Evaluating the Sensitivity of Agricultural Model Performance to Different Climate Inputs*

MICHAEL J. GLOTTER AND ELISABETH J. MOYER

Department of the Geophysical Sciences, The University of Chicago, Chicago, Illinois

ALEX C. RUANE

NASA Goddard Institute for Space Studies, New York, New York

JOSHUA W. ELLIOTT

Computation Institute, The University of Chicago, Chicago, Illinois

(Manuscript received 24 April 2015, in final form 4 September 2015)

ABSTRACT

Projections of future food production necessarily rely on models, which must themselves be validated through historical assessments comparing modeled and observed yields. Reliable historical validation requires both accurate agricultural models and accurate climate inputs. Problems with either may compromise the validation exercise. Previous studies have compared the effects of different climate inputs on agricultural projections but either incompletely or without a ground truth of observed yields that would allow distinguishing errors due to climate inputs from those intrinsic to the crop model. This study is a systematic evaluation of the reliability of a widely used crop model for simulating U.S. maize yields when driven by multiple observational data products. The parallelized Decision Support System for Agrotechnology Transfer (pDSSAT) is driven with climate inputs from multiple sources—reanalysis, reanalysis that is bias corrected with observed climate, and a control dataset—and compared with observed historical yields. The simulations show that model output is more accurate when driven by any observation-based precipitation product than when driven by non-bias-corrected reanalysis. The simulations also suggest, in contrast to previous studies, that biased precipitation distribution is significant for yields only in arid regions. Some issues persist for all choices of climate inputs: crop yields appear to be oversensitive to precipitation fluctuations but undersensitive to floods and heat waves. These results suggest that the most important issue for agricultural projections may be not climate inputs but structural limitations in the crop models themselves.

1. Introduction

Understanding future food production is critical in conditions of changing climate and growing population (Porter et al. 2014). Meeting agricultural food demand that is estimated to increase by ~60% by 2050 (OECD/FAO 2012) presents a significant challenge for future society. Changing demand coupled with changing production will likely have significant impacts on food availability, and subsequently affect food prices (Nelson et al. 2014) and migration patterns (Feng et al. 2012). However, estimating future food production requires understanding past food production. Agricultural models are needed to simulate yield in future climate conditions that are outside the range of historical experience. These models must themselves be validated by ensuring that they can reproduce past observed yields. Model validation efforts have increased in recent years with increasing recognition of the importance of food security to decision-making about climate change. For example, the Agricultural Model Intercomparison and Improvement Project (AgMIP; Rosenzweig et al. 2013) began in 2010 as an international effort to improve agricultural

impacts assessments and to understand how climate inputs and model structure introduce uncertainty. Historical assessments play a key role in characterizing that uncertainty.

Reliable historical validation requires both accurate agricultural models and accurate data inputs. The most common type of model for agricultural projections under climate change are process-based agricultural models, which attempt to explicitly represent biophysical processes affecting crop growth. These models have been shown to reproduce crop yields in many cases but have some known limitations (e.g., models do not represent damages from pests). Recent intercomparisons between models and field experiments suggest that models may also misrepresent some crop responses to changes in climate (Asseng et al. 2013; Bassu et al. 2014). In addition, the inputs used to drive agricultural models for historical assessments may themselves be problematic. Local weather station observations can be used for studies at single locations (e.g., Asseng et al. 2013), but large-scale agricultural assessments require wider coverage. Because observational networks are irregularly spaced and spatially and temporally sparse, some approach must be used to fill data gaps. Gridded weather records interpolate local observations; retrospective analyses (reanalyses) use a hybrid of models and observational data (e.g., Saha et al. 2010). Both approaches may involve biases that can affect agricultural assessments.

Reanalyses are particularly useful for agricultural studies in developing nations where surface measurements are scarce. Reanalyses provide a continuous global record of physically consistent and high-resolution climate information derived from numerical weather prediction models assimilated to observations of state variables (e.g., wind, humidity, pressure) from multiple data sources (including satellites and balloons as well as ground stations). Several groups have developed reanalyses with different approaches for assimilating data into weather prediction models. These reanalyses compare well to observations, but show differences from each other and some common biases (e.g., Bosilovich et al. 2009).

Reanalyses have known weaknesses in reproducing certain variables important for crop growth. Reanalyses typically do not assimilate precipitation, solar radiation, or near-surface air temperature. Precipitation is especially problematic because weather prediction models must rely on simplified parameterizations of cloud processes (e.g., Stensrud 2007). Crops are particularly sensitive to changes in precipitation (Hatfield et al. 2011), so inaccurate rainfall inputs are a significant concern and could compromise validations (Watson and Challinor 2013).

Although many studies have evaluated reanalyses for applications in the water sector (e.g., Mo et al. 2011; Reichle et al. 2011; Xia et al. 2012; Rodell et al. 2004), few studies have evaluated their utility for agricultural projections. Agricultural studies differ in that crops are sensitive to local and short-term climate conditions, whereas hydrological studies are typically interested in variables such as runoff that integrate conditions across an entire watershed and over longer time periods. Existing studies of reanalyses used in agricultural assessments have drawn inconsistent conclusions. Van Wart et al. (2013) compare crop simulations in localized areas driven by gridded weather products (including reanalysis) or station observations. They find significant differences in yield estimates but do not compare with observed yields. Berg et al. (2010) and Challinor et al. (2005) simulate yields over a wider area using reanalysis with and without bias-corrected precipitation. They find that bias correcting precipitation is necessary to reliably estimate observed yields. De Wit et al. (2010) predict single-year future yields with a crop forecasting model driven by interpolated station observations and by reanalysis with some variables (but not precipitation) bias corrected. They compare with observed yields and conclude that reanalysis without adjustments to precipitation prove suitable (as compared with simulations driven by interpolated station observations). Because no study compares all climate variables separately, these results are difficult to parse. It remains unclear which aspects of reanalysis may be problematic in large-scale agricultural assessments.

