Downscaling of Seasonal Precipitation for Crop Simulation

ANDREW W. ROBERTSON, AMOR V. M. INES, AND JAMES W. HANSEN

International Research Institute for Climate and Society, The Earth Institute at Columbia University, Palisades, New York

ABSTRACT

A nonhomogeneous hidden Markov model (NHMM) is used to make stochastic simulations of March–August daily rainfall at 10 stations over the southeastern United States, 1923–98. Station-averaged observed daily rainfall amount is prescribed as an input to the NHMM, which is then used to disaggregate the rainfall in space. These rainfall simulations are then used as inputs to a Crop Estimation through Resource and Environment Synthesis (CERES) crop model for maize. Regionally averaged yields derived from the NHMM rainfall simulations are found to correlate very highly ($r = 0.93$) with those generated by the crop model using observed rainfall; stationwise correlations range between 0.44 and 0.74. Rainfall and crop simulations are then constructed under increasing degrees of temporal smoothing applied to the regional rainfall input to the NHMM, designed to exclude the submonthly weather details that would be unpredictable in seasonal climate forecasts. Regional yields are found to be remarkably insensitive to this temporal smoothing; even with 90-day low-pass-filtered inputs to the NHMM, resulting yields are still correlated at 0.85 with the baseline simulation, whereas stationwise correlations range between 0.18 and 0.68. From these findings, it is expected that regional maize yields over the southeastern United States will be largely insensitive to year-to-year details of subseasonal rainfall variability; they should be downscalable, in principle, using an NHMM from climate forecasts archived at daily resolution, with the important caveat that the latter need to be skillful enough at the 90-day time scale. As a by-product of the analysis, subseasonal-to-interdecadal summer rainfall variability over the southeastern United States is interpretable in terms of six discrete weather states indicative of a monsoonlike climate regime. Low-simulated-yield years are found to be associated with delayed summer rainfall onset.

1. Introduction

Seasonal predictions of precipitation made with general circulation models (GCMs) are often skillful for some regions and seasons, particularly during El Niño–Southern Oscillation (ENSO) events (e.g., Goddard et al. 2003). These predictions are typically expressed probabilistically, for example, in terms of tercile categories of 3-month-averaged precipitation anomalies. The advent of seasonal climate prediction has raised the possibility of harnessing these predictions for use in decision making in agriculture and other areas of risk management. However, there is a mismatch between the temporal and spatial scales on which the forecast is typically issued and the scales often needed in climate risk management. The grid spacing of GCMs currently used for seasonal prediction is typically about 3° latitude and longitude, while the skillful spatial scale of these models has been argued to be on the order of several grid boxes, that is, about 10° (von Storch et al. 2000). Short time scales (days–weeks) are generally dominated by atmospheric "weather noise," whereas the predictable "signal" of seasonal climate evolves according to surface ocean and land conditions on longer (monthly–seasonal) scales where the signal-to-noise ratio becomes larger (Palmer and Anderson 1994). Crops are known to be sensitive to the caprices of weather at a particular locality, and the frequency and length of dry spells. In clayey soils, they can also be sensitive to water logging, in which case the frequency and length of wet spells are important too.

There is, nonetheless, mounting evidence of the utility of seasonal climate forecasts for agriculture (Hammer et al. 2001; Ingram et al. 2002; Jagtap et al. 2001; Patt et al. 2005; Challinor et al. 2005). While seasonal rainfall totals are often only moderately correlated with crop yields, the latter may be more closely related to

Corresponding author address: Andrew W. Robertson, IRI–Monell 230, Lamont Campus, 61 Route 9W, Palisades, NY 10964-8000.
E-mail: awr@iri.columbia.edu

the frequency of dry and wet spells (Frere and Popov 1986). The crop acts as a nonlinear temporal integrator of weather across its growing season. In certain situations this integration may enhance the seasonally predictable signal of climate, beyond that present in the seasonal rainfall total. Regional averaging of yield is likely to enhance the signal-to-noise ratio further.

Crops are especially sensitive to weather conditions during particular windows of time during the growing season, such as flowering (see Doorenbos and Kassam 1979). Ines et al. (2002) found that even with adequate rainfall at the beginning of the growing season, reduced crop yield is expected if water is not available during the middevelopment to maturity stages of the crops.

Stochastic weather generators, based on a Markov chain assumption for the daily occurrence probability of rainfall, have been shown to be effective at simulating the seasonal statistics of run lengths of dry and wet days (Wilks and Wilby 1999; Wilks 2002), although higher-order chains may be needed in order to capture dry-spell distributions accurately (Wilks 1999). The nonhomogeneous hidden Markov model (NHMM) has proved to be a promising approach to constructing multistation weather generators (Hughes and Guttorp 1994). Over northeastern Brazil, Robertson et al. (2004) found that interannual variability in the frequency of occurrence of 10-day dry spells could be simulated reasonably, using an NHMM with GCM seasonal mean large-scale precipitation as a predictor. Similar downscaling results were obtained over Queensland, Australia (Robertson et al. 2006). The NHMM has been applied to two other locations in Australia in downscaling studies (Charles et al. 2003, 2004).

The hidden Markov model (HMM) factorizes the joint probability distribution of daily rainfall sequences at a network of stations by introducing a small number of discrete rainfall states. The station rainfall occurrence and amount recorded on a particular day is assumed to be conditional only on the state that is active, with only one state active on any given day. The states are termed "hidden" in the sense that they are not directly observable, and a Markov chain is used to probabilistically model the temporal transitions between them. In the nonhomogeneous HMM, the probabilities of transitions between states are modeled as a function of exogenous atmospheric variables, or "predictors." By linking synoptic-scale predictors (i.e., on the scale of the rainfall network) to station-scale daily rainfall, the NHMM can serve to downscale—or disaggregate—in space. In the case of seasonal forecasting the predictors are generally slowly varying in time, in which case the NHMM can act to downscale temporally as well. As a potentially useful by-product, the model's hidden states can provide a synoptic rainfall climatology for the study region, including atmospheric circulation patterns through compositing.

In this study, we test the ability of an NHMM to disaggregate regionally averaged observed rainfall in space and time for crop simulation. In addition to evaluating the NHMM as a downscaling method, our goal is to apply it to determine the effective temporal resolution required of seasonal climate forecasts, in order for them to be useful to agriculture. To do this, we drive the NHMM with observed regionally averaged rainfall, which is progressively smoothed in time, thereby eliminating variability on shorter time scales. The extent to which the NHMM is able to mimic this weather variability stochastically is then evaluated in terms of simulated crop yield. This is an important issue in deciding the temporal resolution that seasonal forecasts need to address from an agricultural perspective (e.g., weekly, monthly, or seasonal). One of our main conclusions is that while seasonal forecasts are currently typically issued as 3-month averages, a 90-day low-pass-filtered daily time series would be more useful for predicting crop yields.

The study is conducted using a network of 10 daily rainfall station records over the southeastern United States. Simulated maize yields obtained using observed rainfall serve as a baseline for evaluating yields derived from NHMM rainfall simulations, made from regionally averaged observed rainfall. This may be interpreted as a perfect model approach, in which the regionally averaged rainfall is taken to be perfectly simulated by a GCM, and errors in the crop modeling are neglected. The HMM, NHMM, crop model, and data used are described in section 2. The HMM states of daily rainfall amounts are derived in section 3. Our main rainfall and maize yield simulation results are presented in section 4. A meteorological description of HMM states is given in section 5, together with an interpretation of subseasonal-to-interdecadal rainfall variability over the southeastern United States in terms of these states. The summary and conclusions are reported in section 6.

2. Data and models

a. Observed datasets

We use daily rainfall amounts at five stations in northern Florida and five in southern Georgia, for the 184-day 1 March–31 August season for 1923–98. These data were obtained from the National Climatic Data Center (see information online at http://www.ncdc.noaa.gov). Stochastic infilling was used to fill data gaps, using the weather generator WGEN (Richardson and
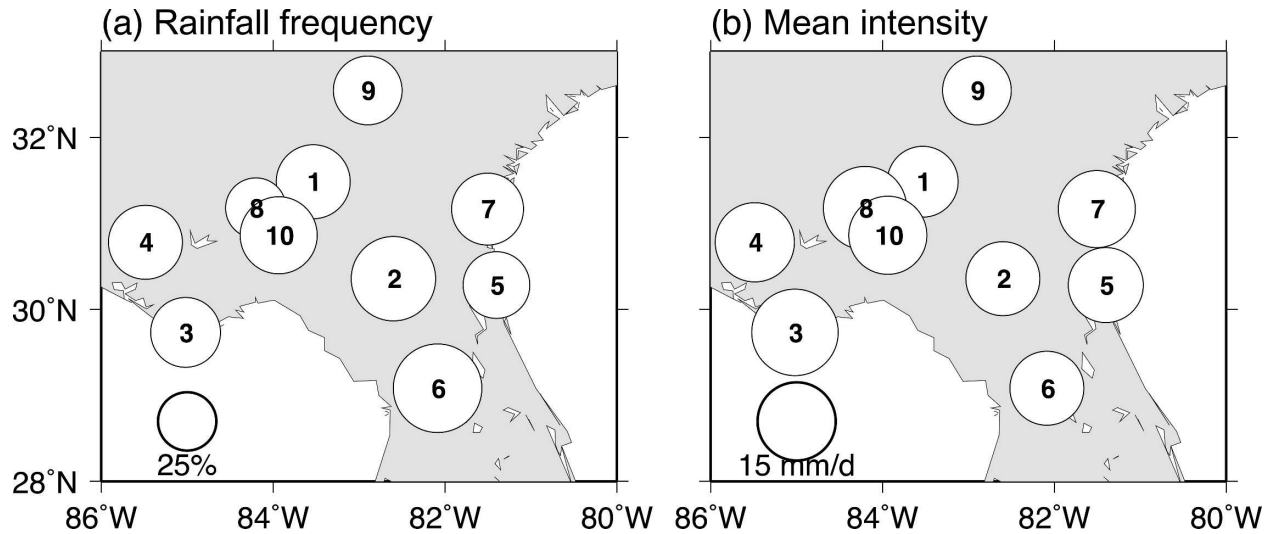
FIG. 1. Rainfall station locations. Circle radius denotes (a) the 1 Mar–31 Aug 1923–98 climatological mean daily rainfall frequency and (b) the corresponding mean rainfall intensity, i.e., mean amount on days receiving ≥1 mm. The stations are (1) Tifton (31.5°N, 83.5°W), (2) Gainesville (30.4°N, 82.6°W), (3) Apalachicola (29.7°N, 85.0°W), (4) Chipley (30.8°N, 85.5°W), (5) Jacksonville Beach (30.3°N, 81.4°W), (6) Ocala (29.1°N, 82.1°W), (7) Brunswick (31.2°N, 81.5°W), (8) Camilla (31.2°N, 84.2°W), (9) Dublin (32.6°N, 82.9°W), and (10) Thomasville (30.9°N, 83.9°W).

