Paper:
Challinor, AJ, Wheeler, TR, Slingo, JM, Craufurd, P and Grimes, D (2005) *Simulation of crop yields using ERA-40: limits to skill and nonstationarity in weather–yield relationships.* Journal of Applied Meteorology, 44 (4). 516 – 531

http://dx.doi.org/10.1175/JAM2212.1

Simulation of Crop Yields Using ERA-40: Limits to Skill and Nonstationarity in Weather–Yield Relationships

A. J. CHALLINOR

Centre for Global Atmospheric Modelling, Department of Meteorology, and Department of Agriculture, The University of Reading, Reading, United Kingdom

T. R. WHEELER

Department of Agriculture, The University of Reading, Reading, United Kingdom

J. M. SLINGO

Centre for Global Atmospheric Modelling, Department of Meteorology, The University of Reading, Reading, United Kingdom

P. Q. CRAUFURD

Department of Agriculture, The University of Reading, Reading, United Kingdom

D. I. F. GRIMES

Department of Meteorology, The University of Reading, Reading, United Kingdom

(Manuscript received 26 December 2003, in final form 25 September 2004)

ABSTRACT

Reanalysis data provide an excellent test bed for impacts prediction systems, because they represent an upper limit on the skill of climate models. Indian groundnut (*Arachis hypogaea* L.) yields have been simulated using the General Large-Area Model (GLAM) for annual crops and the European Centre for Medium-Range Weather Forecasts (ECMWF) 40-yr reanalysis (ERA-40). The ability of ERA-40 to represent the Indian summer monsoon has been examined. The ability of GLAM, when driven with daily ERA-40 data, to model both observed yields and observed relationships between subseasonal weather and yield has been assessed. Mean yields were simulated well across much of India. Correlations between observed and modeled yields, where these are significant, are comparable to correlations between observed yields and ERA-40 rainfall. Uncertainties due to the input planting window, crop duration, and weather data have been examined. A reduction in the root-mean-square error of simulated yields was achieved by applying bias correction techniques to the precipitation. The stability of the relationship between weather and yield over time has been examined. Weather–yield correlations vary on decadal time scales, and this has direct implications for the accuracy of yield simulations. Analysis of the skewness of both detrended yields and precipitation suggest that nonclimatic factors are partly responsible for this nonstationarity. Evidence from other studies, including data on cereal and pulse yields, indicates that this result is not particular to groundnut yield. The detection and modeling of nonstationary weather–yield relationships emerges from this study as an important part of the process of understanding and predicting the impacts of climate variability and change on crop yields.

1. Introduction

Research efforts over recent years have taken advantage of increases in computational power and climate/weather model skill in order to couple climate and impact models. Studies span a range of impacts (crop productivity, e.g., Mearns et al. 1999, 2001; health, e.g., Hoshen et al. 2003; hydrology, e.g., Davis et al. 2003). There are many inherent uncertainties in such studies (see below), and, in exploring the issue of the predictability of impacts, it is important that these are quantified as accurately as possible. One way of exploring the uncertainties that come from a climate model [or gen-

Corresponding author address: Dr. Andrew Challinor, GCAM, Department of Meteorology, The University of Reading, P.O. Box 243, Earley Gate, Reading RG6 6BB, United Kingdom.
E-mail: ajc@met.rdg.ac.uk

eral circulation model (GCM)] is to use reanalysis data (Bengtsson and Shukla 1988; Kalnay et al. 1996; Gibson et al. 1996, 1997). Following the methodology proposed by Challinor et al. (2003) the present study uses reanalysis data with a crop model in order to explore the potential predictability of crop yields and the associated uncertainties.

Reanalysis data are GCM output with weather data from many sources assimilated into the climate simulation. They constitute the most accurate complete description of the weather at the resolution of the GCM, so that when used to drive an impacts model, they provide an upper limit on the accuracy of the combined (climate plus impacts) modeling system (Challinor et al. 2003). While there are numerous published studies that use reanalysis to study climate (e.g., Annamalai et al. 1999; Betts et al. 2003), there are few studies using reanalysis with impacts models (e.g., Palmer et al. 2004; Hoshen et al. 2003). In this study the European Centre for Medium-Range Weather Forecasts (ECMWF) 40-yr reanalysis (ERA-40; http://www.ecmwf.int/research/era/) is used to drive the General Large-Area Model (GLAM) for annual crops (Challinor et al. 2004b, hereinafter GLAM2004), which was designed specifically for use with GCMs and regional climate models.

Direct use of GCM output is not the only method of using GCM simulations with an impacts model. An alternative method summarizes daily GCM output in the form of monthly or seasonal climate; this information is then used as an input to a synthetic weather generator (e.g., Wilks 2002; Wilks and Wilby 1999; Semenov and Brooks 1999; Corte-Real et al. 1999). This method can be used to downscale climate information. It can also be used to generate time series for climate change scenarios based on changes, between the current and future climate, in the mean and the variability of weather (Semenov and Barrow 1997). This method has the advantage of not relying on the correct simulation by the GCM of the basic mean state. It has the disadvantage of relying on a set of assumptions, embedded in the weather generator, regarding the relationship between mean climate and weather and between different weather variables. Such weather statistics may or may not remain constant as the climate changes. The methods used for this study make no such assumptions. Successful results using these methods in current climates would suggest that an extension to use direct GCM output in the study of climate change is plausible.

In addition to the input weather data, there are many sources of error that contribute to disagreement between observed and modeled yields (for an excellent overview, see Hansen and Jones 2000). No crop model formulation is a complete description of the crop and its interaction with the environment. Model parameterizations are simplifications of crop processes that inevitably result in some error. Any yield data will also have associated error, and the necessary separation of the time series into underlying technology trends (defined as a monotonic increase in yield over time due to nonclimatic factors, such as improved yield varieties and an increased use of fertilizer) and interannual variability also adds uncertainty (e.g., Just and Weninger 1999; Yu et al. 2001). Input management data, such as planting date, plant population density, and crop variety, have associated uncertainties that will impact the ability of the model to reproduce reality (i.e., model skill). Aggregation error due to the large spatial scale on which the model is run will also contribute to errors in simulated yield. Finally, the input soil and crop data will have inaccuracies. All of these sources of error will limit the accuracy with which yields can be simulated.

This study restricts its analysis of error sources principally to the weather input data and yield data. The impact of varying soil hydrological properties is discussed in GLAM2004. Here, three principal topics are explored: (i) the ability of ERA-40 to simulate Indian monsoon rainfall (section 3a), (ii) the relationship between monthly mean weather and yield (both simulated and observed; section 3b), and (iii) the relationship between observed and simulated yield (section 3c). This study aims to assess the capability of GCM output (specifically, reanalysis data), together with a large-area crop model, to simulate yield. Through bias correction of the reanalysis rainfall (toward gridded observed values), this study also seeks to assess any potential increase in the accuracy of yield prediction in response to improvements in GCM skill.

The nonlinear response of crops to climate means that changes over time in the relationship between yield and climate (i.e., nonstationarity) may be observed. For example, changes in the fraction of crop under irrigation, or in cultivar-specific properties, such as sensitivity to water stress or pests, could change the response of a crop to the climate. Assumptions regarding the yield technology trend and the planting window (the period of time from which a planting day is chosen) will have direct implications for the observation of any changes, as will any changes over time in the accuracy of the data used. However, trends in yield may also be attributed to trends in climate (Pathak et al. 2003). As well as the impact of seasonal mean climate, the impact of subseasonal weather variability on crop yields can be significant (Gadgil et al. 2002; Hansen and Jones 2000). Our understanding of the crop–climate system will depend upon the extent to which climatic and nonclimatic effects can be separated. The prospects for yield prediction, particularly in the context of climate change, rely, in part, on this understanding. It is for this reason that GLAM is used in this study; it is process based, rather than empirical, and does not rely upon the assumption of stationary yield–weather statistics. It is also designed to simulate the impact of weather on yields, and it can operate on the spatial scales of the reanalysis data. Section 4 contains further discussion on the issue of nonstationarity.

2. Modeling and analysis methods

a. The crop model

The crop model used in this study was the General Large-Area Model for annual crops. It is fully described in GLAM2004. It is a model of intermediate complexity, with parameterizations that seek to avoid a large crop model input data requirement, while capturing the interactions between climate and crop. It is designed to run on any spatial scale at which a relationship exists between crop and climate (Challinor et al. 2003), and it has been run successfully over India at a 2.5° resolution using observed gridded data (GLAM2004) and GCM ensemble output (Challinor et al. 2004a).