In this work, we build on results from previous studies and systematically evaluate the ability of a commonly used crop model to estimate yields when driven by a variety of climate inputs. We evaluate both crop model and climate inputs to assess whether current technology is sufficient and fit for the purpose of agricultural assessments. We drive large-scale crop simulations with many different data products, including gridded weather data, reanalysis, and reanalysis with different variables bias corrected or substituted with observational data. We then compare directly with observed yields to evaluate both yield sensitivities to climate inputs and fundamental model performance. We restrict our study to the United States, where high-quality subnational observational yield records allow us to differentiate misrepresentations due to climate input errors from those due to crop model limitations. In the remainder of this paper, in section 2 we describe the model, climate inputs, and methods; in sections 3a and 3b we evaluate the impact of climate input data choices on yields; in section 3c we extend the analysis in the context of detecting extreme events; in section 3d we evaluate the ability of the model to reproduce observed

interannual yield variations; and in section 4 we discuss implications for agricultural assessments.

2. Materials and methods

a. Crop model and yield comparison

We simulate maize yields with a variety of climate inputs from the period 1980–2009 using the Decision Support System for Agrotechnology Transfer, version 4.5 (DSSAT v4.5), Crop Estimation through Resources and Environmental Synthesis (CERES) maize model (Jones et al. 2003; G. Hoogenboom et al. 2010, unpublished conference presentation). We use a parallelized version of DSSAT (pDSSAT) under the parallelized System for Integrating Impacts Models and Sectors (pSIMS; Elliott et al. 2014) to simulate yield at 1/4° resolution over the contiguous United States. The pSIMS framework allows for large-scale simulations by using the Swift parallel scripting language (Zhao et al. 2007) to process DSSAT concurrently on several clusters. We use fixed management practices to isolate crop model responses to weather and climate. Specifically, we assume temporally homogeneous (but spatially heterogeneous) cultivar choice and planting windows (as defined by Elliott et al. 2015), and uniform fertilizer application (of 150 kg ha^{-1}). The growing season is calibrated in each grid cell to reproduce historical average growing seasons [from U.S. Department of Agriculture (USDA) survey data] using phenology parameters (fixing the accumulated thermal units or growing degreedays between planting and maturity). We simulate the dominant cropped soil type in each grid cell using fixed soil definitions from the Harmonized World Soil Database (Nachtergaele et al. 2009). For irrigated areas, DSSAT applies water automatically once soil moisture falls below a set threshold. We simulate rainfed and irrigated maize yields separately for the entire spatial domain, and aggregate yields to the county level using harvested land-area estimates for each category from the Spatial Production and Allocation Model (SPAM) dataset (You et al. 2010). Unless otherwise noted, modeled county yields represent a combination of rainfed and irrigated maize production. We simulate yields using the Penman–Monteith methodology to approximate evapotranspiration, an estimate based on measurements of temperature, wind speed, humidity, and solar radiation. Simulations using the Priestley–Taylor methodology, a simplification that does not require estimates of wind or humidity, are included in the online supplementary material for comparison.

We compare modeled yields to historical yields taken from county-level survey data from the USDA National Agricultural Statistics Service (NASS; Statistical Methods Branch 2012). Because the purpose of this exercise is to assess the importance of observational climate for yield estimates and not our ability to model technological and management advancements, we primarily compare interannual variations between modeled and historical yields. For this purpose we both detrend and normalize national yields to remove some drivers of yield change not associated with changes in weather (e.g., technology). (Historically, advances in U.S. farming practices have increased maize yields by >50% from 1980 to 2009; see supplementary Fig. S1). We detrend by subtracting the least squares linear regression (while preserving mean yield) and normalize by scaling modeled yields by time-averaged NASS yield. In section 3d, we also show model yields modified by a simple variance adjustment, scaling year-to-year variation by that in the historical NASS record [Eq. (1)]:

$$Y_{\text{VA}}^t = Y_N^t \left(\frac{\sigma_{\text{NASS}}}{\sigma_{\text{model}}} \right), \tag{1}$$

where Y_N^t and Y_{VA}^t are normalized and variance-adjusted (respectively) detrended yield at year t, and σ is the standard deviation of U.S. national (modeled or NASS) yields.

For illustrative purposes, we also convert yields to total maize value in dollars. We first calculate national maize production by multiplying national average yield [in metric tons per hectare (t ha^{-1})] with 1980–2009 national average maize harvested hectares (as estimated by SPAM). To convert maize production to dollars, we assume a fixed price of maize around 2010 values of $204 per metric ton. [In reality, production value is also affected by external factors such as insurance coverage, technological advancements, and changes in management practices (Elliott et al. 2013).]

b. Climate inputs

We drive pDSSAT both with reanalysis output—numerical weather prediction models assimilated with observations—and with reanalysis output combined with observation-based weather products. Reanalysis inputs are derived from the National Centers for Environmental Prediction Climate Forecast System Reanalysis (CFSR; Saha et al. 2010), which has been shown to improve representations over previous reanalyses (Higgins et al. 2010; Wang et al. 2011). CFSR simulates the climate system using a coupled atmosphere–ocean model with a sea ice component at a high T382 global resolution [~38 km, or ~(0.313° × 0.313°)]. When run in reanalysis mode, CFSR is constrained by subdaily measurements related to atmospheric motion, satellite radiances, oceanic temperature and salinity profiles, and

TABLE 1. DSSAT simulations with different climate inputs, including production-weighted 30-yr growing-season average climate variables and corresponding yields. (See text for description of simulations.) Quotation marks indicate the same value as noted above. The T_{mean} is the average of T_{max} and T_{min}, as defined by DSSAT for measuring crop growth progress. Climate averages may vary for simulations with the same input source because growing seasons are not fixed. (DSSAT determines growing seasons on the basis of cumulated temperature and soil moisture, which may differ across simulations.) Yields vary by ±10% across different choices of climate inputs. We approximate DSSAT management circa 2000 when NASS yields are ~8.6 t ha^{-1}, but we do not expect an exact match to NASS. Detrended and normalized NASS yields have a mean of ~7.9 and std dev of ~0.7. Overbars indicate mean quantities.