Wright 1984) in Weatherman (Pickering et al. 1994). WGEN simulates rainfall for an individual station using a first-order Markov model for rainfall occurrence, and a gamma distribution for intensity on days with rainfall. Four of the station records do not go back to 1923 (Apalachicola, Florida, from 1931; Chipley, Florida, from 1939; Jacksonville Beach, Florida, from 1944; Camilla, Georgia, from 1938), and the stochastic infilling was used to fill these early parts of the records, as well as other small gaps in the records.

Figure 1 shows the locations of the 10 stations together with the March–August climatological daily probability of rainfall occurrence (defined as days with ≥1 mm), and the average wet-day amount. The rainfall occurrence probability is estimated as the relative frequency of daily rainfall over a season, and will be referred to as rainfall *frequency* from now on. The average rainfall amount on wet days will be referred to as rainfall *intensity*. Average rainfall frequency and intensity exhibit similar geographical distributions that are fairly uniform across the 10 stations.

The mean seasonal variation in frequency and intensity is depicted in Fig. 2, in terms of 76-yr averages for each pentad. Frequency decreases to a minimum toward the end of April, and then increases strongly to reach a summer maximum in June and July. Stations 2 and 6 in the north-central part of the Florida panhandle have the largest summer rainfall frequencies. Mean intensities are fairly uniform across the March–August period, so that rainfall seasonality is primarily controlled by rainfall frequency; the "onset" of the summer

rainfall season occurs around mid-June. The multiyear pentad averages of intensity are much noisier than those of frequency, consistent with the findings of Moron et al. (2006).

Relationships with atmospheric circulation are explored using the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis data (Kalnay et al. 1996), using the years 1948–98.

b. The crop model

A crop growth simulation model is a simplified representation of crop growth based on knowledge of ecophysiological processes. They have become essential tools for understanding and predicting crop response to interactions between climate, soil, and management. Crop Estimation through Resource and Environment Synthesis (CERES)-Maize is one of the crop models available in the Decision Support System for Agrotechnology Transfer version 3.5 (DSSATv3.5) (Jones et al. 1998). It simulates the duration of growth, growth rate, and partitioning of new biomass among the economic (ears and grain) and other (leaf, stem, roots) components of the plant (Ritchie et al. 1998). Biomass growth is based on solar radiation intercepted and radiation-use efficiency. Biomass partitioning is a function of the stage of development and source–sink relationships. Yield is determined as the product of plant density, grain numbers per plant, and average kernel weight at maturity. To account for impacts of water deficits on the crop, CERES-Maize simulates the soil water bal-
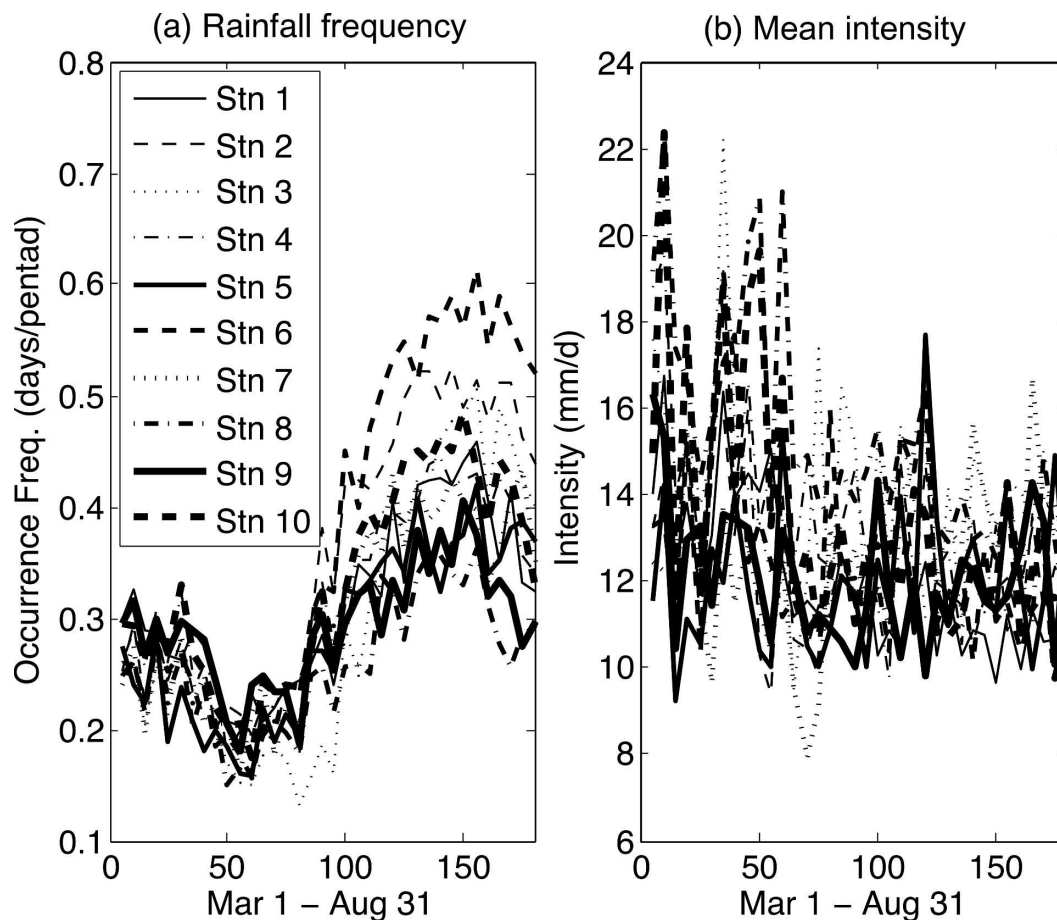
FIG. 2. The mean seasonal cycle of (a) rainfall frequency (days pentad^{-1}) and (b) rainfall intensity (mm day^{-1}) at each station, computed using pentad block means.

ance at a daily time step, as a function of precipitation, irrigation, soil evaporation, transpiration, runoff, and drainage (Ritchie 1998). It uses a modified tipping-bucket approach to account for movement between soil layers. When the capacity of the soil and root system to supply water to the plant constrains transpiration to less than the calculated potential transpiration rate (Priestley and Taylor 1972), potential biomass accumulation is reduced proportionally. The resulting reduction of assimilation affects overall growth. Because the availability of stored carbohydrates near the time of flowering determines the number of grains per plant, yield is particularly sensitive to water deficits during this period. Because leaf area expansion and partitioning between shoots and roots are even more sensitive than biomass accumulation to water deficits, water stress during early vegetative growth can reduce final yields by limiting capacity to intercept solar radiation later in the growing season.

We use CERES-Maize to simulate maize yields for the 10 selected locations during the summer cropping season of March–August for the 74-yr period of 1923–96. Note that the last 2 yr of the rainfall dataset (1997–98) were not used in the crop modeling. Because of our focus on the impact of rainfall variability, daily maximum and minimum temperatures and solar radiation are set to their monthly climatological means, conditioned on the occurrence of rainfall (here ≥0.1 mm). These monthly values for Gainesville (30.4°N, 82.6°W) are used as surrogates for the other sites in Florida, and data from Tifton (31.5°N, 83.5°W) are used for the sites in Georgia. We used soil properties from the Millhopper fine sand, with a plant-extractable soil water (PESW) capacity of 30.9 mm for the top 50 cm for the Florida sites, and from Tifton loamy sand, with a PESW = 48.9 mm for the sites in Georgia. The soil depths used for crop simulations in Georgia and Florida sites are 170 and 180 cm, respectively. The soil columns are assumed to drain freely. The crop cultivar used is McCurdy 84aa. Sowing was on 5 March for the Florida sites and 1 April for the sites in Georgia. Crop growth was simulated without irrigation.

The yields simulated using observed daily rainfall serve as a baseline for this study. In the 10-station average, the resulting simulated yields are only moderately correlated with the observed seasonal total station-averaged rainfall ($r = 0.57$). Thus, only about 32% of the interannual variance of yields (given by r^2) can be represented by a linear regression model of the dependence between seasonal rainfall total and crop yield. Much of this paper is concerned with accounting for the remaining two-thirds of the simulated yield variability.

c. The HMM

The HMM used here follows the approach of Hughes and Guttorp (1994) to model daily rainfall occurrence, while additionally modeling rainfall amounts; it is fully described in Robertson et al. (2004, 2006). In brief, the time sequence of daily rainfall measurements $\mathbf{R}_{1:T}$ on a network of stations is assumed to be generated by a first-order Markov chain of hidden (unobserved) weather states $\mathbf{S}_{1:T} = (S_1, \ldots, S_T)$, where S_t takes values from 1 to K, and K is the chosen number of states. The second defining assumption of the HMM is that the instantaneous rainfall $\mathbf{R}_t$ for a particular day t is assumed to be independent of both (a) all other states, and (b) rainfall on all other days. We further assume that the M station components of the vector of rainfall amounts at time t are conditionally independent of each other given the hidden state S_t; some spatial dependence is, however, captured implicitly via the state variable.

Daily rainfall amount at each station is modeled as a finite mixture of components, consisting of a delta function to model dry days, and a combination of two exponentials to describe rainfall amounts on days with nonzero rainfall. Previous studies have demonstrated that a mixture of two exponentials well represents daily rainfall amounts (e.g., Wilks and Wilby 1999). Fitting the mixture parameters is accomplished as an integral part of the HMM, through the expectation-maximization (EM) algorithm (Dempster et al. 1977), similar to the approach of Bellone et al. (2000). Details of the EM estimation algorithm were presented by Robertson et al. (2003) for a model that is similar, except that precipitation occurrence data are modeled instead of amounts. The additional EM equations required to handle estimation of the parameters for the state-dependent amount models above are described in Kirshner (2005).