GLAM is a process-based crop model with a daily time step, allowing it to resolve the impacts of subseasonal variability in weather. It has a soil water balance with 25 layers, which simulates evaporation, transpiration, and drainage. Transpirative demand is simulated according to Priestly and Taylor (1972), and the supply of water according to root water uptake. Roots grow with a constant extraction-front velocity and a profile that is linearly related to leaf area index (LAI). LAI evolves using a constant maximum rate of change of LAI that is modified by a soil water stress factor (the ratio of water supply to transpirative demand). Separate simulation of biomass accumulation, by the use of transpiration efficiency [an input parameter, normalized by the vapor pressure deficit (VPD) and based on observations], allows specific leaf area (SLA; the mass of leaf per unit area of leaf) to be used as an internal consistency check: leaf area and leaf mass can be derived independently of each other and can be used to calculate values of SLA that can be compared to the typical observed values. The objective of the model is to reproduce the impact of weather on observed crop yield. This aim leads to two particular model characteristics: first, complexity at a level that is far removed from the yield-determining processes is omitted (see Sinclair and Seligman 2000). In general, simple parameterizations are favored over more complex methods. Second, of the impacts on yield due to factors other than weather (biotic stresses such as pests, diseases, weeds, and abiotic management factors), only two are modeled explicitly—(i) planting date, which occurs on the first day within the prescribed planting window for which the available soil moisture is over 50% of the maximum, or at the end of the window if no such day is found; and (ii) soil hydrological properties, which are derived from the prescribed saturated upper limit and drained lower limit of the soil. GLAM uses a single parameter to account for the yield gap (the reduction in yield from climatically determined attainable values to actual values due to the impact of biotic stresses and suboptimal management). While this is a simplification, it allows the model to focus on the crop–climate interaction.

b. Weather and yield data

Weather inputs for the crop model are daily mean values of vapor pressure and mean temperature, and daily total rainfall and solar irradiance. These inputs came from ERA-40 (http://www.ecmwf.int/research/era/). ERA-40 is a global reanalysis at T159 resolution, which equates approximately to a 1.1° × 1.1° square grid. The data cover the period from September 1957 to August 2002. Solar radiation and precipitation are accumulated over 0000–2400 coordinated universal time (UTC) and are input to the crop model without modification. Daily mean temperature is calculated as the average of the ERA-40 six-hourly temperature. Six-houly dewpoint temperature is used to calculate daylight mean vapor pressure by averaging over data between sunrise and sunset only. This method resulted in the use of at least two values for each daily value.

Precipitation is a key determinant of rain-fed groundnut yields (e.g., Camberlin and Diop 1999). It is also a difficult variable for climate models to reproduce accurately. Hence, this variable merits particular attention. During the ERA-40 data period there are differing amounts of observational data used in the analysis: ERA-40 can be characterized as having three separate periods, each with more assimilated data than the last as more satellite data became available. One such period starts in 1973, and another in 1987. The introduction of new data has resulted in differences in the hydrological cycle between the three identified periods (Andersson et al. 2005). These differences are most pronounced over the ocean (Troccoli and Kållberg 2004), although changes also occur over land. The results of Andersson et al. (2005) suggest that the largest change in precipitation over land is likely to be around 1973. This issue is revisited briefly in section 3a, where the ERA-40 dataset is compared with the dataset of the Indian Institute of Tropical Meteorology (IITM; http://www.tropmet.res.in/). Rather than regrid either dataset and degrade the information, ERA-40 grid cells were assigned uniquely to an IITM grid cell according to the region of the greatest overlap. Both data sources have associated errors, but the IITM dataset is based solely on observations, and is expected to provide a better estimate of rainfall over the area covered.

The humidity analysis of the ECMWF causes overprediction of the tropical precipitation used in ERA-40. Rapid adjustment of the Hadley circulation, and adjustment of the precipitation to observed values, take place in the first 6–12 h of each forecast. This “spindown” problem can result in daily precipitation being higher than that observed (for a detailed analysis see Anders-

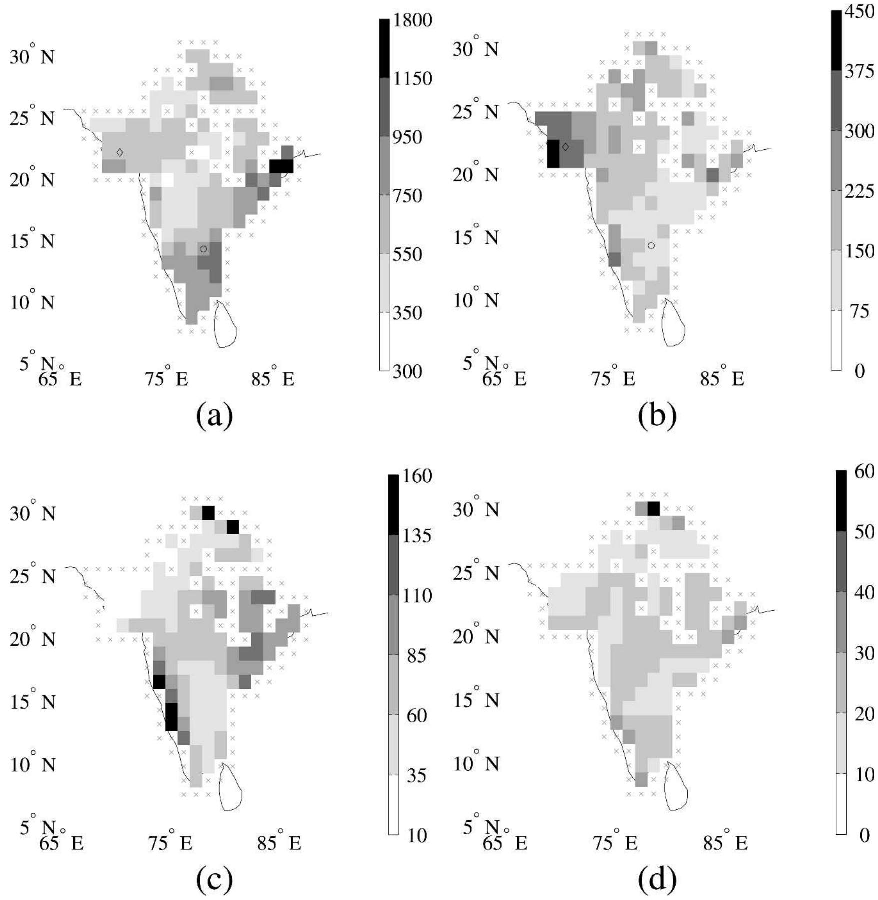

Fig. 1. Observed (a) mean and (b) standard deviation of (linearly) detrended groundnut yields (kg ha^{-1}) in India, for the period of 1966–89, on the ERA-40 grid. For clarity, missing data points that are adjacent to plotted data are marked (x). Missing data imply either no data or grid cells with a mean (1966–89) area of cultivation under 15 kha. Also shown are the (c) mean and (d) standard deviation of ERA-40 JJAS precipitation (cm) over the same period and for the same grid cells. Two grid cells used in later analysis are marked in (a) and (b): GJ with a diamond and AP with a circle.

son et al. 2005). However, comparisons with the IITM data show that the ERA-40 high precipitation bias is not generally a problem for the Indian summer monsoon over this period (section 3a).

The yield data for calibration and evaluation of the model came from the database of agricultural returns compiled by the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) in Patancheru, India. Yields tend to increase over time, as improved varieties and management methods are employed. This trend will not be predicted by the model, and some assumptions must be made in order to remove this trend. In this study we have taken the common step of assuming a linear technology trend. Hence, all yields shown and referred to, unless otherwise stated, have been linearly detrended. The results of this study may depend to some degree on how the technology trend is modeled; it is, therefore, worth considering other options. Because explicit modeling of the underlying technological processes would require a knowledge of crop variety and management, two practical options for modeling the yield technology trend remain—(i) fitting a given trend parameterization (e.g., polynomial) using least squares (Just and Weninger 1999), and (ii) assuming the form of the probability density function (PDF) of detrended yield, and of the form of the technology trend, with subsequent determination of the trend using maximum likelihood (Moss and Shonkwiler 1993; Ramirez et al. 2003). Because the form of the PDF of yield cannot be determined a priori (e.g., Atwood et al. 2003), we use the first of these methods to determine alternative technology trends; linear ($y = a + bt$), quadratic ($y = a + bt + ct^2$), and cubic ($y = a + bt + ct^2 + dt^3$) trends were fitted to either the whole time series or piecewise to each half of the time series (section 3b).

The weather and yield datasets overlap for the period of 1966–89 for 132 grid cells across India. Hence, this is the spatiotemporal domain used for this study. Figure 1 shows the mean and standard deviation of ERA-40 rainfall, as well as the same statistics for the yield data.

c. Crop model calibration

GLAM is used here to simulate groundnut (*Arachis hypogaea* L.) yield. Soil hydrological properties were derived from FAO/UNESCO (1974) and planting windows were from Reddy (1988), following GLAM2004. Control planting windows are of a 30-day duration. The crop-specific parameters for the control run in this study came directly from GLAM2004. In that study, these parameter values minimized root-mean-square error (rmse) in yield for simulations of groundnut yield across India. These optimal parameters all fall within the range of the observed values. The parameters are not genotype specific, although many are broadly based on the TMV-2 cultivar, which is commonly grown in India.

A sensitivity analysis using the ERA-40 data showed that rmse in crop yield could not be systematically reduced across all model grid cells by changing any parameters in the crop model from the GLAM2004 values. This implies that this parameter set is suitable for use with ERA-40 data. Note, however, that calibration of each grid cell individually would reduce rmse in yield, because it would result in spatially variable optimal parameters. This spatial variability in optimal parameters is in contrast to the results using the 2.5° grid in GLAM2004, where the optimal parameter set was relatively constant over space. Likely reasons for the spatial variability of optimal parameters observed here include the finer spatial and temporal resolution of the ERA-40 data. For the current study, a global (India-wide) groundnut parameterization was retained because there is insufficient data on the spatiotemporal pattern of groundnut varieties that are grown to justify local parameterization.

GLAM2004 found that the impact of the inclusion of the irrigated fraction of groundnut crop on the skill of the simulations was small. The irrigated fraction has a mean, over the period of 1966–89, of less than 15% in 48 of the 82 grid cells where there are data (ICRISAT; see, e.g., GLAM2004), and a mean of less than 30% in 64 grid cells. Two cells [at Andhra Pradesh (AP) and Gujarat (GJ)] are considered in particular detail in the analyses that are presented (see Fig. 1 for their location): irrigation in grid cell AP accounts for an average of 12% of the area under cultivation. Because irrigated yields tend to be a factor of 0.4–1 times greater than rain-fed yields in this region (Virmani and Shurpali 1999), irrigation is not expected to significantly alter the results. Mean irrigation in GJ is even less significant, at 1%. Hence, the irrigated fraction was not simulated for this study.