| | Climate source | | | $\overline{T}_{mean}$ | $\overline{T}_{max}$ | $\overline{T}_{min}$ | Precipitation | $\overline{Solar}$ | $\overline{Yield}$ | σ |
Simulation	Temperature	Precipitation	Solar	(°C)	(°C)	(°C)	(cm)	(W m^{-2})	(t ha^{-1})	(t ha^{-1})
CFSR	CFSR	CFSR	CFSR	21.0	27.5	14.4	36.9	262	8.1	1.9
CFSR$_{+s}$	CFSR	CFSR	SRB	”	”	”	”	243	8.1	1.8
CFSR$_{+p}$	CFSR	CPC	CFSR	20.9	27.4	”	42.6	262	9.4	1.6
CFSR$_{+p,s}$	CFSR	CPC	SRB	”	”	”	42.7	243	9.3	1.5
AgCFSR	AgCFSR	AgCFSR	AgCFSR	20.7	27.1	14.2	42.1	238	10.0	1.2
AgCFSR$_{+p}$	AgCFSR	CPC	AgCFSR	”	”	”	42.8	”	10.1	1.2

sea ice concentration, made from weather stations, satellites, and atmospheric platforms. CFSR uses observed precipitation to simulate soil moisture and temperature in its land surface model, but does not directly constrain 2-m temperature or precipitation themselves or cloud cover that affects solar radiation (Saha et al. 2010). Because we simulate yields at ¼° resolution, we regrid coarser climate inputs using nearest-neighbor interpolation.

DSSAT requires, at minimum, four daily weather inputs: maximum and minimum temperature, precipitation, and solar radiation. Temperature is used in part to model crop development through growing degree-days (GDD), a measure of accumulated growing-season temperature above a defined baseline:

$$\text{accumulated GDD}^{i,t} = \sum_{n=0}^{t} T_{mean}^{i,n} - T_{base}$$
$$= \sum_{n=0}^{t} \frac{T_{min}^{i,n} + T_{max}^{i,n}}{2} - T_{base}, \qquad (2)$$

where t indicates day of the growing season, $n = 0$ is the start date of the GDD calculation, and T is the mean, minimum, maximum, or baseline daily temperature (°C). (For DSSAT maize, $T_{base} = 8$°C.) DSSAT does not directly use daily mean temperature in simulations but estimates T_{mean} through T_{max} and T_{min} variables for GDD calculations.

In our analyses CFSR is the default for all climate variables, but in different simulations we manipulate CFSR by either substituting observation-based products for individual input variables, bias correcting the CFSR variables, or both (Table 1). In our notation we identify simulations involving data substitution with subscripts: $+p$, $+s$, and $+p, s$ describe substitution of observation-based precipitation, solar radiation, or both variables, respectively. We identify simulations involving a bias

correction by the prefix "Ag," as in "AgCFSR." Substituting individual variables may affect multivariable correlations, but such practice is common so an assessment is critical.

Observational precipitation is taken from the Climate Prediction Center U.S. Unified Precipitation, version 1.0, dataset (hereinafter CPC). CPC precipitation is derived from ~8000 U.S. rain gauge stations from 1980 to 1991 and ~13 000 from 1992 to 2009. Station data are interpolated to a uniform ¼°-resolution grid using the Cressman method (Cressman 1959), a distance-weighting interpolation. Observational solar radiation is taken from the NASA/GEWEX Surface Radiation Budget project (hereinafter SRB; Stackhouse et al. 2011), which is shown to compare favorably to U.S. weather station measurements (White et al. 2011). The availability of satellite data limits SRB output to years 1984–2007 with global coverage at 1° × 1° resolution. We use solar radiation output from CFSR for years 1980–83 and 2008–09 in order to analyze yields over the full 30-yr period where all other data products are available.

The bias correction we use (AgCFSR; Ruane et al. 2014a) was developed as part of AgMIP to provide consistent temporal and spatial coverage over the world's agricultural regions during the 1980–2010 period. AgCFSR relies on CFSR as its base daily weather and incorporates information from several other datasets. AgCFSR pegs monthly mean temperature to an ensemble of ½° × ½° gridded station datasets [Climatic Research Unit (CRU; Harris et al. 2014) and University of Delaware Willmott–Matsuura (WM; Willmott and Matsuura 1995)] and imposes the average monthly diurnal cycle from CRU. A similar procedure is followed for precipitation [utilizing gauge-based products from CRU, WM, and Global Precipitation Climatology Centre (GPCC); Schneider et al. 2011], although additional resolution is provided by imposing ¼° × ¼° monthly average spatial