In the nonhomogeneous HMM the state transition matrix $\mathbf{\Gamma}$ is no longer stationary, and the transition probabilities are defined to be a function of a (potentially) multivariate "predictor" input time series $\mathbf{X}_{1:T}$, corresponding (e.g.) to other variables that can influence the evolution of the weather-state sequence $\mathbf{S}_{1:T}$. In this paper, the transition probabilities are defined as a logistic function of the 10-station-averaged observed daily rainfall amount, used as the (univariate) predictor variable. The regional average is standardized by subtracting its time mean over the 76 seasons, and dividing by the daily standard deviation. The logistic function then maps this real-valued daily predictor series onto a probability value, bounded between 0 and 1 (see Fig. 15 of Robertson et al. 2004). More complete details on this type of model are provided in Hughes et al. (1999) and Robertson et al. (2003).

The stochastic simulations of daily rainfall amount are used as inputs to the crop model, simulating the crop growth for each of 100 NHMM simulations, over the period 1923–96. The resulting yields are compared with baseline yields derived from the crop model when using observed station rainfall itself.

3. States of daily rainfall amounts

a. Number of states

As in Robertson et al. (2004), cross validation is used to evaluate the quality of the fitted HMMs in terms of log-likelihood as a function of K, the number of states. Here, 5-yr blocks of data were withheld, the model was trained on the remaining 70 yr (omitting the last year of the dataset), and the simulations were compared with observed rainfall for the fourteen 5-yr validation periods. In each case the EM algorithm was run 10 times from different initial seeds, selecting the run with the highest log likelihood. The resulting log-likelihood values for each model were examined for $K = 2$–10 and were found to increases monotonically with K, despite the use of cross validation (not shown). Similar results were obtained for the Bayesian information criterion (BIC), a penalized likelihood measure that is often used to determine the appropriate number of states in HMMs. We chose $K = 6$, where the log-likelihood values start to flatten out. Using a manageable number of states (here 6) enables a more parsimonious description of the rainfall variability that is better suited to the interpretation in section 5. The results for $K = 3$ were also inspected.

b. Estimation of the model parameters

Having chosen the six-state model, its parameters were estimated from the entire 76-season rainfall record. The resulting rainfall parameters are illustrated in Fig. 3, in terms of the probability of rain (Figs. 3a–c, h–j) and the mean rainfall intensity (Figs. 3d–f, k–m);
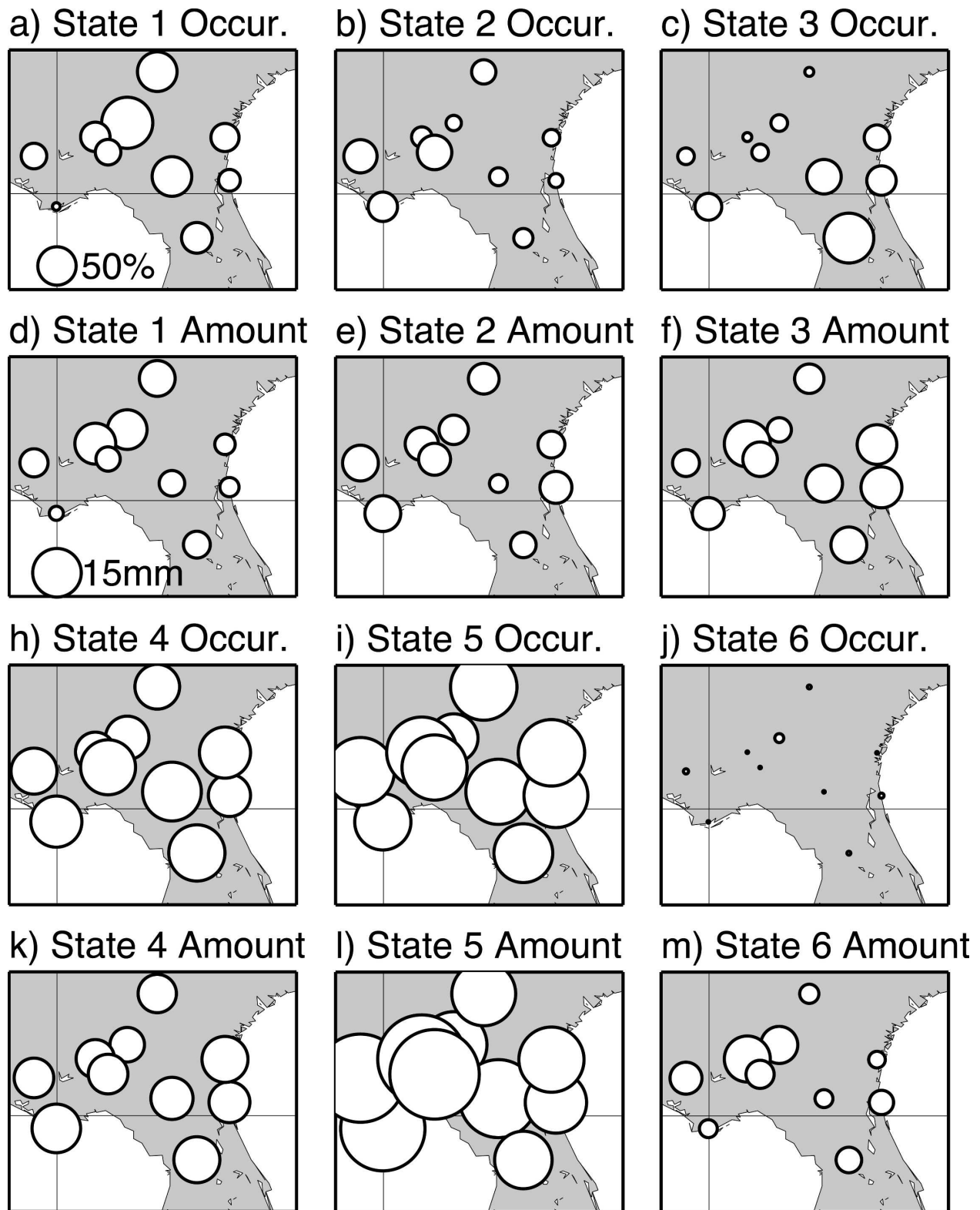
FIG. 3. HMM rainfall parameters: (a)–(c), (h)–(j) rainfall occurrence probabilities (circle radius), and (d)–(f), (k)–(m) mean amounts on days receiving greater than 0 rainfall, calculated from the parameters of the mixed exponential distribution.

TABLE 1. Transition probabilities between HMM hidden states.

		To state					
		1	2	3	4	5	6
From state	1	0.12	0.18	0.16	0.05	0.00	0.48
	2	0.20	0.36	0.06	0.19	0.11	0.08
	3	0.00	0.11	0.60	0.18	0.01	0.09
	4	0.23	0.01	0.10	0.60	0.05	0.01
	5	0.34	0.00	0.01	0.12	0.34	0.18
	6	0.00	0.19	0.04	0.02	0.01	0.74

the latter was computed from the parameters of the mixed exponential distribution.

State 6 is the "dry" state with very small rainfall probabilities. States 4 and 5 are both "wet" with rainfall probabilities around or exceeding 0.5, uniformly at all stations. State 5 is characterized by very large mean wet-day amounts (>20 mm day^{-1}). States 1–3 are characterized by spatial contrasts, particularly visible in occurrence probabilities; states 2 and 3 tend to mirror each other, with northwest–southeast contrasts in probability. State 1 has larger occurrence and wet-day amounts in the northwest.

The state transition matrix is given in Table 1. The larger self-transition probabilities on the main diagonal indicate substantial temporal persistence, particularly for states 6, 3, and 4. On the other hand, persistence is very low for state 1, which has a larger probability of transitioning to states 2, 3, and especially 6, than persisting. There are other preferred transitions, such as from states 2, 4, and 5 to state 1, and we will comment on their meteorological interpretation in section 5 below.

4. Simulations

In this section, we use the NHMM to make spatially disaggregated rainfall simulations at each of the 10 stations, using the station average of observed daily rainfall as input to the NHMM. The 10 stations have similar climatological frequencies and intensities (Fig. 1), justifying a simple arithmetic average among stations. The resulting stochastic rainfall simulations are then passed to the crop model, which is integrated for each station in turn. In addition to examining the NHMM's ability to spatially disaggregate regionally averaged rainfall, we consider temporally low-pass-filtered versions of the input time series. In this way, we examine the NHMM's temporal disaggregation. We evaluate the model in terms of both its rainfall simulations and crop yields derived from these simulations, using 100 stochastic simulations of the 1923–98 rainfall record. No cross validation is used. However, we have repeated the cases with unfiltered and 90-day low-pass-filtered inputs, using leave-6-yr-out cross validation. The correlations re-

ported below are found to be almost unchanged under cross validation.
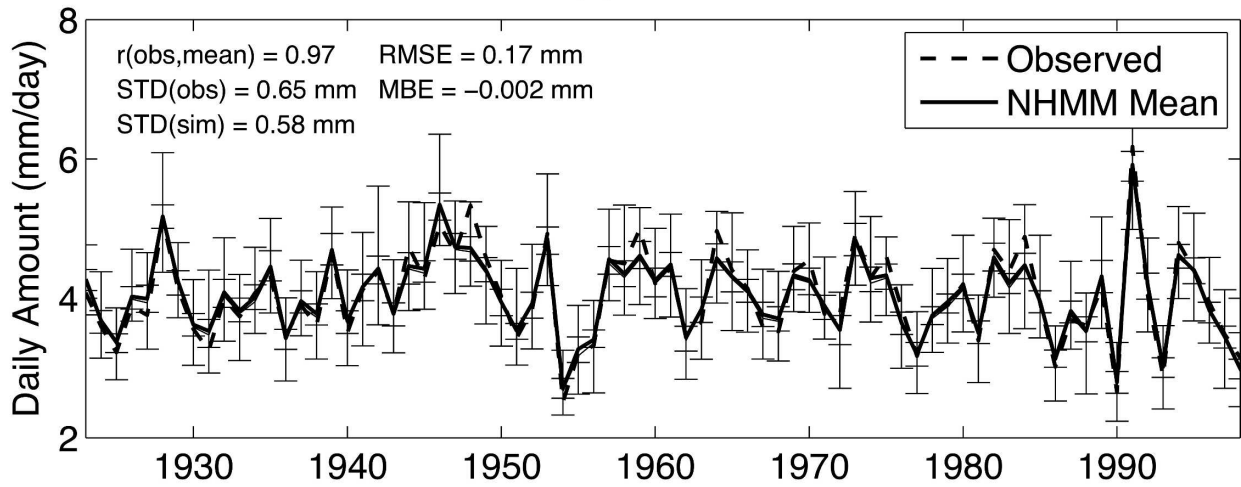
a. Rainfall simulations