Model calibration follows the same procedure as GLAM2004. The yield gap parameter (YGP) is determined for each individual grid cell by cross validation, with the period of 1966–77 being used to determine the 1978–89 YGP, and vice versa. Calibrated values of YGP are values that minimize the rmse between the observed and simulated crop yields.

d. Crop growth simulations

In addition to the control run that is described above, a number of other simulations were carried out. Key parameters relating to the crop, its management, and the input weather data were varied. To examine the impact of crop duration in the southern peninsula the four genetic coefficients, which determine the thermal time for each of the four growth stages in GLAM, were altered. Each was increased by an equal fraction such that the mean duration from planting to harvest increased from 90–100 days (control run) to 120–130 days. These increased values allow the simulation of an extended duration crop (simulation E1).

To examine the impact of the choice of planting window, a delayed window was used in grid cells in two key regions: Gujarat (where the control planting window of 30 days starts on 15 June, and the delayed planting window P1 is 1–10 August) and Andhra Pradesh (where the control planting window is 1–30 June, and the delayed planting window P2 is 1–30 July). These two regions have the largest groundnut production over the study period (Challinor et al. 2003).

Because ERA-40 precipitation is not expected to be as accurate as the observed values, two bias corrections of the ERA-40 precipitation were carried out: simulation B1 corrects the mean climatology (it shifts the ERA-40 June–September rainfall totals by a constant factor for each location, such that the 1966–89 mean rainfall agrees with the IITM data), and simulation B2 corrects the interannual variability (it shifts the seasonal total by a time-varying amount such that it agrees with the IITM data). These simulations allow the impact of precipitation inaccuracies to be assessed. Simulations with input precipitation bias correction using the extended duration crop were also carried out (E1B1). Because YGP is calibrated on rmse in yields, it will have some component of bias correction, which may include input weather data bias (Challinor et al. 2004a). Because all of the simulations (control, P1, P2, B1, B2, E1, E1B1) differ in terms of the exact input weather time series, each simulation has its own calibrated spatial distribution of YGP.

e. Analysis of simulated yields

The relationship between monthly mean weather during the growing season and the observed yield (i.e., the observed relationship between yield and weather) may be different than the relationship between the monthly mean weather and the simulated yield (i.e., the modeled relationship between yield and weather). Comparisons between modeled and observed yield–

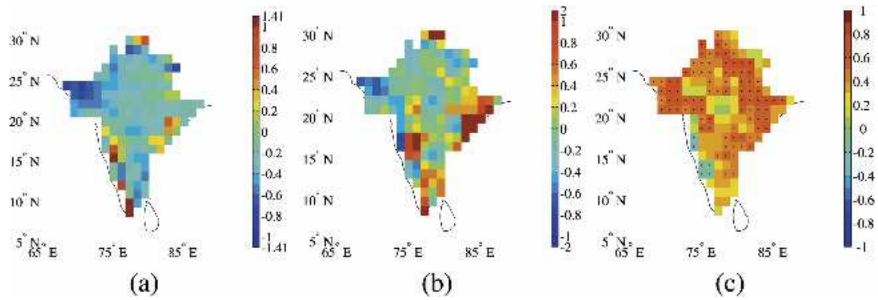

FIG. 2. Fractional difference between ERA-40 and IITM data in (a) mean and (b) standard deviation of JJAS precipitation. (c) The correlation in JJAS precipitation between the two datasets. Statistically significant correlations ($p < 0.05$) are marked with a thick dot.

weather relationships enable the assessment of the crop model's ability to correctly simulate the response of the crop to weather. Section 3b compares the simulated and observed weather–yield relationships, using the ERA-40 weather variables. Section 3c goes on to compare the observed and simulated yields directly, using correlation coefficients and rmse of the former relative to the latter.

3. Results

a. *Evaluation of the Indian summer monsoon in ERA-40*

To assess the potential impact of the introduction of satellite data into the reanalysis, the June–July–August–September (JJAS) precipitation from the IITM and ERA-40 were compared by averaging each across two periods—1966–72 and 1974–89 (see section 2b). The difference between the mean ERA-40 and mean IITM values (d) is an indication of systematic bias. A change in the sign of d is an indication of a potentially serious drift in the ERA-40 data; such a drift represents a change in the bias of the ERA-40 precipitation (note, however, that bias correction B2 does correct for this). Over 80% of the 159 grid cells over India (which include the 132 for which there are yield data) have the same sign of d for the two periods. A Kolmogorov–Smirnov test on the two distributions of total JJAS precipitation shows that 80% of the grid cells for which there are yield data cannot be said to have different distributions for the two periods (at a significance level of 5%). Hence, the impact of the changes in the ERA-40 hydrological cycle on the simulation of the Indian summer monsoon appears not to seriously question the validity of the precipitation data.

Examining now the study period as a whole, Fig. 2 compares the seasonal total ERA-40 monsoon rainfall with that of the IITM. The mean June–September ERA-40 rainfall is mostly lower than IITM values, and many standard deviations are lower also. ERA-40 overestimates the frequency of light rains, and underestimates heavy rains (Fig. 3). Overall, the number of rain days in ERA-40 is greater than that of the IITM data. Mean monsoon onset is in broad agreement over much of India, although the southern peninsula shows a tendency for earlier onset in ERA-40 (Fig. 3c). Onset here is defined after Zhang et al. (2002): a criterion of rainfall >5 mm day^{-1} is fulfilled first for 5 days, then consecutively for 10 out of 20 days.

The subseasonal variability is compared for two ERA-40 grid cells (in regions where much of India's

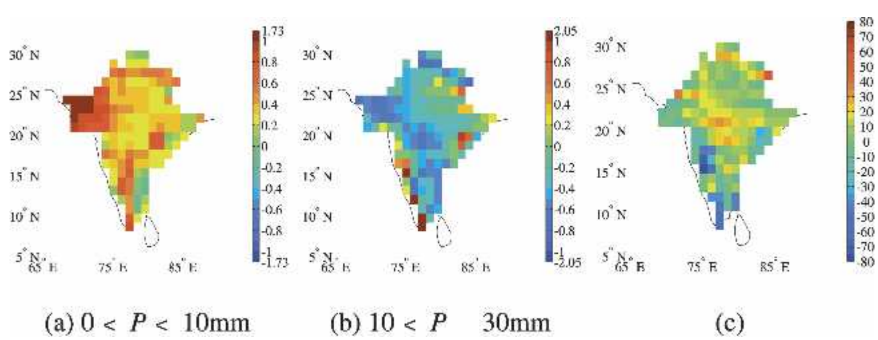

FIG. 3. Fractional difference between ERA-40 and IITM in the mean (1966–89) number of JJAS days with precipitation in the ranges shown. (c) The difference in days (ERA-40 − IITM) between the mean monsoon onset (see text for definition) over the period of 1966–89. Five missing data points have less than 5 yr in which the onset criterion is satisfied.

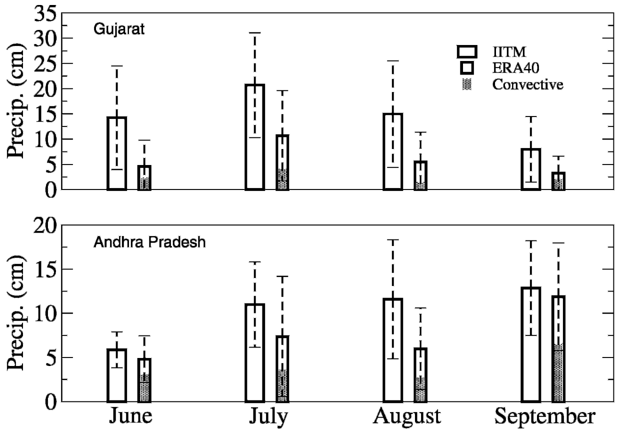

FIG. 4. The ERA-40 and IITM seasonal cycle of precipitation for two grid cells in India. Bars show the 1966–89 mean, and whiskers show one standard deviation. The Gujarat and Andhra Pradesh grid cells (GJ and AP, respectively) are marked on Fig. 1a. The ERA-40 convective precipitation is also shown; the remaining precipitation is large scale.

groundnut is grown) in Fig. 4. For one of these cells (GJ), the seasonal cycle is well simulated, despite the mean values that are underestimated. In cell AP, June and September totals are well simulated, but July and August are deficient; thus, the seasonal cycle is poorly simulated. These differences in the subseasonal variability have consequences for the model simulations in these regions (section 3c).

The subseasonal variability of the precipitation also changes over time, as the following example shows. The seasonal cycle of rainfall in western Gujarat (the 13 westernmost grid cells in Fig. 3c) was tested for changes across the periods of 1966–77 and 1978–89, using the Wilcoxon rank sum test (Wilcoxon 1945). Five grid cells showed a change, at a 5% significance level, in the 14-day-summed values of precipitation that were centered on either 17 August (which showed increased precipitation) or 31 August (which showed decreased precipitation). All of the IITM grid cells for western Gujarat showed one of these changes, which was significant at 5%.

b. Relationships between weather and yield

Figure 5 shows that modeled yields are more rainfall dependent than the observed yields. Given that rainfall is one of the few input variables in the model, but one of many in reality, this is not surprising. The same pattern of overcorrelation can be seen between simulated yield and VPD (Fig. 6), although some of the lower correlations in the far south are reproduced by the model. Correlations between yield and VPD do not imply causality because VPD and rainfall are highly correlated (Fig. 6c). However, looking at the differences between correlations in the first and second half of the time series (Table 1) reveals that for GJ, observed correlations with VPD can be more robust than correlations with precipitation. Note that for both GJ and AP the changes in correlation are not related to changes in the calibration parameter YGP, because YGP is constant over time in both cases.

