthors, 2001: Regional climate change information—Evaluation and projections. *Climate Change 2001: The Scientific Basis,* J. T. Houghton et al., Eds., Cambridge University Press, 583–638.

Gleckler, P., cited 2001: Atmospheric Model Intercomparison Project. [Available online at http://www-pcm-di.llnl.gov/amip.]

Grell, G. A., 1993: Prognostic evaluation of assumptions used by cumulus parameterizations. *Mon. Wea. Rev.,* **121,** 764–787.

Gutowski, W. J., and Coauthors, 2000: Project to intercompare regional climate simulations: Advancing the CLIVAR agenda. *CLIVAR Exchanges,* Vol. 5, No. 4, International CLIVAR Project Office, Southampton, United Kingdom, 13–15.

Haas, T. C., 1990: Lognormal and moving window methods of estimating acid deposition. *J. Amer. Stat. Assoc.,* **85,** 643–652.

Hewitson, B. C., and R. G. Crane, 2002: Self-organizing maps: Applications to synoptic climatology. *Climate Res.,* **22,** 13–26.

Higgins, R. W., J. E. Janowiak, and Y. Yao, 1996: A gridded hourly precipitation data base for the United States (1963–1993). NCEP/Climate Prediction Center Atlas 1, U.S. Dept. of Commerce, 47 pp. [Available from NCEP Climate Prediction Center, 5200 Auth Road, Camp Springs, MD 20746.]

Holtslag, A. A. M., E. I. F. de Bruijn, and H. L. Pan, 1990: A high resolution air mass transformation model for short-range weather forecasting. *Mon. Wea. Rev.,* **118,** 1561–1575.

Hsie, E.-Y., R. A. Anthes, and D. Keyser, 1984: Simulations of frontogenesis in a moist atmosphere using alternative parameterizations of condensation and precipitation. *J. Atmos. Sci.,* **41,** 2701–2716.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–471.

Kittel, T. G. F., N. A. Rosenbloom, T. H. Painter, D. S. Schimel, and VEMAP Modeling Participants, 1995: The VEMAP integrated database for modeling United States ecosystem/vegetation sensitivity to climate change. *J. Biogeogr.,* **22,** 857–862.

——, and Coauthors, 1997: A gridded historical (1895–1993) bioclimate dataset for the conterminous United States. Preprints, *10th Conf. on Applied Climatology,* Reno, NV, Amer. Meteor. Soc., 219–222.

——, and Coauthors, cited 2000: VEMAP Phase 2 Historical and Future Scenario Climate Database. [Available online at http://www.cgd.ucar.edu/vemap.]

Kohonen, T., 1995: *Self-Organizing Maps.* 2d ed. Springer-Verlag, 426 pp.

MacAvaney, B. J., and Coauthors, 2001: Model evaluation. *Climate Change 2001: The Scientific Basis,* J. T. Houghton et al., Eds., Cambridge University Press, 471–523.

Main, J. P. L., 1997: Seasonality of circulation in southern Africa using the Kohonen self organizing map. M.S. thesis, Dept. of Environmental and Geographical Science, University of Cape Town, South Africa, 84 pp. [Available from Dept. of Environmental and Geographical Science, University of Cape Town, Private Bag, Rondebosch 7701, South Africa.]

Michaelides, S. C., C. S. Pattichis, and G. Kleovoulou, 2001: Classification of rainfall variability by using artificial neural networks. *Int. J. Climatol.,* **21,** 1401–1414.

NRCS, cited 2000: State of the land—Wetlands. Natural Resources Conservation Service. [Available online at http://www.nrcs.usda.gov/technical/land/wetlands.html.]

Otieno, F., 2001: Evaluation of the seasonal predictive potential of RegCM2 from a 10-yr simulation. M.S. thesis, Dept. of Geological and Atmospheric Sciences, Iowa State University, 63 pp.

Pan, Z., R. W. Arritt, W. J. Gutowski, and E. S. Takle, 2001a: Soil moisture in regional climate models: Simulation and projection. *Geophys. Res. Lett.,* **28,** 2947–2950.

——, J. H. Christensen, R. W. Arritt, W. J. Gutowski Jr., E. S. Takle, and F. Otieno, 2001b: Evaluation of uncertainties in regional climate change simulations. *J. Geophys. Res.,* **106,** 17 735–17 752.

Pitman, A. J., 1991: A simple parameterization of sub-grid scale open water for climate models. *Climate Dyn.,* **6,** 99–112.

Roads, J. O., and A. K. Betts, 2000: NCEP–NCAR and ECMWF reanalysis surface water and energy budgets for the Mississippi River basin. *J. Hydrometeor.,* **1,** 88–94.

——, S.-C. Chen, M. Kanamitsu, and H. Juang, 1999: Surface water characteristics in the NCEP global spectral model and reanalysis. *J. Geophys. Res.,* **104,** 19 307–19 327.

Sammon, J. W., 1969: A non-linear mapping for data structure analyses. *IEEE Trans. Comput.,* **18,** 401–409.

Schimel, D. S., and Coauthors, 2000: Contribution of increasing CO2 and climate to carbon storage by ecosystems in the United States. *Science,* **287,** 2004–2006.

Takle, E. S., and Coauthors, 1999: Project to Intercompare Regional Climate Simulations (PIRCS): Description and initial results. *J. Geophys. Res.,* **104,** 19 443–19 461.

Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences.* Academic Press, 467 pp.