The NHMM simulations of rainfall made from daily unfiltered regionally averaged observed rainfall are shown in Fig. 4, plotted in terms of the seasonal averages of amount (Fig. 4a), rainfall frequency (Fig. 4b), and intensity (Fig. 4c), averaged over the 10 stations. The ensemble mean of the 100 simulations is plotted versus the observed, together with the interquartile and full range of the 100-member simulation distribution. Some summary statistics are included in each panel. Interannual variability of rainfall amount is very well simulated, with an anomaly correlation $r = 0.97$ and root-mean-square error (RMSE) of 0.17 mm, demonstrating that the NHMM is successful at the regional scale. The simulations of interannual variability of daily rainfall frequency ($r = 0.77$, RMSE = 3.8 days per season) and intensity ($r = 0.67$, RMSE = 0.99 mm) are less good.

The interannual variance of the simulated ensemble mean is underestimated by 25% for amount and intensity, and 21% for rainfall frequency. This bias in amount and frequency is largely a result of averaging over the 100 simulations: on average, the individual simulations underestimate the variance of amount by only 6%, and the frequency by 9%. However, the individual simulations overestimate the variance of intensity by 20%. The mean bias errors are very small (all within about 1%), although some nonstationarity is visible in rainfall frequency, with the simulations tending to overestimate the occurrence of rain prior to about 1940, and to underestimate it thereafter. This nonstationarity largely disappears if the four stations with data missing during the early part of the record are omitted from the analysis, that is, Apalachicola, Chipley, Jacksonville Beach, and Camilla (not shown).
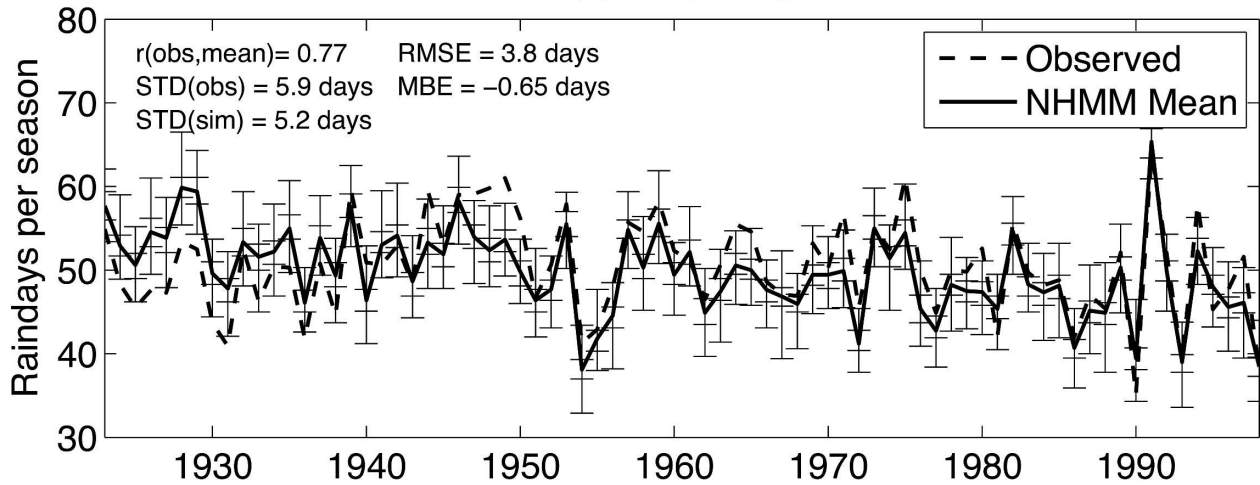
The anomaly correlation in Fig. 4 is highest for amount, which is to be expected because rainfall amount is used to drive the NHMM. If instead station-averaged rainfall frequency is used as the NHMM's input, the resulting interannual correlations for amount, frequency, and intensity become $r = 0.80$, 0.98, and 0.15, respectively. Thus, frequency is a very poor predictor of intensity.

Because crop water stress is influenced by the timing and frequency of dry spells, we examined the 10-day dry-spell frequency. Anomaly correlations between simulated and observed seasonally averaged 10-day dry-spell frequency are tabulated in Table 2. Here a dry spell is defined as a run of at least 10 dry (i.e., rainfall ≤ 1 mm) days, with no more than one intervening wet
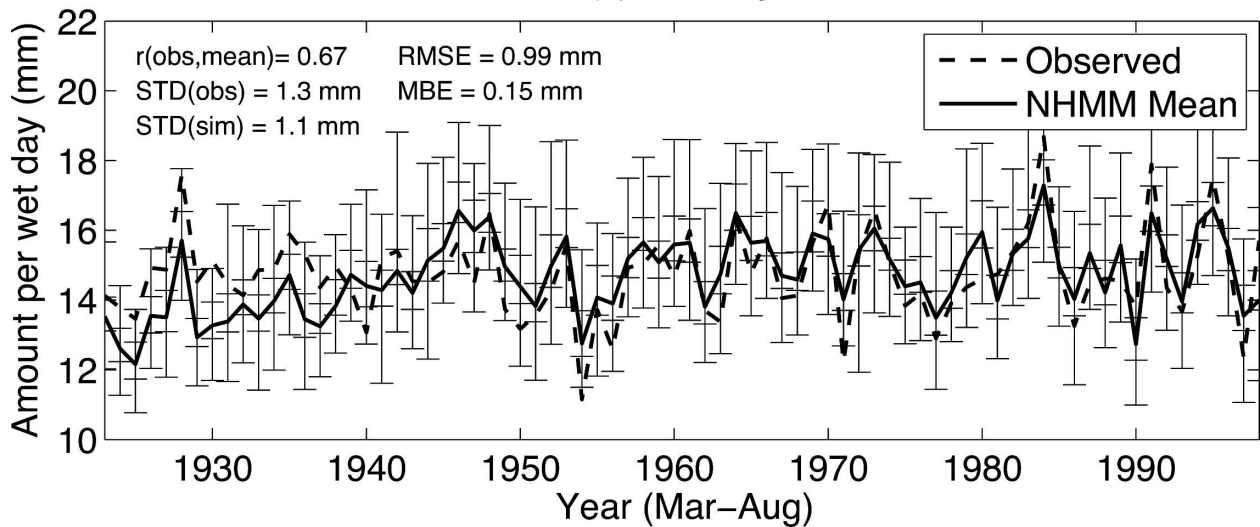
(a) Amount

r(obs,mean) = 0.97   RMSE = 0.17 mm
STD(obs) = 0.65 mm   MBE = −0.002 mm
STD(sim) = 0.58 mm

- - - Observed
—— NHMM Mean

(b) Frequency

r(obs,mean)= 0.77   RMSE = 3.8 days
STD(obs) = 5.9 days   MBE = −0.65 days
STD(sim) = 5.2 days

- - - Observed
—— NHMM Mean

(c) Intensity

r(obs,mean)= 0.67   RMSE = 0.99 mm
STD(obs) = 1.3 mm   MBE = 0.15 mm
STD(sim) = 1.1 mm

- - - Observed
—— NHMM Mean

Year (Mar–Aug)

TABLE 2. Correlations between simulated and observed seasonal rainfall amount, 10-day dry-spell frequency, and crop yield; shown are NHMM simulations with daily regional rainfall input, and 90-day low-pass-filtered input. Note that Avg* in last row is the correlation between station-averaged quantities.

Station	Rainfall		Dry spells		Crop yield	
	Daily	90-day	Daily	90-day	Daily	90-day
1	0.63	0.62	0.34	0.27	0.70	0.51
2	0.68	0.75	0.47	0.1	0.63	0.62
3	0.62	0.66	0.36	0.29	0.43	0.18
4	0.55	0.56	0.24	0.26	0.44	0.57
5	0.52	0.50	0.34	0.17	0.62	0.37
6	0.56	0.53	0.11	0.15	0.55	0.48
7	0.72	0.71	0.4	0.15	0.61	0.61
8	0.60	0.66	0.35	0.19	0.56	0.55
9	0.60	0.52	0.22	0.092	0.66	0.51
10	0.56	0.58	0.49	0.13	0.76	0.68
Avg*	0.97	0.97	0.68	0.39	0.93	0.85

day. Allowing an intervening wet day in the definition substantially reduces the NHMM's tendency to underestimate the number of long dry spells.[1] The mean bias error in the 10-day dry-spell frequency is always less that 15% at the individual stations, and is 2%–5% for the station average, depending on the input series. The number of observed 10-day dry spells varies between 1 and 9 per season, across the 10 stations, with dry spells occurring predominantly during spring.

The interannual variability of station-averaged dry-spell frequency is less well simulated ($r = 0.68$) than rainfall frequency ($r = 0.77$), and the interannual variance of the simulated series of dry-spell frequency is only 41% of the observed one. These correlation values improve when rainfall frequency is used as the predictor, in place of amount, reaching 0.82 for dry spells and 0.98 for rainfall frequency. The poorer simulation of dry-spell counts can largely be attributed to the effects of statistical sampling: if the individual simulations are compared with the 100-member ensemble mean (again in terms of station averages), it is found that 95% of them have anomaly correlations of dry-spell counts less

---

[1] The first-order Markov chain tends to underrepresent long dry spells. In addition, because the Markovian dependence in the NHMM is modeled at the state level, the conditional dependence of station rainfall on the state may further reduce wet-/dry-spell lengths (Robertson et al. 2004).

than or equal to 0.84, while the corresponding figure for rainfall frequency is 0.96. Thus, the individual simulations naturally differ more from each other in terms of dry-spell counts than they do in terms of rainfall frequency.