Yields in the GJ region have been shown to have a stronger climate signal (defined as the strength of the detectable impact of climate on yield, as measured by a correlation coefficient, e.g.) than yield in the AP region (GLAM2004; Challinor et al. 2003), and this is reflected in correlations between the June–September meteorological variables and yields (Table 1). All of the observed statistically significant correlations over the 1966–89 time period are reproduced as statistically significant correlations by the model. Four of the six observed significant correlations for the two 12-yr subperiods are reproduced as statistically significant correla-

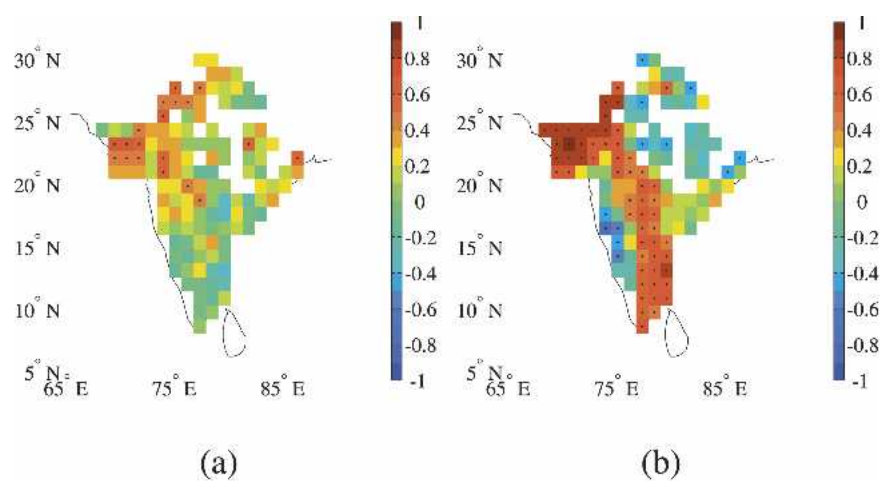

FIG. 5. Correlations between ERA-40 May–Nov precipitation and (a) observed and (b) modeled yields. Significant correlations ($p < 0.05$) are marked with a dot.

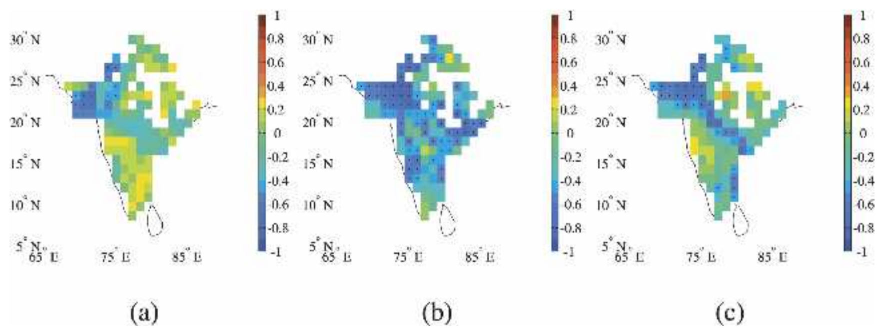

FIG. 6. Correlations between ERA-40 Jul VPD and (a) observed and (b) modeled yields. (c) Also shown is the correlation between seasonal precipitation and mean VPD, over the growing season defined by the control run. Significant correlations ($p < 0.05$) are marked with a dot.

tions by the model. However, differences between the first and second half of the time series are not well represented. This issue is revisited below.

Correlating, in turn, monthly mean values of precipitation, the VPD and net radiation from ERA-40 with yield in the Gujarat region (Fig. 7) show that the simulations pick out correctly the months during which weather has the most impact on yield (July, August, and September). The crop model overpredicts precipitation and radiation correlations early in the season. This suggests that a later planting date may be used in the simulations for this region, because reduced early season correlations would result. Using a planting window of 1–10 August (simulation P1) reduces the June and July weather–yield correlations (Fig. 7). The same simulation also produces a change in the correlation between rainfall and yield over the two halves of the time series (0.46 for 1966–77 and 0.65 for 1978–89), which agrees more closely with the observed change (c.f. Table 1). A similar result is obtained with bias-corrected input weather data in simulation B2 (Fig. 7); B2 has a correlation between yield and rainfall for 1966–77 of 0.43, and the value for 1978–89 is 0.65.

The observation of a change in the relationship between weather and yield in India at this time (1977–78) is not unique to this study, and has also been observed by Parthasarathy et al. (1992), Challinor et al. (2003), Selvaraju (2003), and Kulkarni and Pandit (1988). Further evidence of such changes in weather–yield relationships can be found by examining the correlation of all-India groundnut yield (produced by the Food and Agriculture Organization of the United Nations) with El Niño region sea surface temperatures (SSTs). For example, using the Niño-1 and -2 regions (0°–10°S, 90°–80°W) September SSTs from the Climate Prediction Center (http://www.cpc.noaa.gov), and linearly detrending the FAO yield data between 1966 and 1989, the correlation for the period of 1966–77 is −0.78, and the 1978–89 value is −0.26. A similar analysis with all-India rainfall data (Parthasarathy et al. 1995) reveals a change in correlation between rainfall and yield, across the same two time periods, from 0.96 to 0.67.

The skewness of the yield and rainfall time series also suggests a change in the nature of the relationship between weather and yield: the skewness of yield shows more variability over both space and time than the skewness in rainfall (Fig. 8). The change in yield skewness suggests a systematic shift in the PDF of yields. That the rainfall does not show a change in skewness over time, while the yield does, suggests that the shift in the PDF has a nonclimatic component (or, at least, a component that is unrelated to rainfall). Changes in irrigation levels or in the form of the technology trend are two of the numerous possible nonclimatic factors that could contribute to the change in skewness.

To determine whether the observed nonstationarity

TABLE 1. Correlations between observed and modeled (control run) yields and ERA-40 Jun–Sep values of (i) total precipitation, (ii) mean VPD, and (iii) net radiation for three time periods and two grid cells. The grid cells correspond to those used in Fig. 4. Statistically significant correlations ($p < 0.05$) are shown in boldface.

		Precipitation		VPD		Net radiation	
Cell	Period	Obs	Modeled	Obs	Modeled	Obs	Modeled
AP	1966–89	0.15	**0.62**	0.22	−0.19	−0.08	**−0.48**
AP	1966–77	−0.13	**0.79**	0.23	−0.06	−0.09	−0.51
AP	1978–89	**0.58**	0.38	0.22	−0.42	−0.07	**−0.64**
GJ	1966–89	**0.49**	**0.83**	**−0.69**	**−0.44**	**−0.70**	**−0.84**
GJ	1966–77	0.41	**0.94**	**−0.73**	−0.33	**−0.75**	**−0.80**
GJ	1978–89	**0.67**	**0.78**	**−0.72**	**−0.88**	**−0.80**	**−0.89**

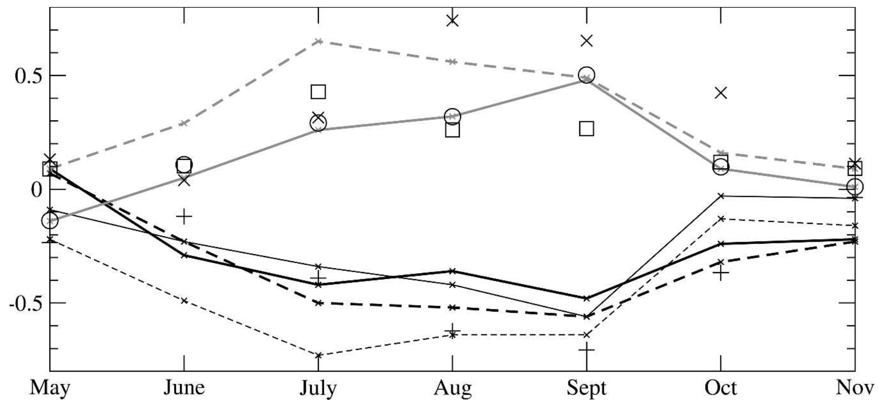

FIG. 7. Correlations, averaged over western Gujarat (the 13 westernmost grid points), between mean monthly weather variables and observed (solid lines) and simulated (dashed lines) yields. The mean monthly weather variables are (i) net radiation (thin black lines), (ii) VPD (thick black lines), and (iii) precipitation (thick gray lines). Crosses show observed correlations between precipitation and yield for a delayed planting window (simulation P1). Pluses mark the corresponding correlation with net radiation. Circles show the correlations between precipitation and observed yield for bias-corrected simulation B2, and squares show the corresponding correlations with simulated yields.

was due solely to the chosen form of the yield technology trend, a number of other forms of trend were considered. Linear, quadratic, and cubic trends, together with the yield data and the yield residuals, are shown in Fig. 9. Running means over 5, 7, or 9 yr were also used to detrend. The analysis was applied to the GJ and AP grid cells, and to all-India yields used in the SST and rainfall correlation analysis above. The character of the results (i.e., the changes in correlation) was not altered in any case. Most of the computed trends in GJ and AP were not statistically significant, and so the all-India case is chosen to illustrate this (Table 2).