a) growing season precipitation
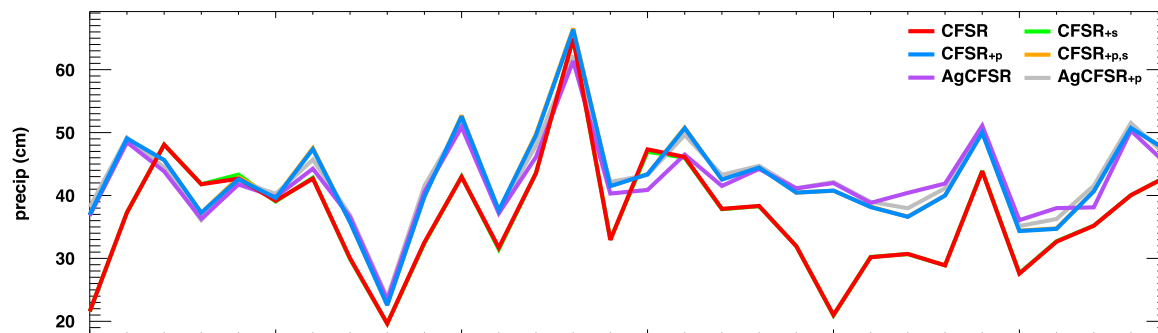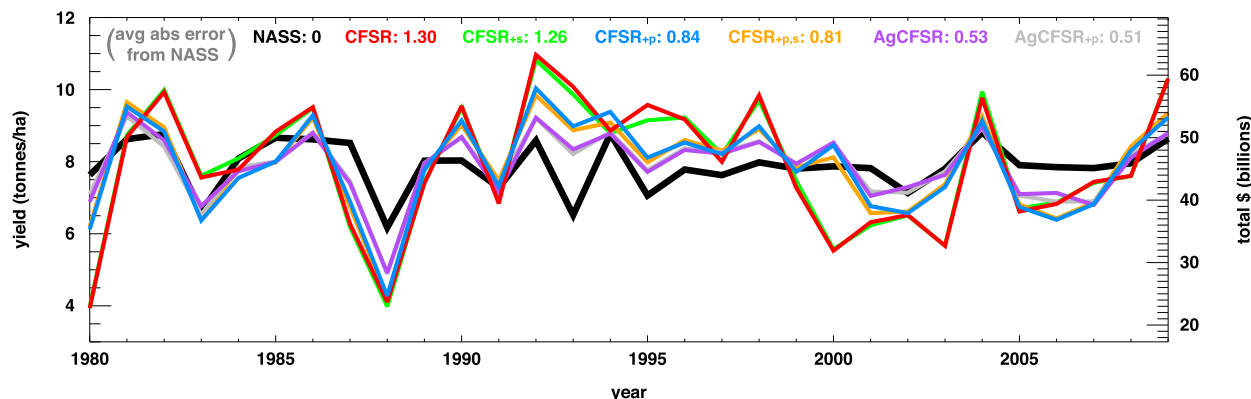

b) detrended, normalized national yield

FIG. 1. Interannual variations in 1980–2009 U.S. (a) precipitation and (b) maize yields for DSSAT simulations with various climate inputs. All data are weighted by harvested maize hectares and averaged over the growing season, and all yields are detrended and normalized. Detrended NASS survey yields are shown in black. Total dollar deviations on the right axis are calculated using 1980–2009 national average maize harvested hectares and a fixed price of maize around 2010 values ($204 per metric ton). For reference, 2010 NASS yields were $\sim 9.6 \, \mathrm{t \, ha^{-1}}$ with a total production value of $\sim$$64 billion ($\sim$$70 billion in 2014 dollars). Variations in yield are largely correlated with variations in precipitation. Simulations do not capture yield losses from 1993 flood. (See section 3c for a discussion.)

patterns from an average of three high-resolution precipitation products (Huffman et al. 2007; Hsu et al. 1997; Joyce et al. 2004). (Because CRU, WM, and GPCC are limited to monthly aggregates, we do not use them as direct inputs to pDSSAT.) The number of rainy days is then adjusted to match CRU, with the driest days removed first when CFSR indicates too many rainy days and small rainfall totals added to the cloudiest days when CFSR indicates too few rain events. Finally, AgCFSR replaces CFSR solar radiation with SRB output, and adjusts CFSR over the missing-data period (i.e., 1980–83, and 2008–10) according to distribution fits from the 1984–2007 SRB output. (Note that we do not adjust CFSR solar radiation over the missing-data period for the SRB output used in CFSR$_{+s}$ and CFSR$_{+p,s}$ simulations, but do compare results over the entire time period.)

CFSR reanalysis output generally reproduces temperature well over the regions and seasons most relevant to

maize yields but shows biases in precipitation and solar radiation (Table 1 and supplementary Fig. S2). The 30-yr average national maximum and minimum temperatures in CFSR differ from AgCFSR by $\sim 0.4°C$, and mean temperatures follow this pattern. Annual temperatures in CFSR, however, can differ from AgCFSR by up to $\sim 2°C$. While CFSR accurately reproduces annual deviations in solar radiation from trend, it uniformly overestimates U.S. mean insolation by $\sim 20 \, \mathrm{W \, m^{-2}}$, or $\sim 10\%$ (Fig. S2). CFSR also underestimates 30-yr average U.S. precipitation by $\sim 6 \, \mathrm{cm}$ per growing season, or $\sim 15\%$. In an individual year, growing-season precipitation errors can reach 20 cm, or $\sim 50\%$ (Fig. 1a). AgCFSR improves on precipitation and solar radiation variables, with year-to-year national estimates nearly identical to observation-based products (Fig. S2).

Last, we use station measurements not as inputs to pDSSAT but to provide a ground-truth comparison of the precipitation products used in this study (CFSR

reanalysis, AgCFSR bias-corrected reanalysis, and CPC gridded station data). Because CPC and CRU products are interpolated and not original data, we compare with raw U.S. station measurements from the Hadley Centre Integrated Surface Database (HadISD), version 1.0.1.2012p (Dunn et al. 2012), where less than 1% of growing-season precipitation data are missing (totaling 157 of 1212 stations). Note that some (if not all) of these stations are likely also used to generate the CPC dataset, but precipitation estimates can differ because of station interpolations.

3. Analysis

a. Simulated yield means and interannual variations

The choice of climate inputs significantly affects mean simulated maize yields, with different choices altering national production-weighted yields by ~25% (8.1–10.1 t ha^{-1}; Table 1). Changes result primarily from differences in precipitation inputs; yields are negligibly sensitive to differences in insolation despite significant disparity between sources. (Little sensitivity for solar radiation means normalization across climate variables would have minimal effect.) Because we simulate fixed management practices, the absolute magnitudes of simulated yields are not directly useful for validation, but the yield differences from trend highlight the sensitivity of pDSSAT yields to climate inputs. [For direct comparison of pDSSAT output with absolute NASS yields, see Elliott et al. (2013).] Interannual variations in simulated national-average yields suggest again that the dominant factor driving yield differences is precipitation. (See Fig. 1 and supplementary Fig. S3 for results using Penman–Monteith and Priestley–Taylor estimates of evapotranspiration, respectively, and see Fig. S2 for equivalent figures of other climate variables.) Interannual variations in simulated yields closely follow interannual differences in precipitation inputs (correlation coefficients of 0.7–0.8).