The NHMM's performance at the individual station level is tabulated in Table 2 in terms of anomaly correlation for seasonal rainfall amount and dry-spell frequency. The last row of the table shows the correlation between the station-averaged quantities. Clearly, while near optimal for the station-averaged simulation, the simulations at the individual stations are much less successful, with interannual correlations with the observed of $r = 0.52$–0.72 for rainfall amount, but reaching only $r = 0.11$–0.49 for dry-spell counts.

Table 2 also shows the results when driving the NHMM with 90-day low-pass-filtered rainfall amount. There is little impact on simulated rainfall amount, while dry-spell frequency is seriously degraded. A similar degradation takes place even when rainfall frequency is used to drive the NHMM. Despite the large sampling uncertainty in dry-spell counts, the implication is that the number of dry spells is significantly impacted by year-to-year differences in weather details that cannot be replicated stochastically by the NHMM. This may be associated with the short memory of the geometric distribution implied by the first-order Markov model.

b. Crop yield simulations

Figure 5 shows the simulated station-averaged yields averaged over the 100 simulations. The curves in Fig. 5 show the yield obtained using either daily unfiltered or temporally smoothed inputs to the NHMM. The results are summarized in Table 3 in terms of anomaly correlation, mean bias error, and RMSE. For comparison, the bottom two rows of Table 3 show the results obtained 1) without any downscaling and 2) with a simple downscaling using local bias correction. In the first approach, we simply use the 10-station average of the observed daily rainfall to drive the crop model at each station. In the second approach, we use a local bias correction of the regionally average observed daily rainfall, according to each station's cumulative distribution function (Ines and Hansen 2006).

←

FIG. 4. Interannual variability of NHMM-simulated seasonal rainfall (a) amount, (b) frequency, and (c) intensity, using station-averaged daily rainfall amount as input. Plotted is the mean of 100 NHMM simulations averaged over all 10 stations (solid) vs the observed (dashed). The error bars indicate the entire range of the 100 simulations, with the interquartile range given by the inner ticks. The correlations between the curves are shown in the corner of each subplot, together with their variances. The interannual rainfall statistics given in each panel are anomaly correlation (r), std dev (STD), RMSE, and mean bias error (MBE).
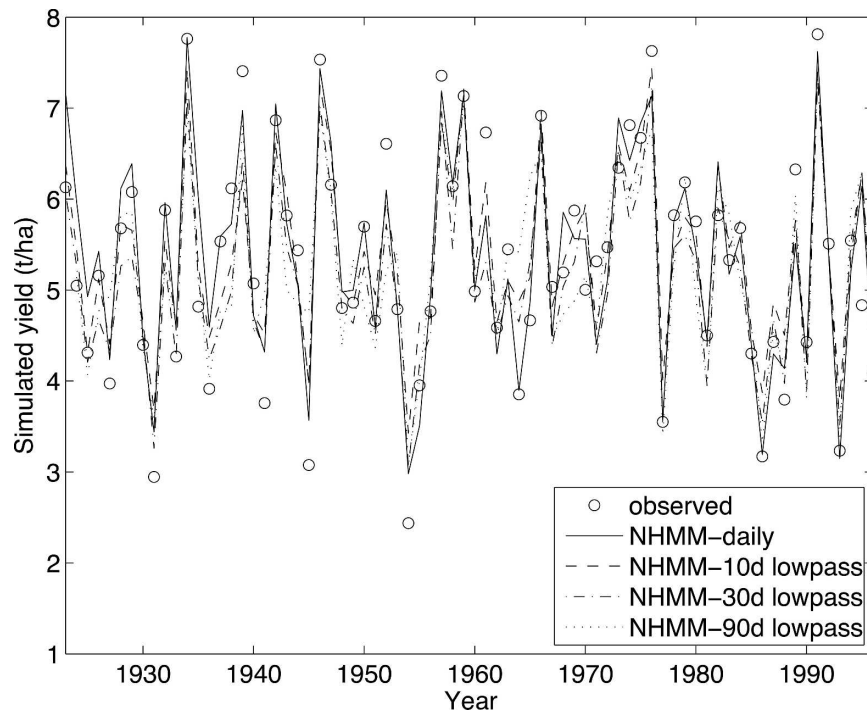
FIG. 5. Interannual variability of simulated station-averaged yield using observed (circles) and NHMM-simulated (lines) rainfall. The latter also show the yields obtained using NHMM rainfall simulations made with low-pass filtering of the regional rainfall time series used to drive the NHMM. The mean bias error is less than 4% in all cases.

Interannual variability of regional yields is well simulated by the NHMM–crop model combination when daily station-averaged rainfall is used as input to the NHMM ($r = 0.94$), while the variance is underestimated by about 13%. The NHMM performs considerably better than the two cases of no downscaling, or local bias correction, in terms of anomaly correlation, mean bias error, and RMSE. The two simple schemes perform comparably, but with the mean bias much reduced in the local bias correction. However, even the latter produced crop yields with a considerable amount of bias in this case.

The simulation of regional yield is scarcely degraded, in terms of correlation and RMSE, when the NHMM input time series is low-pass filtered at either 10 or 30 days, with a slight degradation using a 90-day low-pass-filtered input ($r = 0.85$). However, the yield variance is underestimated by 35%–40% when the low-pass-filtered inputs are used. This loss of yield variance is inevitable because the stochastic high-frequency rainfall variability, generated by the NHMM, becomes averaged out in the ensemble mean. This contrasts with the unfiltered daily input case, in which the single observed daily weather sequence is prescribed to be the input in all realizations.

The right-hand columns in Table 2 show the yield correlations at the individual stations, for daily and 90-day low-pass-filtered inputs to the NHMM. It is notable that the yield correlation values almost reach those of rainfall amount, and are considerably higher than those of 10-day dry-spell frequency.

The insensitivity of the yield simulations to temporally smoothing the regional rainfall input series is striking. It suggests that crop yield is not sensitive to the sequence of daily weather particular to each year but,

TABLE 3. Performance of yield simulations using NHMM downscaling under varying degrees of temporal smoothing of NHMM input, i.e., daily, 10-day, 30-day, and 90-day low-pass filtering. Bottom two rows show crop yields derived from observed (unfiltered) daily rainfall with no downscaling, and with simple local bias correction.

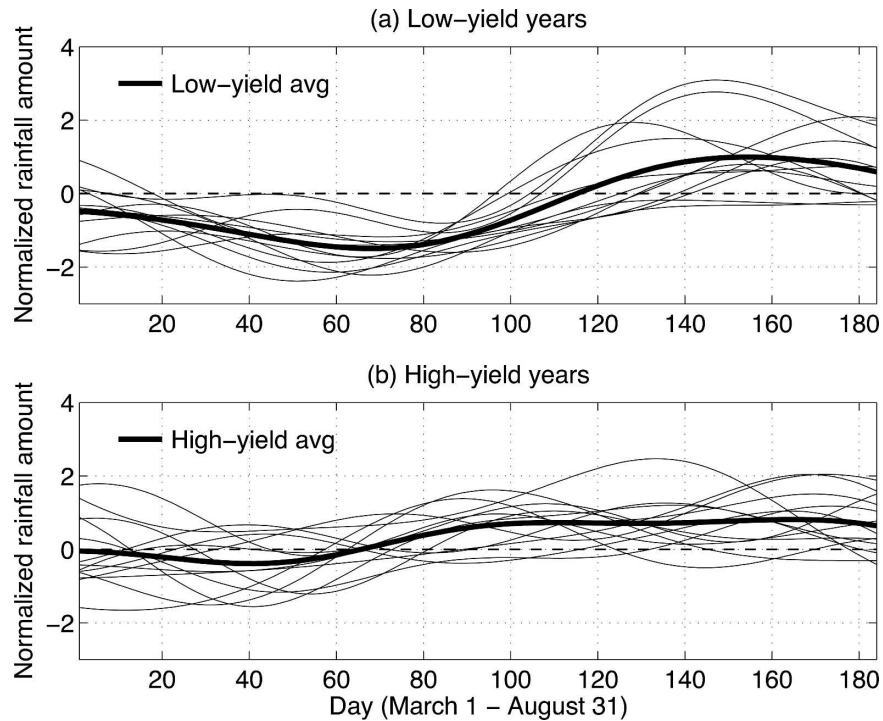|	Yield correlation	Mean bias error (kg ha^{-1})	RMSE (kg ha^{-1})
NHMM daily	0.94	55.4	441.8
NHMM 10 day	0.93	70.6	497.1
NHMM 30 day	0.91	−171.7	545.5
NHMM 90 day	0.85	−89.7	655.6
No downscaling	0.82	1279.0	1484.9
Local correction	0.80	−448.7	1367.8

FIG. 6. Normalized 90-day low-pass-filtered regionally averaged rainfall: (a) low-simulated-yield years and (b) high-simulated-yield years. Years were selected using a threshold of one standard deviation about the 1923–96 mean. The thick line shows the multiyear average in each case.

rather, that this can be represented stochastically, as a function of a 90-day smoothed input. It also suggests that the latter contains most of the predictive value, as far as yield is concerned. Figure 6 contrasts the years with low versus high simulated yields, in terms of the 90-day low-pass-filtered input series, using a 1 standard deviation selection criterion. A very clear distinction in rainfall seasonality emerges between the two sets of years: low-yield years are characterized by anomalously low regionally averaged rainfall during the first 90 days of the season (March–May), while high-yield years tend to have above-average regional rainfall during May–August. The largest difference between the two sets of years occurs around May. Both high- and low-yield years tend to have above-average regional rainfall in July–August.