An analysis of the ratio of yield to rainfall (Fig. 10) shows that over Gujarat and much of the southern peninsula, the apparent water-use efficiency (AWUE) is particularly high—in some places unrealistically so. To place an approximate upper limit on AWUE, we take an average-to-high yield for India (1000 kg ha^{-1}), and an average-to-low value of water use (400 mm for a groundnut crop cycle; Sivakumar and Sarma 1986), giving a value of 30 kg ha^{-1} cm^{-1}. Frequent occurrence of higher values could be indicative of errors in the rainfall, which is underestimated by ERA-40 (Fig. 2). In the southern peninsula, however, AWUE values are high without frequently being unrealistic. This may be due to differences in the seasonal cycle. In the AP grid cell (see Fig. 1), for instance, an average of 56% of the rainfall between (simulated) planting and harvest falls during pod filling, compared to less than 30% for three grid boxes in Gujarat. A similar value can be found in the GLAM2004 simulations (using IITM rainfall) for the grid cell corresponding to AP (59%). A higher percentage of rainfall during pod filling will contribute to higher AWUE.

c. Relationship between observed and simulated yields

The control simulation output is shown in Fig. 11. Mean yields across many regions are well simulated, with the notable exceptions being much of the southern peninsula and Gujarat. The GLAM2004 simulations, which used observed gridded data, showed closer agreement with observations for both the southern peninsula and Gujarat. For the control simulation, 33% of the grid cells have mean yields within 5% of observed values, 53% are within 10% of observations, and 83% are within 50% of observations. The fact that mean yields have been well simulated across a range of environments is not due simply to a tuning of YGP to an appropriate value: narrow ranges of YGP for which mean yields are simulated accurately show diversity in the June–September precipitation characteristics (Fig. 12). Mean simulated LAIs for the control simulation across India are mostly less than one (not shown), which is realistic for an Indian groundnut crop (Kakani 2001).

Standard deviations in yield (σ_y) are harder to simulate than the mean, with values over most of India being underpredicted, and values in the southern peninsula being strongly overpredicted. Similar results were found in GLAM2004. Observed values of σ_y in the current study are higher than in GLAM2004 because of the finer spatial scale of the ERA-40 grid; yield data aggregated to a finer resolution tend to have higher standard deviations (Hansen and Jones 2000). The ERA-40-driven simulations do have higher values of σ_y than the GLAM2004 simulations. This is partly due to the use of daily, as opposed to monthly, VPD, radiation, and temperature data.

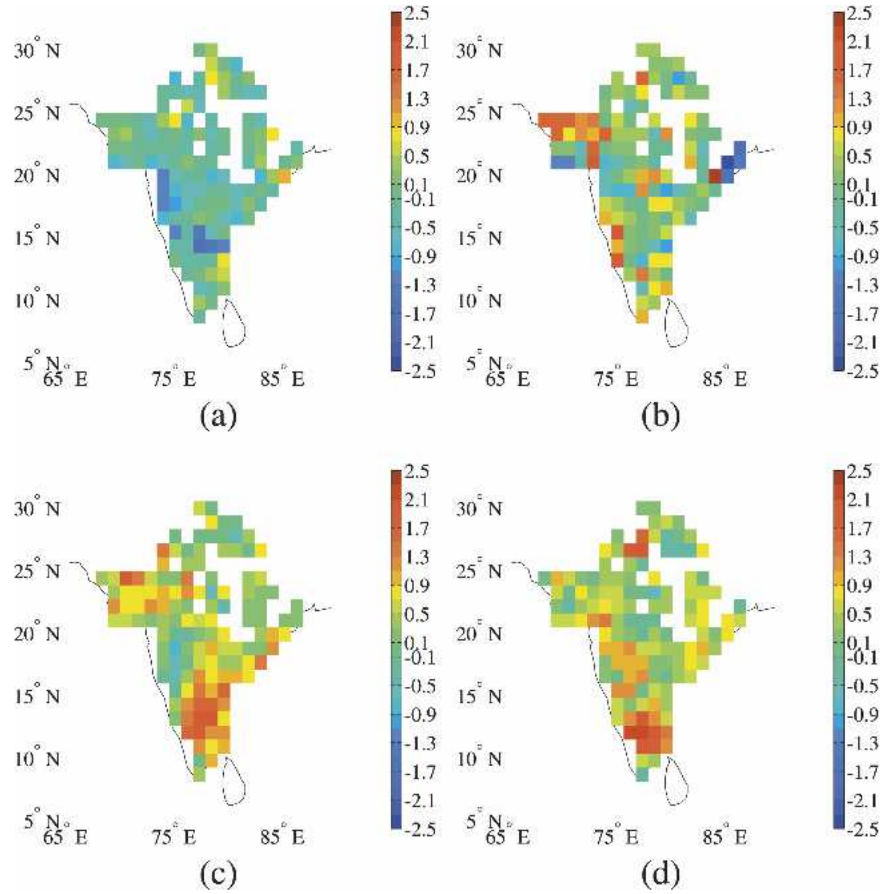

FIG. 8. The skewness of the linearly detrended groundnut yields for the periods of (a) 1966–77 and (b) 1978–89, and the skewness of the Jun–Sep ERA-40 precipitation for (c) 1966–77 and (d) 1978–89.

Overall, correlations between observed and simulated yields are comparable with observed yield–rainfall correlations (cf. Figs. 5a and 11c). However, some grid cells show a correlation between rainfall and observed yield that is not reflected in the correlation between the observed and predicted yields, and some grid cells show the converse. Further examination shows that, at least in some cases, this discrepancy is due to the VPD. If simulated yields are VPD limited, and not rainfall limited, and VPD is not correlated with rainfall, then the observed rainfall signal is not found in the simulations. Conversely, if there is a VPD signal in the observed yields, but no rainfall signal, VPD can impact yields. This effect may be a real (direct) effect, or it may be an indirect effect; larger VPDs imply drier conditions, and independently, via the GLAM formulation, tend to produce higher yields (GLAM2004).

Changes in observed and simulated weather–yield correlations over time (Table 1) are manifest in changes in skill over time. Grid cell GJ has a correlation (observed and simulated yields) of $r_{66} = 0.34$ for the period of 1966–77, and $r_{78} = 0.71$ for 1978–89. The corresponding grid cell in the GLAM2004 simulations showed no such change ($r_{66} = 0.79$; $r_{78} = 0.70$). This difference in behavior is also manifest in the correlation between ERA-40 and IITM rainfall for this location—0.41 and 0.86 for the two respective time periods. Using the August planting window (P1) reverses the sign of the change ($r_{66} = 0.91$; $r_{78} = 0.47$). This situation is common across Gujarat. A reduction in correlation between crop yield and weather variables across these two periods was observed by Parthasarathy et al. (1992) for cereals plus pulses, and Kulkarni and Pandit (1988) for district yields of sorghum. In the latter case the change was attributed to the introduction of new, higher-yielding, less weather-sensitive varieties in 1977–78. In the current study, it is clear that uncertainty in the planting date significantly decreases our ability to either ascribe a cause for this change in correlation, or to quantify sources of error in the input data.

The length of the growing season provides one way to assess the accuracy of the simulations. Across much of the southern peninsula both the IITM data and ERA-40 show that the growing season, as defined by rains, is longer than the crop durations that are predicted by the model. Approximately 30% of the May–

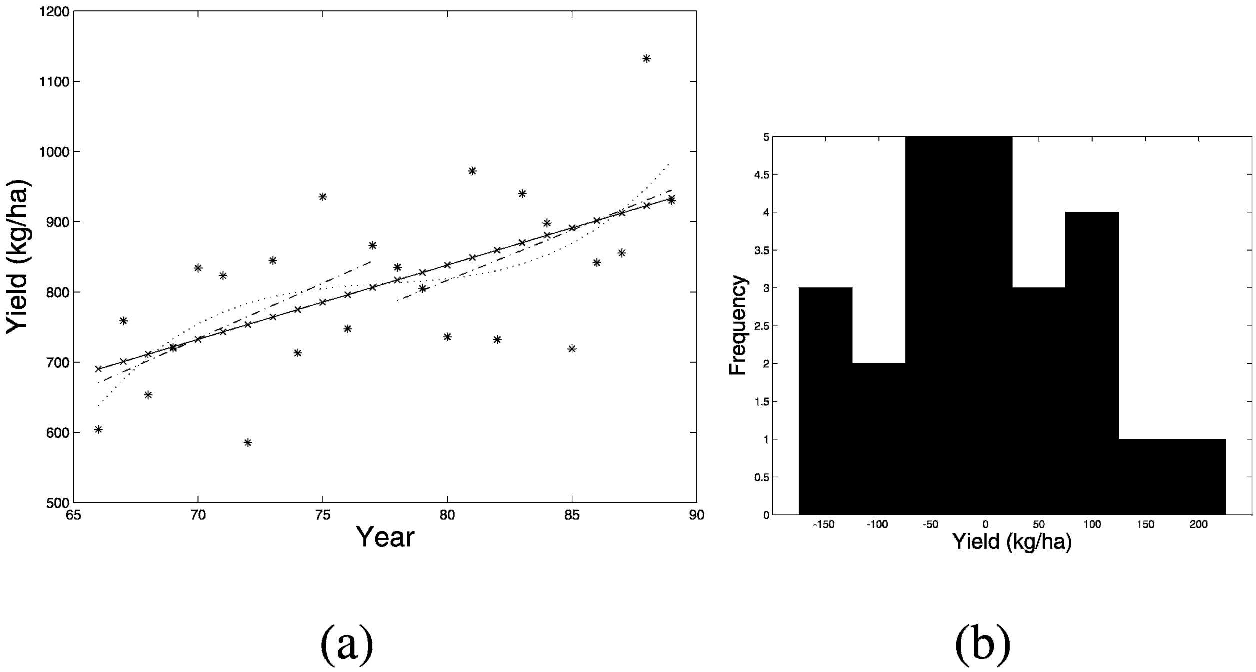

FIG. 9. (a) All-India groundnut yields (stars) with the fitted single linear (unbroken line), quadratic (crosses), cubic (dotted line), and two-piece linear (dot–dashed line) trends; (b) the residuals for the linear fit.