The sensitivity to precipitation means that increased fidelity of precipitation inputs improves simulation estimates of interannual fluctuations in maize yield. Driving simulations with unaltered reanalysis output (CFSR) produces average absolute difference from detrended and normalized national NASS yields of ~1.30 t ha^{-1} (Fig. 1b). Using reanalysis combined with some form of observational precipitation reduces this average difference by 40%–60%: average differences from NASS in simulations using observation-based precipitation (CFSR$_{+s}$) and bias-corrected reanalysis (AgCFSR) are ~0.84 and ~0.53 t ha^{-1}, respectively. (For reference, the mean and standard deviation of detrended, normalized NASS yields are 7.9 and 0.7 t ha^{-1}; see

Fig. S1.) In a single year, using observation-based precipitation can reduce yield difference from NASS by more than 2 t ha^{-1} (over 20% of total yield), which when averaged over all U.S. production has a value of ~$10 billion in 2010 dollars (Fig. 1b, right axis). Differences in simulations driven by AgCFSR and CPC precipitation (i.e., AgCFSR and AgCFSR$_{+p}$) are small because precipitation inputs are similar. Note that in all cases, pDSSAT simulations produce stronger interannual variations in yield than is seen in historical data (see section 3d).

The benefits from using observation-based precipitation for agricultural impacts assessments become more evident when yields are considered at the county level (Fig. 2, which shows correlation coefficients between simulated and NASS interannual yield variations). Simulations driven with reanalysis inputs are unable to reproduce historical yield variations in the regions where most maize is grown (i.e., correlation coefficients are low in the Corn Belt). Precipitation plays an especially significant role in this region, since maize cultivation is almost entirely rainfed. (In contrast, cultivation west of the Corn Belt in, e.g., Nebraska, is largely irrigated. Irrigated yields, especially after detrending, show little interannual variation so correlations between modeled and NASS yields are generally low.) While substituting reanalysis with observation-based solar radiation has little impact, substituting reanalysis with observation-based precipitation leads to large increases in correlation coefficients across the United States (supplementary Fig. S4).

To demonstrate that changes in precipitation do in fact drive most of the improvements in simulated interannual yield variations, we compare differences in the various climate variables between each simulation scenario and the default CFSR with the resulting differences in yield. That is, we compare spatial patterns between differences in time-averaged county climate and yield. Differences in precipitation inputs are spatially correlated with differences in yield, as expected (Fig. 3). Observed precipitation is higher in the Corn Belt and lower in the Southeast than precipitation in CFSR, so using any observation-based precipitation source produces higher yields in the Corn Belt and lower yields in the Southeast. (See Fig. 3 for AgCFSR and supplementary Fig. S5 for CFSR$_{+p,s}$ cases.) Patterns of yield change differ from those of insolation and minimum temperature, suggesting that those variables play a less significant role. (Note that because CFSR solar radiation is consistently biased high, differences in precipitation and solar radiation may not correlate.) The fact that precipitation differences account for most of the differences in yield variation is borne out by
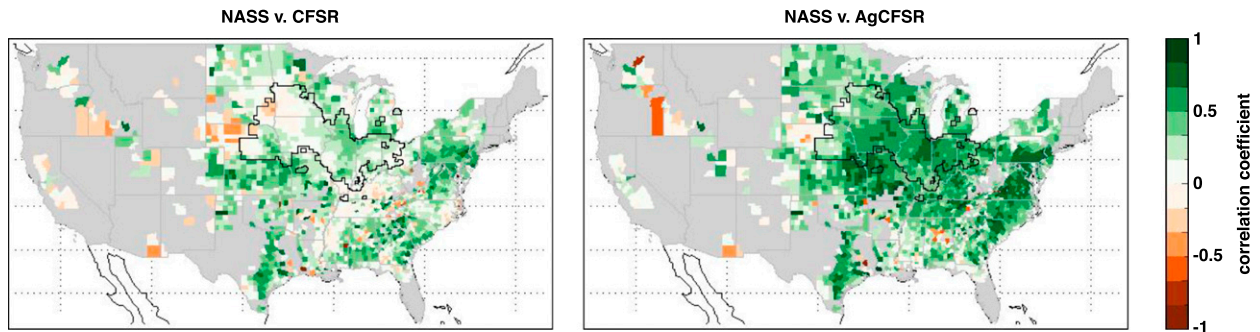
FIG. 2. Correlations between annual modeled and observed yield over years 1980–2009. Aggregation is at the county level, and gray regions indicate counties with <0.1% of land harvested for maize. (left) Correlations between CFSR and surveyed yield are neutral to weakly positive, especially in maize-growing regions (with black outline marking counties with ≥1/4 of land harvested for maize). (right) Bias correction significantly improves estimates in the Corn Belt, where most maize is rainfed.

examining coefficients of determination R^2 of 1) differences in a given climate variable between a simulation scenario and the default CFSR and 2) the resulting differences in yield. (Here, climates and yields are nationally averaged and weighted by maize-growing regions, dominated by rainfed yields in the Corn Belt.) Precipitation and yield differences are highly correlated with $R^2 \sim 0.8$ for all simulations (Table 2).