We have also driven the NHMM with monthly and seasonal (184 day) totals of rainfall, prescribing the monthly (or seasonal) value on each day of the respective month (or season) as input to the NHMM, again using standardized values. Using monthly totals, we obtain an anomaly correlation of 0.87, which is comparable to the 90-day low-pass value. Using (184 day) seasonal rainfall totals, the respective correlation is only 0.57. However, recall that this value is equal to the correlation between the seasonal rainfall totals themselves and the baseline yields reported in section 2b. Consistently, we find that seasonal rainfall totals are almost perfectly correlated ($r = 0.97$) with yield simulated from seasonal rainfall total via the NHMM/crop model. A scatterplot of this relationship (not shown) reveals a near-linear dependence within the range of observed rainfall totals, with a slight flattening out for the highest seasonal totals. The modest correlation of $r = 0.57$ between rainfall totals and baseline yields can thus be attributed almost entirely to the omission of subseasonal time-scale rainfall variability, as opposed to any nonlinearity in the relationship. CERES-Maize uses two linear water stress factors that are functions of the ratio between the water-supplying ability of the soil and root system and the evaporative demand (Vaux and Pruitt 1983). It does not capture the direct effects of water logging on the root system, but will still give declining simulated yields at high rainfall amounts resulting from the reduced solar irradiance under very rainy conditions.

5. Interpretation of the HMM states

Beyond its ability to generate daily sequences of local rainfall, conditioned on large-scale rainfall, the HMM
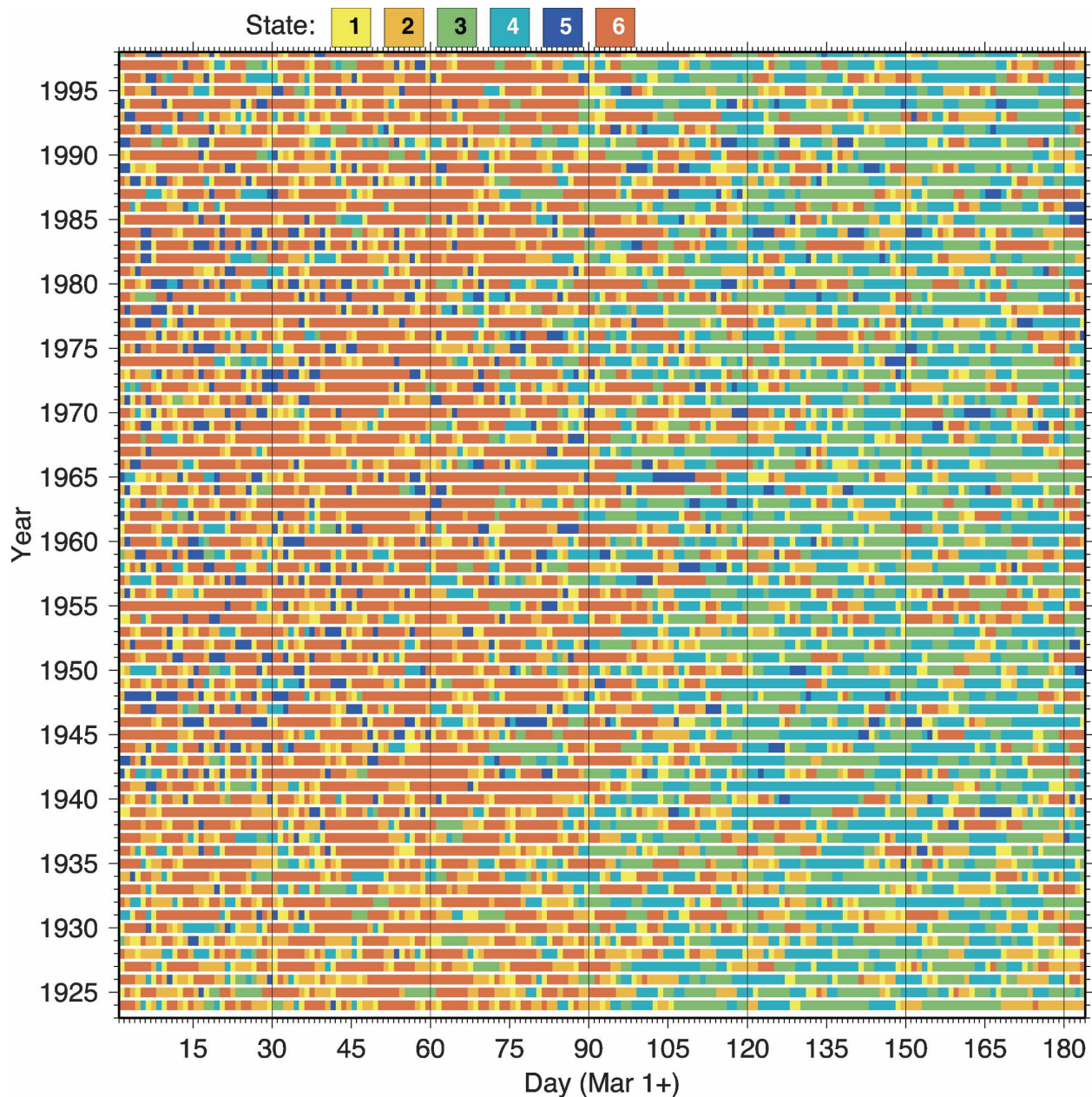
FIG. 7. The estimated HMM state sequence, with colors running from red (driest) to dark blue (wettest).

can provide potential insight into the rainfall process, through inspection of the hidden states and their chronological sequence. These rainfall states provide a diagnostic of large-scale weather conditions across the region on a daily basis.

a. The estimated state sequence

Once the parameters of the HMM have been determined from the rainfall observations, the most probable daily sequence of the six states can be estimated using the Viterbi algorithm (e.g., Rabiner 1989). This allows for an interpretation of the 76-yr rainfall record in terms of these states by assigning each day to the state that was most probable on that day. The sequence is plotted in Fig. 7, from which the relative frequencies of the six states can be simply counted; they are 6.1%, 8.2%, 10.1%, 12.0%, 3.5%, and 24.5% respectively.

The state sequence exhibits strong seasonality, with the dry state (state 6) dominating during March–May, and states 3 and 4 dominating in June–August; state 4 is wet, and state 3 is fairly wet in the south. The wettest state (state 5) occurs on only 3.5% of days, with a slight
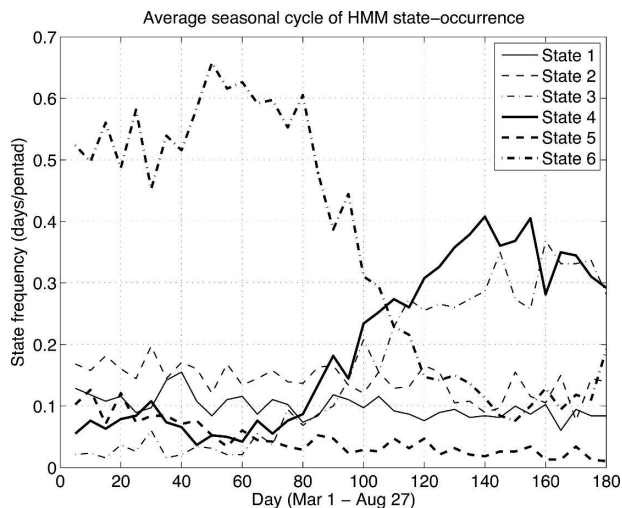
FIG. 8. The mean seasonal cycle of HMM state occurrence, computed from pentad means (days pentad^{-1}).

preference for the spring. States 1 and 2 exhibit little seasonal change in occurrence. States 1, 2, and 5 are highly transient, consistent with the low persistence seen in Table 1, while states 3, 4, and 6 exhibit persistent spells. The average seasonality is plotted in Fig. 8, and suggests a description of the average seasonal evolution in terms of the rainfall states. Figure 7 indicates that the onset of the rainy season is abrupt—near the beginning of June—but that there is also a substantial amount of within-season and year-to-year variability. Thus, the dry state 6 can occur even in the peak of summer.

Because low- and high-yield years exhibit distinctly different rainfall seasonality (Fig. 6), we examine this contrast in terms of state frequency. Figure 9 shows the seasonal cycle of state frequency, averaging over each of these two sets of years. There is a clear distinction in the frequency of states 4 (wet) and 6 (dry) between days 80 and 100 (mid-May–mid-June), with a much higher (lower) prevalence of the dry (wet) state in the low-yield years. Thus, the seasonal summer rainfall onset is delayed in low-yield years.

b. Interannual variability

The interannual variability in state frequency is plotted in Fig. 10. The prevalence of the two dominant states [6 (dry, spring) and 4 (wet, summer)] tends to vary inversely, indicating interannual differences in the length of the summer rainfall season or the within-season intermittency of these states.

An interdecadal trend toward drier conditions in also visible since the 1950s, again mostly in states 4 and 6. There are also trends in the early part of the record,
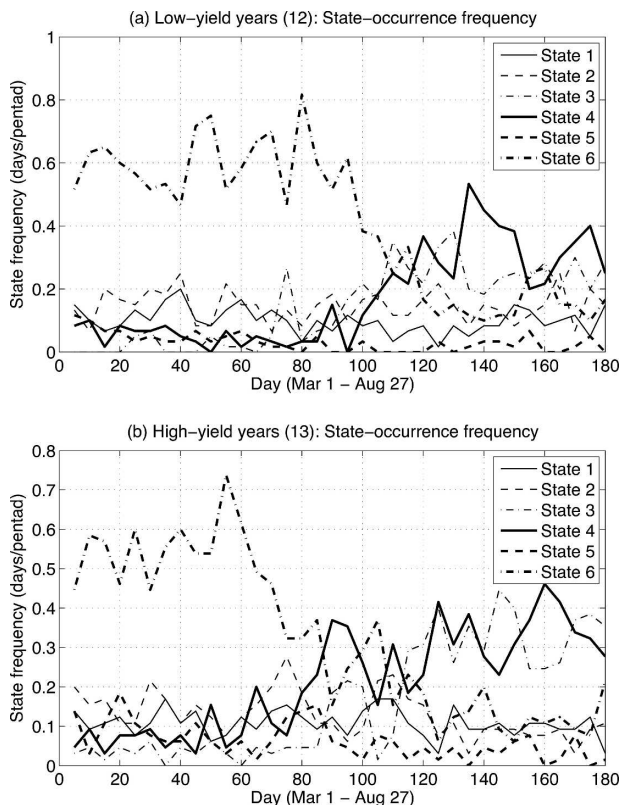