November rain falls in the growing season for the control simulations in the south. In southern India, use of a planting window 1 month later than the control (simulation P2) increases this fraction to 40%. The extended-duration crop (simulation E1) has up to 50% of the May–November rainfall during the growing season. Both P2 and E1 reduce the error in mean yields in the southern peninsula. Note that a study in the Anantapur region (~15°N, 75.5°E) by Gadgil et al. (1999) reported model results suggesting that the broad planting window used by farmers in this region (22 June–17 August) minimized the risk of failure, while a later (mid-July onward) planting window increased yields in 78% of the years. Thus, the use of a later planting window in this study does not result in planting dates beyond locally observed values.

Bias correcting the ERA-40 rainfall to IITM values can improve on the accuracy of the control simulations

TABLE 2. Correlation coefficients between all-India yield and (i) all-India precipitation (Parthasarathy et al. 1995), and (ii) Niño-1 and -2 region (0°–10°S, 90°–80°W) Sep SSTs. The yield data were detrended according to the parameterizations shown (RM is the running mean). The results point consistently to a change in correlation over time. Correlations refer to the years shown, except where otherwise noted: * denotes 1969–77 and 1978–87, and ** denotes 1970–77 and 1978–85. Statistically significant correlations ($p < 0.01$) are shown in boldface.

	Precipitation		SST	
Trend	1966–77	1978–89	1966–77	1978–89
Linear	**0.96**	0.67	**−0.78**	−0.26
Quadratic	**0.96**	0.67	**−0.78**	−0.26
Cubic	**0.91**	0.65	**−0.85**	−0.17
Two-piece linear	**0.92**	0.67	**−0.82**	−0.23
5-yr RM*	**0.93**	0.61	**−0.92**	−0.03
9-yr RM**	**0.97**	0.58	**−0.84**	0.16

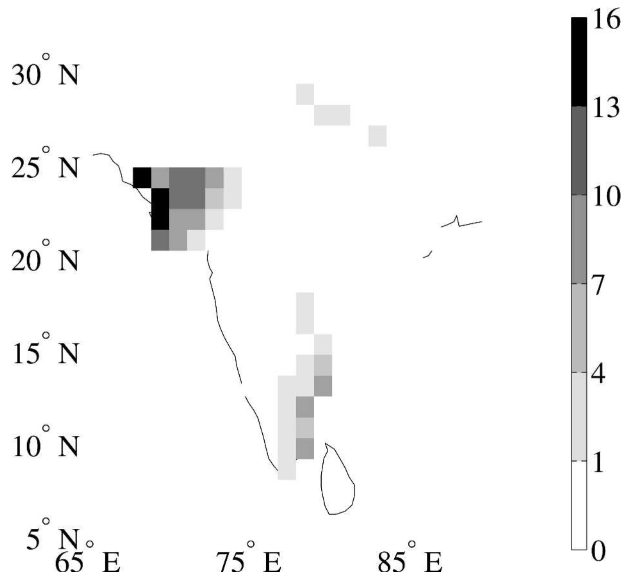

FIG. 10. Analysis of apparent water-use efficiency: the number of years in the period of 1966–89 for which the ratio of observed groundnut yield to JJAS rainfall is greater than 30 kg ha^{-1} cm^{-1}. A usual value for this ratio would be 1000/40 = 25 (see text). The same plot for simulated yields (not shown) has only four grid boxes with values of two or more.

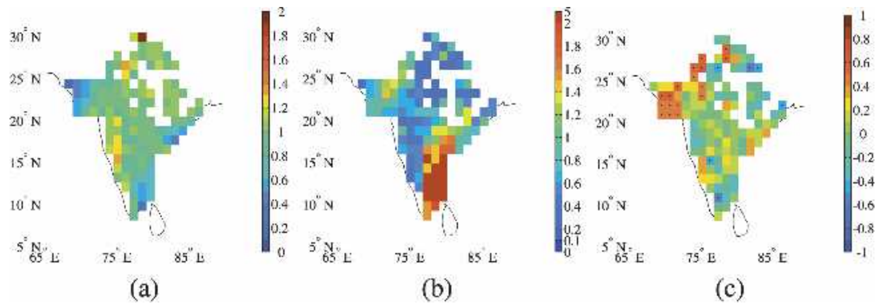

FIG. 11. Comparison between simulated (control run) and observed yields for the period of 1966–89: ratio of simulated to observed (a) mean and (b) standard deviation, and (c) correlation between simulated and observed yields. Significant correlations are marked with a dot.

(Fig. 11). Figure 13a shows the results for bias correction of the mean climatology (B1). Error in mean yields in many regions, particularly Gujarat, is reduced. In Gujarat, correlations between observed and predicted yields (not shown) rise: of the 10 significant correlations in B1 in GJ, 7 are higher than those in the control run. In other regions (e.g., AP) correlations are not improved. The lack of improvement in AP is, in part, due to the smaller climate signal. The accuracy of the ERA-40 seasonal cycle of precipitation (higher in GJ than in AP; Fig. 4) may play a part in this. Simulation B2 is a further improvement on B1 across much of Gujarat and the southern peninsula (Fig. 13b). The additional improvement seen in B2 over B1 demonstrates the importance of interannual variability, which B2 corrects but B1 does not. For simulation B2, 40% of the grid cells have mean yields within 5% of observed values, 62% are within 10% of observations, and 88% are within 50% of observations.

The simulations of the extended-duration crop (E1) show a reduced rmse in yield compared to the control run (Fig. 13c). The E1B1 simulation produces a lower rmse than either B1 or E1 alone in all but two grid cells (Fig. 13d). Note that in regions with lower air tempera-

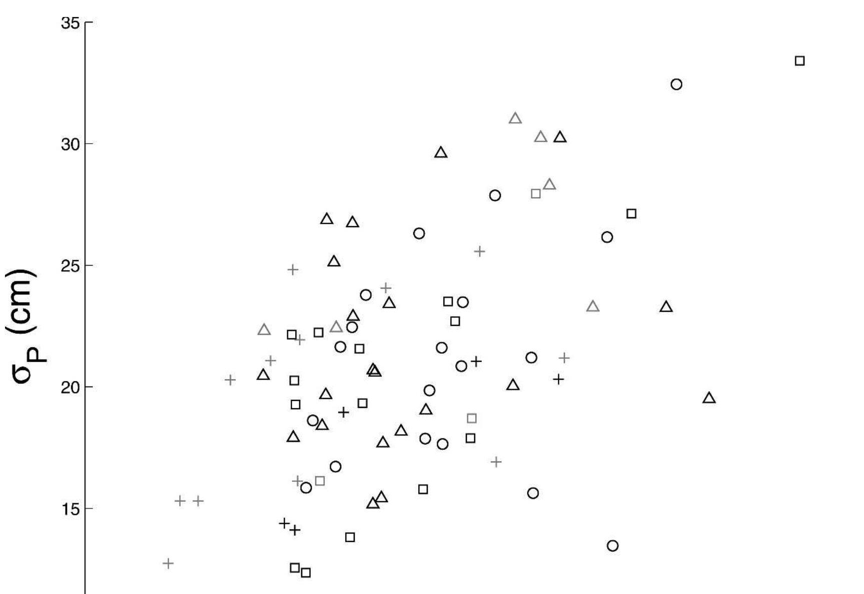

FIG. 12. Error in crop yield simulation (circles indicate 0%–5%, squares indicate 5%–10%, triangles indicate 10%–25%, "+" indicates 25%–50%, and "x" indicates 50%–100%) for two YGP intervals (0.05–0.15 in black and 0.90–1.00 in gray) plotted on a graph of mean vs standard deviation of Jun–Sep precipitation for the period of 1966–89. Only points for which both the 1966–77 and 1978–89 values of YGP fall into the designated intervals are shown. An additional point (x) at (155, 55) is omitted for clarity.

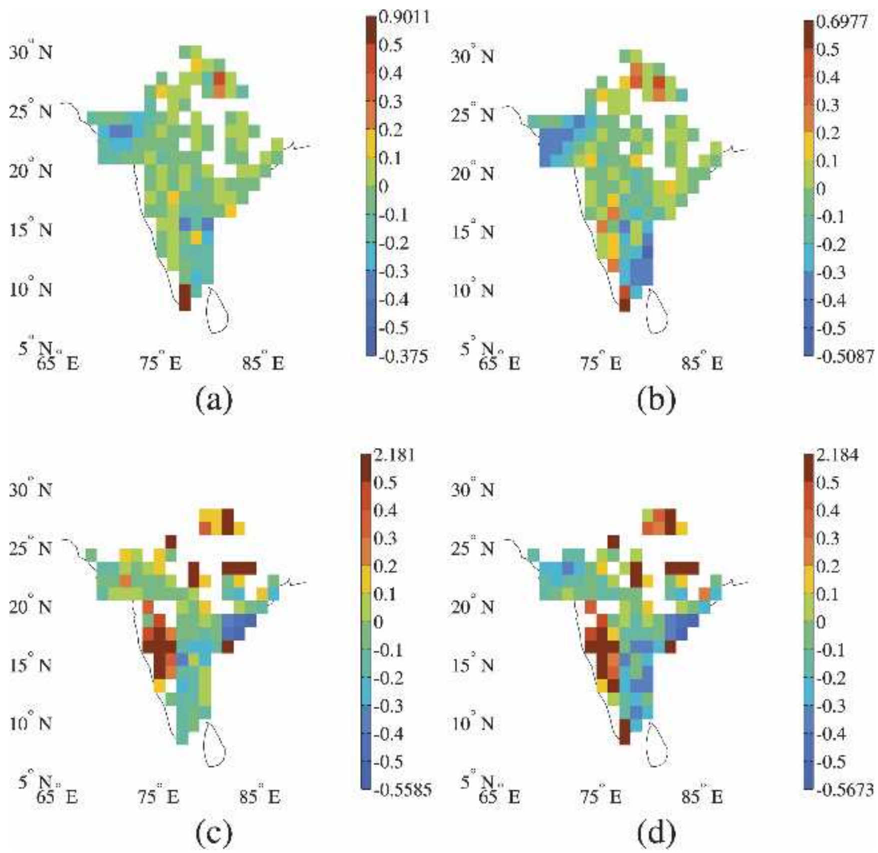

FIG. 13. Fractional changes in the rmse in yield, from the control run baseline, for three model runs: (a) bias correction B1, (b) bias correction B2, (c) extended duration crop E1, and (d) extended duration crop with bias correction B1; B1 shifts linearly the ERA-40 1966–89 precipitation so that its climatology matches the IITM climatology for that period; B2 shifts the ERA-40 1966–89 precipitation so that each JJAS total agrees with the IITM JJAS total. Regions where E1 results in a crop that has not matured before the end of Nov are blanked out.

tures, E1 can result in an unrealistically long duration. This highlights one problem with running a generalized crop parameterization across large areas.