The influence of temperature is more difficult to quantify, since we do not consider scenarios where temperature alone is altered from reanalysis (because reanalysis temperatures tend to be more accurate). In the AgCFSR and AgCFSR$_{+p}$ simulations, both temperature and precipitation differ from CFSR. For these simulations, differences in maximum temperature are well correlated with differences in yield ($R^2 \sim 0.7$ in Table 2), but this relationship may not be causal, as differences in maximum temperature are spatially correlated with those in precipitation (Fig. 3). It is not possible from these simulations alone to disentangle the relative importance of precipitation and temperature inputs for maize yields, since both inputs differ across simulations. We can, however, isolate and assess the importance of temperature inputs by comparing the CFSR$_{+p,s}$ and AgCFSR$_{+p}$ simulations, since these essentially only differ in temperature (see section 3b). In the national production-weighted yield time series shown in Fig. 1, the bias-corrected temperatures in AgCFSR$_{+p}$ reduce the average absolute differences from national NASS yields by $\sim$0.28 t ha^{-1}, or $\sim$30% [i.e., (ΔCFSR$_{+p,s}$ − ΔAgCFSR$_{+p}$)/ΔCFSR$_{+p,s}$, where Δ indicates the time average of the absolute differences from NASS]. Most of that effect appears related to differences in maximum temperature ($R^2 \sim 0.7$ in Table 2) rather than minimum temperature ($R^2 < 0.2$). (The greater importance of maximum than minimum temperature is consistent with

spatial patterns in Fig. 3.) Higher correlation with maximum temperature is not surprising. Maximum temperatures typically occur during the day when photosynthesis is at its peak. Maximum temperatures also appear to drive high correlations for mean temperature, which largely control crop development.

b. Rainfall distribution

Because precipitation appears the dominant climate factor driving uncertainty in yield estimates, it is worth considering which aspects of precipitation datasets are most significant for yields. In our study we drive pDSSAT with two observation-based precipitation sources: 1) CPC gridded rain gauge measurements (AgCFSR$_{+p}$), and 2) AgCFSR bias-corrected reanalysis precipitation (AgCFSR). CPC and AgCFSR have nearly identical rainfall averages (Fig. 1a), but distributions are quite different (supplementary Fig. S6a). That distinction allows us to evaluate the relative importance of precipitation means and distributions for making yield estimates.

We compare the rainfall distributions in CPC and AgCFSR with those of 168 U.S. weather stations described in section 2b, using the number of rainy days as our metric of comparison (Fig. 4). If station data are taken as ground truth (averaging $\sim$40 rainy days per growing season), then the bias-corrected AgCFSR tends to underestimate precipitation frequency (by $\sim$8 rainy days per growing season). (For comparison, the non-bias-corrected CFSR has the well-known excess "drizzle problem" and overestimates precipitation frequency by $\sim$26 rainy days per growing season; see supplementary Fig. S7). The observation-based CPC (version 1.0) precipitation dataset also has a large drizzle problem and overestimates precipitation frequency by $\sim$22 rainy days per growing season. The discrepancy likely arises in the distance-weighting
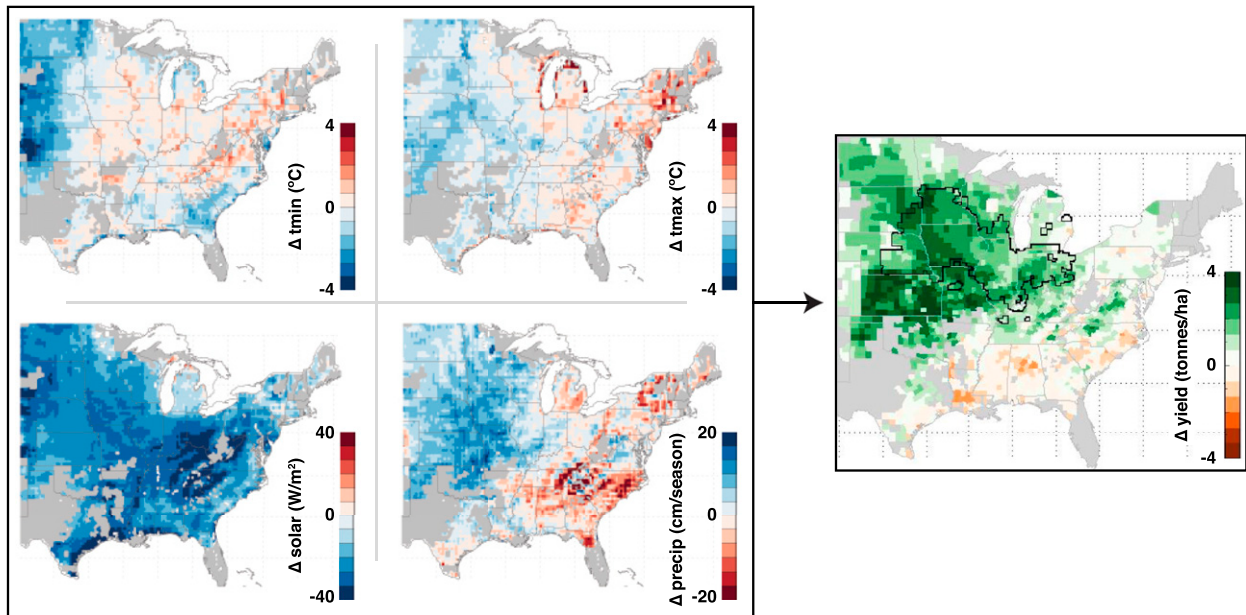
FIG. 3. Differences in 1980–2009 growing-season-averaged bias-corrected and original reanalysis climate indices (AgCFSR − CFSR) and resulting yields. Yields here are not detrended or normalized. The Corn Belt is outlined in black, where most maize is rainfed. Gray regions denote counties with <0.1% of land harvested for maize. Spatial patterns in yield differences are similar to precipitation differences; a wetter AgCFSR Midwest correlates with higher yields. Yield differences are inconsistent with minimum temperature and solar radiation distributions.

interpolation used by CPC to convert spatially heterogeneous station data to a uniform grid, which smears rainfall events across stations. The technique has been shown to produce erroneous assumptions of many days with low-intensity precipitation (Ensor and Robeson 2008), and the CPC precipitation distribution is consistent with that of spatially averaged station data (Fig. S6b). As a consequence, CPC and AgCFSR precipitation distributions differ on average by ~30 rainy days, one-quarter of the growing season.