FIG. 9. The mean seasonal cycle of HMM state occurrence, computed from (a) 12 low-yield and (b) 13 high-yield years (days pentad^{-1}). Details are as in Figs. 6 and 8.

with state 2 much more frequent, and state 5 almost absent prior to about 1940. However, it is not clear if this difference in character in the pre-1940 record is real because the record contains a large amount of missing data at several stations, which were filled using a univariate weather generator, as described in section 2. This may have serious implications for the spatial rainfall patterns during the pre-1944 period.

c. Synoptic conditions

To determine the physical significance of the rainfall states, composites of atmospheric circulation variables from NCEP–NCAR reanalysis data (1948–98) are plotted for each state, computed by averaging over the days assigned to each state. Figure 11 shows composites of 850-hPa winds and 500-hPa isobaric vertical velocity, constructed from unfiltered daily data, with the March–August mean subtracted. The vertical motion composite anomalies are similar to composites of the full fields, without the seasonal mean subtracted, while the wind anomalies are superposed on a strong mean subtropical anticyclonic circulation over the Gulf of Mexico and western subtropical Atlantic Ocean. Note that as-
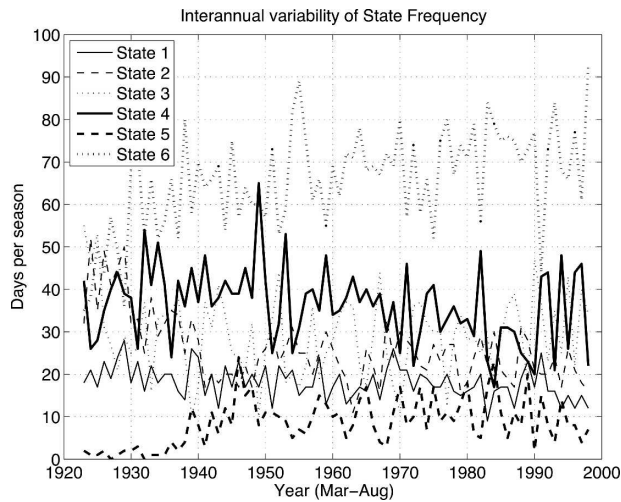
FIG. 10. Interannual variability of state frequency.

cending motion is *negative* in isobaric coordinates ("omega"). All the anomaly composites exhibit synoptic wave patterns in the middle latitudes. These waves dominate the intermittent states with the lowest persistence (states 1, 2, and 5). The strongly preferred transition from state 1 to 6 (Table 1) can be interpreted as an eastward displacement of a ridge from the central to the eastern United States, while the preferred transitions of states 2, 4, and 5 to state 1 reflect the eastward progression of a trough over the eastern United States into the Atlantic.

The vertical motion anomaly fields show large consistency with the rainfall; anomalous descent extends over the southeastern United States during the dry state 6, while ascent anomalies dominate the wet states 3 and 4. The extratropical wave patterns are characterized by large meridional wind anomalies, with southerly anomalies tending to accompany anomalous ascent, and vice versa. Anomalous easterlies predominate in the subtropics of the wet summer states 3 and 4, extending from the Atlantic, indicative of a monsoonal circulation with an intensified subtropical anticyclone to the east. The dry state 6 shows subtropical wind anomalies of the opposite sense, so that there is an effective seasonal reversal of the *anomalous* winds. This is also seen at upper levels (not shown). The abruptness of "onset" (Fig. 7) of the summer rainy season, together with the seasonal reversal of low-level wind anomalies (cf. Zhou and Lau 1998), indicate a monsoonlike climate over north Florida/south Georgia during summer, consistent with Mechoso et al. (2005).

6. Summary and conclusions

We have used a nonhomogeneous hidden Markov model (NHMM) in conjunction with a crop model to

investigate spatial and temporal disaggregation of seasonal rainfall for simulating maize yields over the southeastern United States during the March–August half-year. The observed station-averaged rainfall was used as the single driver of the NHMM, in order to investigate the NHMM's ability to downscale under ideal conditions. The downscaled rainfall simulations were then used to drive a crop model, in order to evaluate the quality of the NHMM's rainfall simulations in terms of crop yields.

When the daily station-averaged rainfall amount was used to drive the NHMM, the simulations were able to recover the interannual variability of station-averaged rainfall amount almost perfectly ($r = 0.97$), providing a regionally averaged consistency check on the NHMM's performance. Station-averaged rainfall frequency and mean daily intensity were less-well captured, however. Interannual variability of rainfall amounts at the individual stations was also less well reproduced, with correlations ranging from 0.52 to 0.72, with an average of 0.60. This provides a measure of the "downscalability" of regional-scale rainfall to the point scale. The difference between the correlation value obtained for station-averaged rainfall ($r = 0.97$) and the mean of the individual station correlations ($r = 0.60$) is consistent with the theoretical analysis of Moron et al. (2006, see their Fig. 5), and would imply an external variance ratio of about 35%, suggesting that the reduction in rainfall correlations at the station scale is due to unpredictable station-scale noise.

Year-to-year differences in 10-day dry-spell counts were found to be relatively poorly simulated by the NHMM ($r = 0.68$ for the regional average), largely because of the sampling uncertainty inherent in the number of 10-day dry spells in any 184-day period (ranging from 1 to 9 in the observed record).

When the input time series was low-pass filtered at 90 days, we found no impact on the simulated seasonal rainfall totals. Thus, subseasonal rainfall anomalies are simply integrated out in the seasonal total. The story is different for 10-day dry spells, where the 90-day low-pass-filtered input lead to very poor simulation of dry-spell counts. This suggests that year-to-year details of weather time-scale variability play an important role in determining the number of simulated dry spells, that is, interannual differences in the latter cannot simply be represented by the geometric distribution given by the first-order Markov chain of the NHMM. Because high-frequency weather variability is unlikely to be predictable at the seasonal scale, this result is further evidence that dry-spell counts are inherently unpredictable. Weather indices based on them should be used with caution.
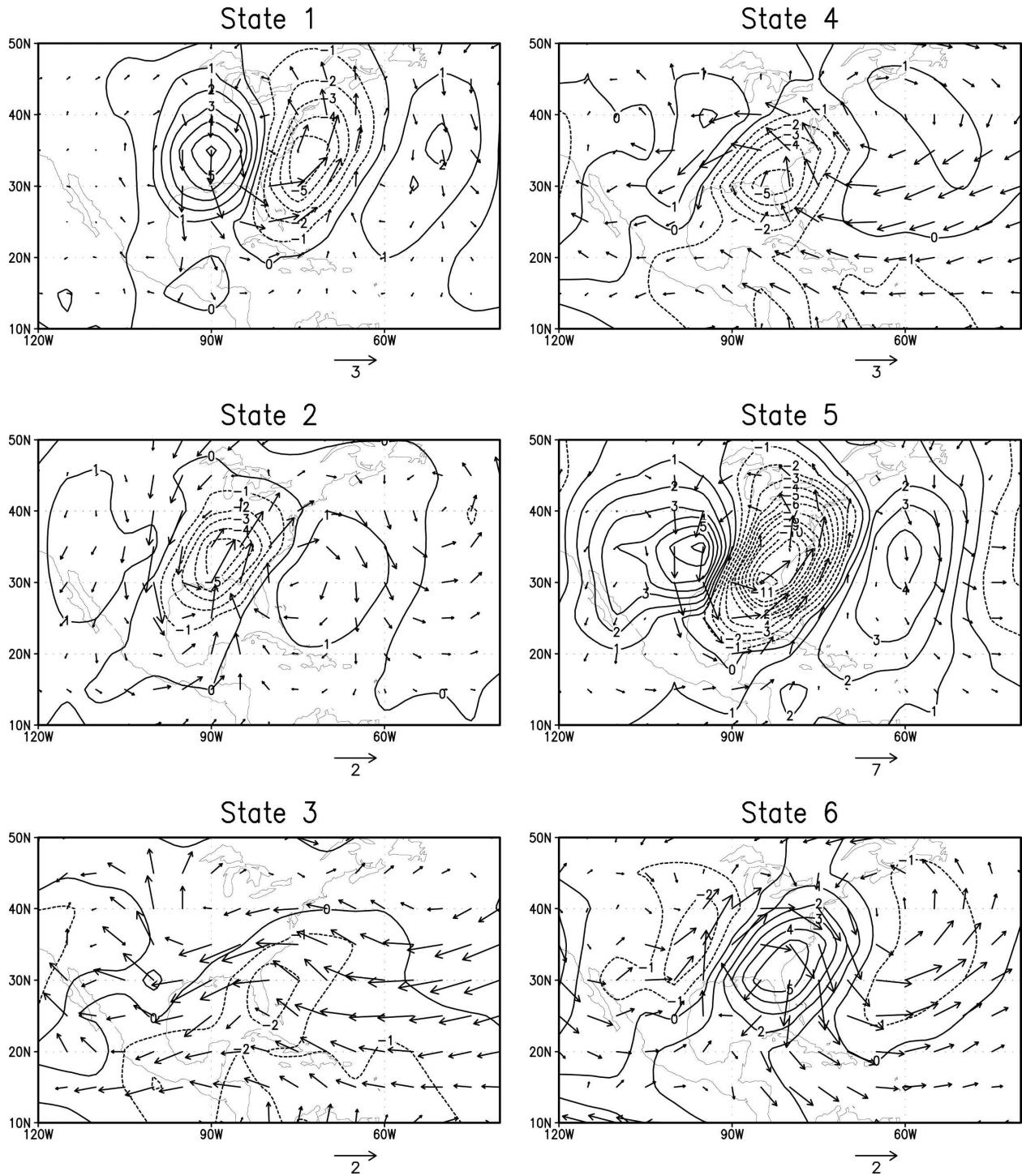
FIG. 11. Anomaly composites with respect to the March–August climatological mean, over the days assigned to each state. Arrows show 850-hPa winds, with contours of 500-hPa omega vertical velocity. Contour interval is 1 Pa s^{-1}, with negative contours dashed (i.e., regions of ascent). Arrow scale is given below each panel (m s^{-1}).