4. Discussion: Nonstationarity in the yield–weather relationship

The time dependence of the yield–weather relationship (section 3b and Table 1) has important implications. The changes in skewness shown in Fig. 8 suggest that the changes in weather–yield correlations are in part due to nonclimatic factors. However, the analysis in section 3b suggests that the observed nonstationarity is not solely a result of the detrending method chosen. It is likely that a combination of effects, such as planting window and yield technology trend assumptions, changes in data accuracy, climatic trends, and random variability in the system, contribute to the detection of nonstationarity. The key question for prediction is *how much* of the nonstationarity in the weather–yield relationship can be accurately simulated. This is an especially important question in the context of climate change. While a mean shift in climate may (or may not) relate linearly to a mean shift in yields, a change in climate variability is less likely to do so, particularly when the impact of extreme weather events (e.g., high temperatures near anthesis; Wheeler et al. 2000) are taken into consideration. The potential contribution of effects such as these to nonstationarity is not known.

This study found that a mean bias correction to rainfall improved (yield) results in many regions (Fig. 13), particularly in Gujarat where there is a strong climate signal, and the seasonal cycle of rainfall is well represented by ERA-40. The simulation of monthly weather–yield correlations in Gujarat (Fig. 7) relies on the accuracy of the seasonal cycle of weather. Changes in climate variability are likely to be hard to correct using a mean bias correction, and the resulting (nonlinear) impact on yield adds further uncertainty. Capturing the impact of subseasonal weather variability is an important challenge that must be faced by the modeling community when seeking to model yields under changing climates. An example of changing subseasonal weather variability was presented in section 3a. Longer time series, and more precise data on planting date, are needed in order to examine the impact of such changes on crop yield.

It was shown in section 3b that the assumption of a planting window that remains constant over time can, where there is a climate signal, result in the accurate simulation of changes over time in rainfall–yield correlations. Any test of the statistical significance of such changes in correlation would rely upon an assumed form for the PDF of the detrended yield. This makes any study of causality speculative. The further evidence of changes in weather–yield relationships presented in section 3b does suggest that this nonstationarity is a real effect in the India-wide context.

Wu and Wang (2002) are among the many authors to explore another example of nonstationarity, namely, ENSO and the Asian summer monsoon. Explanations for the change in the nature of the ENSO–monsoon relationship vary from physical mechanisms (e.g., Kumar et al. 1999) to stochastic processes (e.g., Gershunov et al. 2001). This change is variously cited either as occurring near the early 1980s (as above), or as being part of an interdecadal variation (e.g., Torrence and Webster 1999). Nonstationary relationships can be the subject of much debate. Rather than make premature statements regarding the specific change in yield–weather relationships discussed here, the authors wish to draw attention to this issue of nonstationarity in the response of crops to weather. While research has been directed at the impact of climate variability on both observed (e.g., Lansigan et al. 2000) and simulated (e.g., Legler et al. 1999; Southworth et al. 2000) yields, the issue of nonstationary links between climate and agriculture, while having emerged in a study of disease dynamics (Rodó et al. 2002), has, to the authors' knowledge, not been explored by the crop modeling community. Because it could have implications for the prediction of crop yield, it is a potentially important research topic.

5. Conclusions: Issues and challenges

Some specific observations regarding the use of ERA-40 GLAM for large-area crop simulations have been made in this study. These observations are relevant to the broader methodology for the development of a combined seasonal weather and crop productivity forecasting system, as described in Challinor et al. (2003). The use of a global (domainwide) crop parameterization has associated problems. Simulations using the extended duration crop in southern India improved agreement with observed yields; this shows that it is important to ensure that the simulated crop durations match the observations (either in cropping practice, or a surrogate—observed rainfall). This study has shown that reanalysis data can be used to explore sources of uncertainty in yield simulation. Sufficient uncertainty exists so as to make it difficult for a modeling study on this scale to assess the accuracy of inputs, such as crop duration and the planting window. This uncertainty results in difficulty in ascribing sources of error in the input data. Because two of the simplest adaptations to climate change are changes in (total) crop duration (via the introduction of newly developed or existing cultivars) and planting date, this point is particularly relevant to climate change adaptation studies.

Mean yields have been simulated well across a range of environments, regardless of the strength of the climate signal. A simple mean bias correction further improved the accuracy of yield simulations. In Gujarat there is both a strong climate signal in yields, and the ERA-40 seasonal cycle is reasonably accurate. Correlations between simulated yields and weather agree reasonably with those observed, and this region shows the highest correlations between observed and simulated yields. The fact that the crop model parameters have not been fine-tuned for the ERA-40 data gives further indication that encouraging results can be found when studying crop/climate interactions on large spatial scales using relatively simple mechanistic parameterizations of crop growth. Together with the importance of the region for Indian groundnut production (Kakani 2001), these facts point to the potential for seasonal yield predictions in this region.

One of the challenges facing climate impacts research is to improve the accuracy of yield simulations over large areas, in order to improve assessments of the impact of climate variability and change on crop production. The fact that bias correction of the ERA-40 precipitation reduced rmse in the yield suggests that improved GCM skill could be translated into improved yield estimation. The accuracy of yield simulations can also be improved by minimizing errors from other sources. Remote sensing, for example, enables the assessment of aggregation error by identifying homogeneous crop regions (Basso et al. 2001; Jones and Barnes 2000; Guerif and Duke 2000). Remote sensing may also provide a way to test simulated crop durations, by observing ground cover over the season. Regional climate modeling enables the simulation of crop growth on higher spatial resolutions, hence, reducing aggregation error. Where precise input management and crop cultivar information are known, detailed physiological crop models (e.g., Boote and Jones 1998) can improve the understanding of the impact of climate variability and change on crop yield, and, hence, the ability to simulate it. That some skill exists in the simulation of crop yield, across a range of environments, using fewer inputs with a simpler large-area crop model and reanalysis data suggests that climate change scenarios could be addressed using a similar approach.

A second important challenge is the detection and modeling of nonstationary weather–yield relationships.

This has emerged from this study as being an important part of the process of understanding and predicting the impacts of climate variability and change. Regional studies using detailed crop models may aid our understanding of the causes of nonstationarity. Larger-scale integrated modeling of climate and crops (e.g., Tsvetsinskaya et al. 2001a,b) provides another methodology, with the advantage that full coupling of the climate and the crop allows the crop to impact its own environment and vice versa.

Acknowledgments. The authors thank the International Crops Research Institute for the Semi–Arid Tropics (ICRISAT), the European Centre for Medium-Range Weather Forecasts (ECMWF), and the Indian Institute for Tropical Meteorology (IITM) for the crop productivity and weather data.

REFERENCES

Andersson, E., and Coauthors, 2005: Assimilation and modeling of the atmospheric hydrological cycle in the ECMWF forecasting system. *Bull. Amer. Meteor. Soc.,* **86,** 387–402.

Annamalai, H., J. M. Slingo, K. R. Sperber, and K. Hodges, 1999: The mean evolution and variability of the Asian summer monsoon: Comparison of ECMWF and NCEP–NCAR reanalyses. *Mon. Wea. Rev.,* **127,** 1157–1186.

Atwood, J., S. Shaik, and M. Watts, 2003: Are crop yields normally distributed? A reexamination. *Amer. J. Agric. Econ.,* **85,** 888–901.

Basso, B., J. T. Ritchie, F. J. Pierce, R. P. Braga, and J. W. Jones, 2001: Spatial validation of crop models for precision agriculture. *Agric. Syst.,* **68,** 97–112.

Bengtsson, L., and J. Shukla, 1988: Integration of space and in situ observations to study global climate change. *Bull. Amer. Meteor. Soc.,* **69,** 1130–1143.

Betts, A. K., J. H. Ball, M. Bosilovich, P. Viterbo, Y. Zhang, and W. B. Rossow, 2003: Intercomparison of water and energy budgets for five Mississippi subbasins between ECMWF reanalysis (ERA-40) and NASA Data Assimilation Office fvGCM for 1990–1999. *J. Geophys. Res.,* **108,** 8618, doi:10.1029/2002JD003127.

Boote, K. J., and J. W. Jones, 1998: Simulation of crop growth: Cropgro model. *Agricultural Systems Modeling and Simulation,* R. M. Peart and R. B. Curry, Eds., Marcel Dekker, 651–692.