Simulated U.S. maize yields are, however, largely insensitive to rainfall distributions. For aggregate U.S. yields, the distinction between AgCFSR and AgCFSR$_{+p}$ does not seem relevant (Table 1 and Fig. 1). However, there may be situations where the differences are important. We therefore evaluate the sensitivity of yields to the number of rainy days for subsets of the data with different total growing-season precipitation (Fig. 5). (Total precipitation scales monotonically with the number of rainy days.) In cases where growing-season precipitation exceeds 20 cm, only a negligible fraction ($R^2 < 0.05$) of yield variation between AgCFSR and AgCFSR$_{+p}$ can be explained by differences in estimated number of rainy days. If total growing-season precipitation is less than 20 cm, the number of rainy days becomes significant ($R^2 > 0.6$). Aggregate national yields are largely unaffected by differences in precipitation distributions because most U.S. counties

experience at least 20 cm of precipitation per growing season (Fig. 5, red line). Our results suggest that agricultural assessments of more arid regions are likely more sensitive to precipitation inputs.

c. Estimating extremes

The sensitivity of mean yields to accurate precipitation inputs may have implications for reproducing extreme events. We calculate the probability of a

TABLE 2. Coefficients of determination of differences between each simulation and CFSR in 1980–2009 national average yield and climate. National averages are production weighted, and yields are not detrended or normalized. The R^2 values can be thought of as the fraction of yield change variation that can be described by changes in climate inputs from standard CFSR; these are not shown where simulations use the same climate source (e.g., CFSR and CFSR$_{+p}$ use the same temperature and solar radiation sources). Changing growing-season precipitation explains most of the variation in yield change. Correlations between temperature/precipitation and solar radiation inflate AgCFSR and AgCFSR$_{+p}$ solar radiation R^2 values.

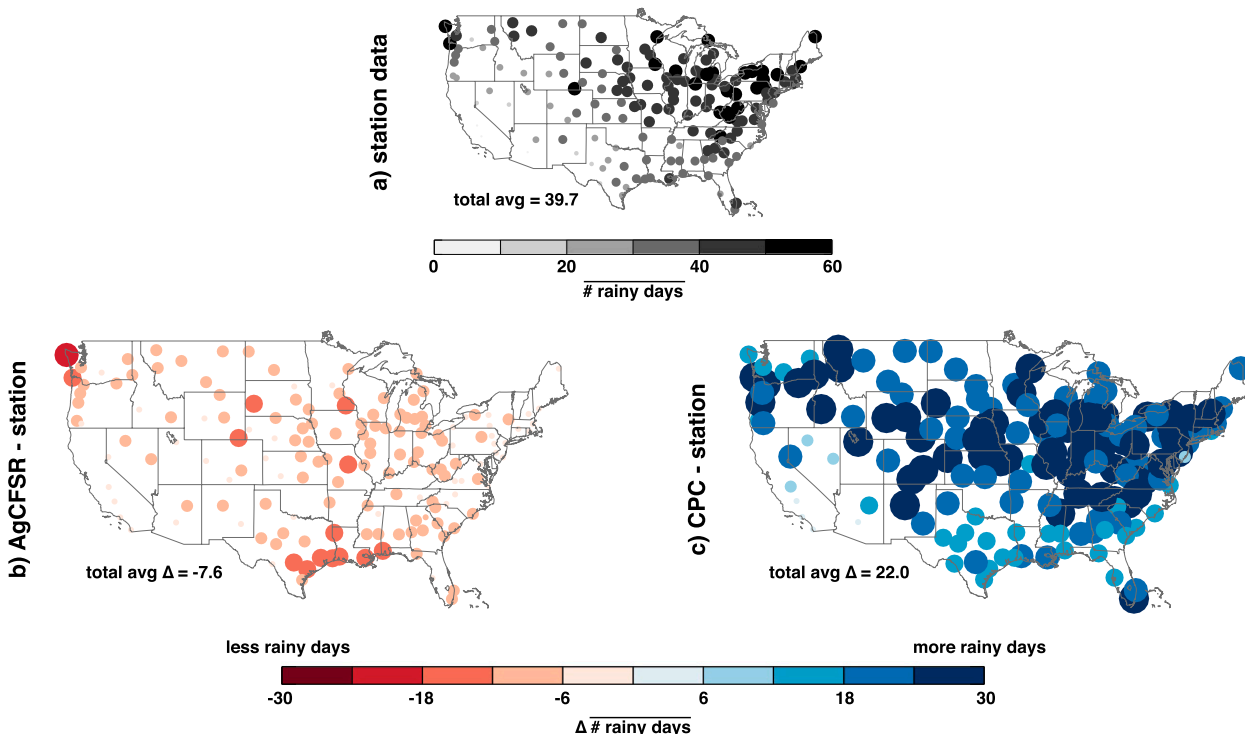|	Δyield variation				
Simulation	ΔT_{mean}	ΔT_{max}	ΔT_{min}	Δprecipitation	Δsolar
CFSR$_{+s}$	—	—	—	—	0.005
CFSR$_{+p}$	—	—	—	0.841	—
CFSR$_{+p,s}$	—	—	—	0.789	0.025
AgCFSR	0.568	0.670	0.143	0.825	0.337
AgCFSR$_{+p}$	0.572	0.662	0.163	0.786	0.354

FIG. 4. Mean growing-season rainy days from station observations for 1980–2009, and differences from AgCFSR/CPC datasets. Rainy days are tallied for daily precipitation ≥0.1 mm over the growing season as defined by AgCFSR yield simulations. Dots are colored and scaled by magnitude. (a) Stations are selected only where ≤1% of growing-season rainfall is missing; the average across all 157 stations is ~40 rainy days. (b) AgCFSR underestimates the number of rainy days, and (c) CPC overestimates the number of rainy days. Average error of CPC is about triple that of AgCFSR.

simulation to detect a year containing one of the worst five yields among the 30 years from 1980 to 2009 NASS county data (hereinafter bottom-five year). We assess only the ability of a model simulation to detect the occurrence of a bottom-five year and disregard the order of bottom-five year and the magnitude of yield loss. We do not expect any model to perfectly detect all extreme events. To validate results from this bottom-five approach, we also calculate the probability of detecting yield events more than one standard deviation from the 30-yr mean (see the supplementary material for a discussion). Our model can capture large-scale maize reductions attributed to droughts and heat waves (Elliott et al. 2013); numerous other factors (e.g., floods and pests) can also cause, or contribute to, large-scale losses.