One of the goals of this study was to assess the impact of subseasonal rainfall characteristics on crop yield simulations. The work extends previous single-site studies, such as Hansen and Ines (2005), to a network of sites. We used a "perfect model" approach by comparing with the yield simulated by the crop model when driven by the observed daily rainfall itself. The station-averaged (i.e., regional) yield of the NHMM–

crop model combination was found to be very well simulated when the NHMM is driven by daily data ($r = 0.93$). NHMM-derived yields at the individual stations were found to exhibit interannual correlations of the order of those of NHMM's seasonal rainfall amount, that is, with both reflecting limitations in the spatial disaggregation. The results of Moron et al. (2006) suggest that this error source reflects station-scale variations (largely in rainfall intensity) that are inherently unpredictable at the seasonal scale. Thus, not surprisingly, regionally average yield predictions derived from GCM seasonal forecasts (using an NHMM/crop model, or otherwise) are likely to be more accurate than those made at individual locations.

At the station aggregate level, the NHMM was found to lead to more accurate crop yield simulations than those obtained without downscaling, or with local bias correction of daily rainfall.

Remarkably, the NHMM-derived yield simulations were found to be scarcely degraded when the input series to the NHMM was low-pass filtered, even at 90 days. In particular, the anomaly correlation of simulated yield was found to be much higher than that of 10-day dry-spell counts, because of the temporal integration inherent in the crop model. A substantial degree of seasonality remains in 90-day low-pass-filtered regional rainfall, which has a large impact on simulated yield. In a linear regression sense (i.e., from the squared correlation values), 32% of the simulated station-averaged yield variability was attributable to seasonal rainfall totals (Fig. 4), and an additional 40% to interannual differences in seasonality retained in the 90-day low-pass-filtered rainfall variability (Table 2).

The results of this study demonstrate that regional maize yields over the southeastern United Sates could, in principle, be simulated successfully from 90-day seasonal time-scale regional precipitation alone, provided that low-pass-filtered daily series are used, rather than 3-month averages. These are the time and space scales on which seasonal climate forecasts have been demonstrated to contain skill over certain regions and seasons (Goddard et al. 2003; Gong et al. 2003). Thus, there are good reasons why crop yields should be predictable from these forecasts, as is proving to be the case in several recent studies (e.g., Challinor et al. 2005).

The NHMM without atmospheric predictors, that is, a homogeneous HMM, was shown to yield an informative spatiotemporal diagnostic of the observed rainfall record, in terms of subseasonal, seasonal, interannual, and longer-term variability of six discrete rainfall states. These states were shown to be associated with distinct atmospheric circulation anomalies indicative of a monsoonlike climate over north Florida/south Georgia dur-

ing summer, with two wet monsoonal states, a dry state, and three transient synoptic wave patterns, and an abrupt transition to a prevalence of the wet states near the beginning of June. Delayed monsoon onset was found to characterize low simulated yield years with a much higher prevalence of the dry state between mid-May and mid-June. A gradual long-term drying trend was found to be expressed as an increased prevalence of the dry state relative to the monsoonal wet state (state 4).

REFERENCES

Bellone, E., J. P. Hughes, and P. Guttorp, 2000: A hidden Markov model for downscaling synoptic atmospheric patterns to precipitation amounts. *Climate Res.,* **15,** 1–12.

Challinor, A. J., J. M. Slingo, T. R. Wheeler, and F. J. Doblas-Reyes, 2005: Probabilistic simulations of crop yield over western India using DEMETER seasonal hindcast ensembles. *Tellus,* **57A,** 498–512.

Charles, S. P., B. C. Bates, and N. R. Viney, 2003: Linking atmospheric circulation to daily rainfall patterns across the Murrumbidgee River Basin. *Water Sci. Technol.,* **48,** 233–240.

——, ——, I. N. Smith, and J. P. Hughes, 2004: Statistical downscaling from observed and modelled atmospheric fields. *Hydrol. Processes,* **18,** 1373–1394.

Dempster, A. P., N. M. Laird, and D. B. Rubin, 1977: Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Stat. Soc.,* **39B,** 1–38.

Doorenbos, J., and A. H. Kassam, 1979: Yield response to water. FAO Irrigation and Drainage Paper 33, 193 pp.

Frere, M., and G. F. Popov, 1986: Early agrometeorological crop yield assessment. FAO Plant Production and Protection Paper 73, 144 pp.

Goddard, L., A. G. Barnston, and S. J. Mason, 2003: Evaluation of the IRI's "net assessment" seasonal climate forecasts: 1997–2001. *Bull. Amer. Meteor. Soc.,* **84,** 1761–1781.

Gong, X., A. G. Barnston, and M. N. Ward, 2003: The effect of spatial aggregation on the skill of seasonal precipitation forecasts. *J. Climate,* **16,** 3059–3071.

Hammer, G. L., J. W. Hansen, J. G. Phillips, J. W. Mjelde, H. Hill,

A. Love, and A. Potgieter, 2001: Advances in application of climate prediction in agriculture. *Agric. Syst.,* **70,** 515–553.

Hansen, J. W., and A. V. M. Ines, 2005: Stochastic disaggregation of monthly rainfall data for crop simulation studies. *Agric. For. Meteor.,* **131,** 233–246.

Hughes, J. P., and P. Guttorp, 1994: A class of stochastic models for relating synoptic atmospheric patterns to regional hydrologic phenomena. *Water Resour. Res.,* **30,** 1535–1546.

——, ——, and S. P. Charles, 1999: A non-homogeneous hidden Markov model for precipitation occurrence. *J. Roy. Stat. Soc.,* **48C,** 15–30.

Ines, A. V. M., and J. W. Hansen, 2006: Bias correction of daily GCM rainfall for crop simulation studies. *Agric. For. Meteor.,* **138,** 44–53.

——, A. D. Gupta, and R. Loof, 2002: Application of GIS and crop growth models in estimating water productivity. *Agric. Water Manage.,* **54,** 205–225.

Ingram, K. T., M. C. Roncoli, and P. H. Kirshen, 2002: Opportunities and constraints for farmers of West Africa to use seasonal precipitation forecasts with Burkina Faso as a case study. *Agric. Syst.,* **74,** 331–349.

Jagtap, S. S., J. W. Jones, P. Hildebrand, D. Letson, J. J. O'Brien, G. Podesta, D. Zierden, and F. Zazueta, 2001: Responding to stakeholder's demands for climate information: From research to applications in Florida. *Agric. Syst.,* **74,** 415–430.

Jones, J. W., and Coauthors, 1998: Decision support system for agrotechnology transfer: Dssat v3. *Understanding Options for Agricultural Production,* G. Y. Tsuji, G. Hoogenboom, and P. K. Thorton, Eds., Kluwer Academic, 157–177.

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

Kirshner, S., 2005: Modeling of multivariate time series using hidden Markov models. Ph.D. dissertation, University of California, Irvine, 202 pp.

Mechoso, C. R., A. W. Robertson, C. F. Ropelewski, and A. M. Grimm, 2005: The American monsoon systems: An introduction. WMO Tech. Doc. 1266 (TMRP Rep. 70), 197–206.

Moron, V., A. W. Robertson, and M. N. Ward, 2006: Seasonal predictability and spatial coherence of rainfall characteristics in the tropical setting of Senegal. *Mon. Wea. Rev.,* **134,** 3248–3262.

Palmer, T. N., and D. L. T. Anderson, 1994: The prospects for seasonal forecasting—A review paper. *Quart. J. Roy. Meteor. Soc.,* **120,** 755–793.

Patt, A., P. Suarez, and C. Gwata, 2005: Effects of seasonal climate forecasts and participatory workshops among subsistence farmers in Zimbabwe. *Proc. Natl. Acad. Sci. USA,* **102,** 12 623–12 628.

Pickering, N. B., J. W. Hansen, J. W. Jones, C. M. Wells, and V. Chan, 1994: WeatherMan—A utility for managing and generating weather data. *Agron. J.,* **86,** 332–337.

Priestley, C. H. B., and R. J. Taylor, 1972: On the assessment of surface heat flux and evaporation using large-scale parameters. *Mon. Wea. Rev.,* **100,** 81–92.

Rabiner, L. R., 1989: A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. IEEE,* **77,** 257–286.

Richardson, C. W., and D. A. Wright, 1984: A model for generating daily weather variables. Tech. Rep. ARS-8, U.S. Dept. of Agriculture, Agriculture Research Service, 83 pp.

Ritchie, J. T., 1998: Soil water balance and plant water stress. *Understanding Options for Agricultural Production,* G. Y. Tsuji, G. Hoogenboom, and P. K. Thorton, Eds., Kluwer Academic, 41–54.

——, U. Sigh, D. C. Godwin, and W. T. Bowen, 1998: Cereal growth, development and yield. *Understanding Options for Agricultural Production,* G. Y. Tsuji, G. Hoogenboom, and P. K. Thorton, Eds., Kluwer Academic, 79–98.

Robertson, A. W., S. Kirshner, and P. J. Smyth, 2003: Hidden Markov models for modeling daily rainfall occurrence over Brazil. Tech. Rep. ICS-TR 03-27, Information and Computer Science, University of California, Irvine, CA, 36 pp.

——, ——, and ——, 2004: Downscaling of daily rainfall occurrence over northeast Brazil using a hidden Markov model. *J. Climate,* **17,** 4407–4424.

——, ——, ——, S. P. Charles, and B. C. Bates, 2006: Subseasonal-to-interdecadal variability of the Australian monsoon over north Queensland. *Quart. J. Roy. Meteor. Soc.,* **132,** 519–542.

Vaux, H. J., and W. O. Pruitt, 1983: Crop-water production functions. *Adv. Irrig.,* **2,** 72–79.

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

Wilks, D. S., 1999: Interannual variability and extreme-value characteristics of several stochastic daily precipitation models. *Agric. For. Meteor.,* **93,** 153–169.

——, 2002: Realizations of daily weather in forecast seasonal climate. *J. Hydrometeor.,* **3,** 195–207.

——, and R. L. Wilby, 1999: The weather generation game: A review of stochastic weather models. *Prog. Phys. Geogr.,* **23,** 329–357.

Zhou, J., and K.-M. Lau, 1998: Does a monsoon climate exist over South America? *J. Climate,* **11,** 1020–1040.