Camberlin, P., and M. Diop, 1999: Inter-relationships between groundnut yield in Senegal, interannual rainfall variability and sea-surface temperatures. *Theor. Appl. Climatol.,* **63,** 163–181.

Challinor, A. J., J. M. Slingo, T. R. Wheeler, P. Q. Craufurd, and D. I. F. Grimes, 2003: Toward a combined seasonal weather and crop productivity forecasting system: Determination of the spatial correlation scale. *J. Appl. Meteor.,* **42,** 175–192.

——, ——, ——, and F. J. Doblas-Reyes, 2004a: Probabilistic hindcasts of crop yield over western India. *Tellus,* in press.

——, T. R. Wheeler, J. M. Slingo, P. Q. Craufurd, and D. I. F. Grimes, 2004b: Design and optimisation of a large-area process-based model for annual crops. *Agric. For. Meteor.,* **124,** 99–120.

Corte-Real, J., H. Xu, and B. D. Qian, 1999: A weather generator for obtaining daily precipitation scenarios based on circulation patterns. *Climate Res.,* **13,** 61–75.

Davis, C. A., K. W. Manning, R. E. Carbone, S. B. Trier, and J. D. Tuttle, 2003: Coherence of warm-season continental rainfall in numerical weather prediction models. *Mon. Wea. Rev.,* **131,** 2667–2679.

FAO/UNESCO, 1974: *Southeast Asia.* Vol. 9, *FAO/UNESCO Soil Map of the World, 1:5,000,000,* UNESCO, 149 pp.

Gadgil, S., P. R. S. Rao, K. N. Rao, and K. Savithiri, 1999: Farming strategies for a variable climate. Indian Institute of Science CAOS Tech. Rep. 99AS7, 46 pp.

——, ——, and ——, 2002: Use of climate information for farm-level decision making: Rainfed groundnut in southern India. *Agric. Syst.,* **74,** 431–457.

Gershunov, A., N. Schneider, and T. Barnett, 2001: Low-frequency modulation of the ENSO–Indian monsoon rainfall relationship: Signal or noise? *J. Climate,* **14,** 2486–2492.

Gibson, J. K., P. Kållberg, and S. Uppala, 1996: The ECMWF Re-Analysis (ERA) project. *ECMWF Newsletter,* No. 73, ECMWF, 7–17.

——, ——, ——, A. Hernandez, A. Nomura, and E. Serrano, 1997: Era description. Re-Analysis (ERA) Project Report Series 1, ECMWF, 72 pp.

Guerif, M., and C. L. Duke, 2000: Adjustment procedures of a crop model to the site specific characteristics of soil and crop using remote sensing data assimilation. *Agric. Ecosyst. Environ.,* **81,** 57–69.

Hansen, J. W., and J. W. Jones, 2000: Scaling-up crop models for climatic variability applications. *Agric. Syst.,* **65,** 43–72.

Hoshen, M. B., E. Worrall, A. P. Morse, and M. C. Thomson, 2003: A weather-driven model of malaria transmission. *Malaria J.,* **3,** doi:10.1186/1475-2875-3-32.

Jones, D., and E. M. Barnes, 2000: Fuzzy composite programming to combine remote sensing and crop models for decision support in precision crop management. *Agric. Syst.,* **65**, 137–158.

Just, R. E., and Q. Weninger, 1999: Are crop yields normally distributed? *Amer. J. Agric. Econ.,* **81,** 287–304.

Kakani, V. G., 2001: Quantifying the effects of high temperature and water stress in groundnut. Ph.D. thesis, University of Reading, 309 pp.

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

Kulkarni, B. S., and S. N. N. Pandit, 1988: A discrete step in the technology trend for sorghum yields in Parbhani, India. *Agric. For. Meteor.,* **42**, 156–165.

Kumar, K. K., B. Rajagopalan, and M. A. Cane, 1999: On the weakening relationship between the Indian Monsoon and ENSO. *Science,* **284,** 2156–2159.

Lansigan, F. P., W. L. De Los Santos, and J. O. Coladilla, 2000: Agronomic impacts of climate variability on rice production in the Philippines. *Agric. Ecosyst. Environ.,* **82,** 129–137.

Legler, D. M., K. J. Bryant, and J. J. O'Brien, 1999: Impact of ENSO-related climate anomalies on crop yields in the U.S. *Climatic Change,* **42,** 351–375.

Mearns, L. O., T. Mavromatis, and E. Tsvetsinskaya, 1999: Comparative response of EPIC and CERES crop models to high and low spatial resolution climate change scenarios. *J. Geophys. Res.,* **104** (D6), 6623–6646.

——, W. Easterling, C. Hays, and D. Marx, 2001: Comparison of agricultural impacts of climate change calculated from the high and low resolution climate change scenarios: Part I. The uncertainty due to spatial scale. *Climatic Change,* **51,** 131–172.

Moss, C. B., and J. S. Shonkwiler, 1993: Estimating yield distributions with a stochastic trend and nonnormal errors. *Amer. J. Agric. Econ.,* **75,** 1056–1062.

Palmer, T. N., and Coauthors, 2004: Development of a European multimodel ensemble system for seasonal-to-interannual prediction (DEMETER). *Bull. Amer. Meteor. Soc.,* **85,** 853–872.

Parthasarathy, B., K. R. Kumar, and A. A. Munot, 1992: Forecast of rainy-season food-grain production based on monsoon rainfall. *Indian J. Agric. Sci.,* **62,** 1–8.

——, A. A. Munot, and D. R. Kothawale, 1995: All India monthly and seasonal rainfall series: 1871–1993. *Theor. Appl. Climatol.,* **49,** 217–224.

Pathak, H., and Coauthors, 2003: Trends of climatic potential and on-farm yields of rice and wheat in the Indo-Gangetic Plains. *Field Crops Res.,* **80,** 223–234.

Priestly, 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.

Ramirez, O. A., S. Misra, and J. Field, 2003: Crop-yield distributions revisited. *Amer. J. Agric. Econ.,* **85,** 108–120.

Reddy, P. S., Ed., 1988: *Groundnut.* Indian Council of Agricultural Research, 583 pp.

Rodó, X., M. Pascual, G. Fuchs, and A. S. G. Faruque, 2002: Enso and cholera: A non-stationary link related to climate change? *Proc. Natl. Acad. Sci.,* **99,** 12 901–12 906.

Selvaraju, R., 2003: Impact of El Niño–Southern Oscillation on Indian foodgrain production. *Int. J. Climatol.,* **23,** 187–206.

Semenov, M. A., and E. M. Barrow, 1997: Use of a stochastic weather generator in the development of climate change scenarios. *Climatic Change,* **35,** 397–414.

——, and R. J. Brooks, 1999: Spatial interpolation of the LARS-WG stochastic weather generator in Great Britain. *Climate Res.,* **11,** 137–148.

Sinclair, T. R., and N. Seligman, 2000: Criteria for publishing papers on crop modelling. *Field Crops Res.,* **68,** 165–172.

Sivakumar, M. V. K., and P. S. Sarma, 1986: Studies on water relations of groundnut. *Agrometeorology of Groundnut (Proceedings of the International Symposium held at ICRISAT Sahelian Center, Niamey, Niger),* M. V. K. Sivakumar and S. M. Virimani, Eds., ICRISAT, 83–98.

Southworth, J., J. C. Randolph, M. Habeck, O. C. Doering, R. A. Pfeifer, D. G. Rai, and J. J. Johnston, 2000: Consequences of future climate change and changing climate variability on maize yields in the midwestern United States. *Agric. Ecosyst. Environ.,* **82,** 139–158.

Torrence, C., and P. J. Webster, 1999: Interdecadal changes in the ENSO–monsoon system. *J. Climate,* **12,** 2679–2690.

Troccoli, A., and P. Kållberg, 2004: Precipitation correction in the ERA-40 reanalysis. ERA-40 Project Report Series 13, ECMWF, 10 pp.

Tsvetsinskaya, E. A., L. O. Mearns, and W. E. Easterling, 2001a: Investigating the effect of seasonal plant growth and development in three-dimensional atmospheric simulations. Part I: Simulation of surface fluxes over the growing season. *J. Climate,* **14,** 692–709.

——, ——, and ——, 2001b: Investigating the effect of seasonal plant growth and development in three-dimensional atmospheric simulations. Part II: Atmospheric response to crop growth and development. *J. Climate,* **14,** 711–729.

Virmani, S. M., and N. J. Shurpali, 1999: Climate prediction for sustainable production of rainfed groundnuts in SAT: Crop establishment risks in the Anantapur region. International Crop Research Institute for the Semi–Arid Tropics Tech. Manual 4, 50 pp.

Wheeler, T. R., P. Q. Craufurd, R. H. Ellis, J. R. Porter, and P. V. V. Prasad, 2000: Temperature variability and the annual yield of crops. *Agric. Ecosyst. Environ.,* **82,** 159–167.

Wilcoxon, F., 1945: Individual comparisons by ranking methods. *Biometrics,* **1,** 80–83.

Wilks, D. S., 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.

Wu, R., and B. Wang, 2002: A contrast of the east Asian summer monsoon–ENSO relationship between 1962–77 and 1978–93. *J. Climate,* **15,** 3266–3279.

Yu, Q., H. Hengsdijk, and J. D. Liu, 2001: Application of a progressive-difference method to identify climatic factors causing variation in the rice yield in the Yangtze Delta, China. *Int. J. Biometeor.,* **45**, 53–58.

Zhang, Y., T. Li, B. Wang, and G. Wu, 2002: Onset of the summer monsoon over the Indochina Peninsula: Climatology and interannual variations. *J. Climate,* **15,** 3206–3221.