We find that crop yield simulations driven by observation-based precipitation are better able to capture extreme losses in yield. On average, simulating yields with observation-based rather than reanalysis precipitation improves the ability to detect a bottom-five year by >10%, with CPC providing a slight advantage over AgCFSR (Fig. 6a). Spatially, simulating yields with observation-based AgCFSR rather than reanalysis improves extreme

event detection in the Midwest by up to ~60% (Fig. 6b). This improved performance holds only in rainfed maize areas, suggesting again that precipitation is the dominant climate factor.

We also separate extreme events by year to more closely analyze drivers of yield loss and specific large-scale historical events (Fig. 6c). All simulations correctly identify the drought of 1988 as a bottom-five year, but those driven with reanalysis output do not identify the 1983 Midwest drought. (The drought in 1988 was much more severe and widespread than the drought in 1983.) Observation-based precipitation provides similar (but smaller) improvements in the less damaging droughts of 1991 and 2002. Detection benefits are due to precipitation alone; use of observation-based solar radiation or maximum–minimum temperature has little consequence for detecting any single extreme event. (See Fig. 6c; for precipitation compare CFSR and CFSR$_{+p}$, for solar radiation compare CFSR and CFSR$_{+s}$, and for temperature compare CFSR$_{+p,s}$ and AgCFSR$_{+p}$.) These results are robust to the definition of extreme loss (see supplementary Fig. S8 for an analogous analysis defining extremes as yields of more than one standard deviation from the county mean). Note that no simulation
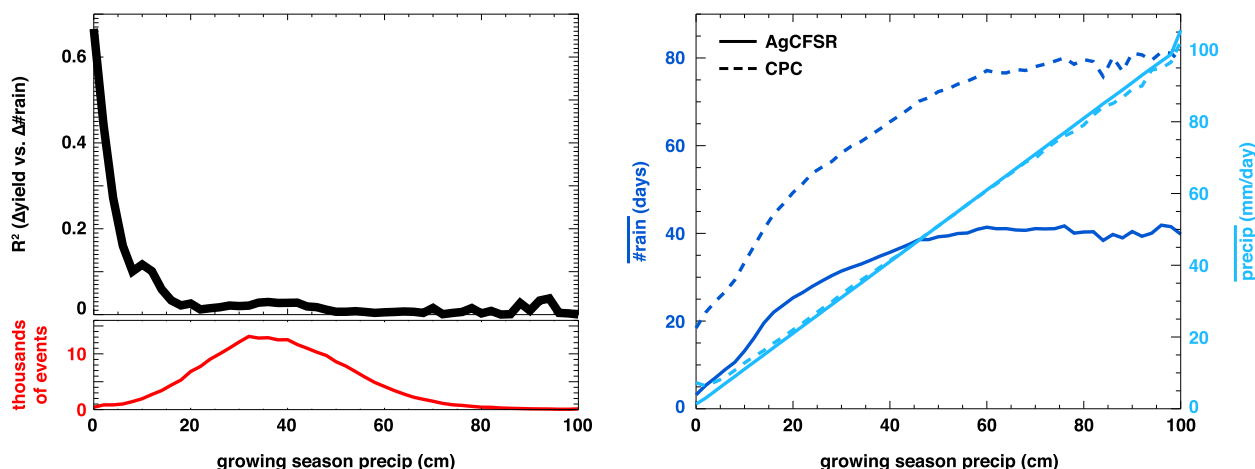
FIG. 5. Coefficients of determination of differences in growing-season county yield and number of rainy days in AgCFSR and AgCFSR$_{+p}$. AgCFSR$_{+p}$ uses CPC precipitation; all other climate variables are identical to AgCFSR. Counties are binned by AgCFSR growing-season precipitation in 2-cm increments for the eastern United States only (where each year within a given county is considered to be a separate "event"). Yield is defined as potential rainfed yield to highlight the model response to rainy days. Total rainfall between both simulations is similar (light blue). Differences in number of rainy days describes most of the differences in local annual yield only when total rainfall is low (black), even though differences in estimated number of rainy days are largest when total rainfall is high (dark blue). National average yields in AgCFSR and AgCFSR$_{+p}$ are insensitive to differences in the number of rainy days because most counties experience rainfall totals of approximately 20–60 cm season^{-1} (red).

captures the extreme events in 1993 and 1995. Losses in 1993 were due to flooding, which is poorly represented in DSSAT, and those in 1995 resulted from a complex series of non-precipitation-related factors.

d. Fundamental issues with model response

Crop models, like climate models, may have systematic biases in yield estimations. In the simulations discussed here we assume fixed management without technological trends, and therefore inherit trend biases in our yield simulations. To compare modeled and observed yields on similar terms, we detrend and normalize NASS and modeled yields. Examination of the detrended time series suggests, however, that the crop model simulations also include some type of variance bias, since crop model output appears overly sensitive to interannual climate variability (Fig. 1b). All simulations overestimate the magnitude of yield changes in anomalous years, even after aggregating to national levels where finescale errors are likely to cancel. This oversensitivity is clearest when large-scale drought reduces yield. In the most damaging drought in 1988, simulated national yield reductions from the trend line are about twice as large as actual reported losses in NASS yields.

To investigate the nature of the model response to changes in climate and to test results under more realistic representations of interannual variability, we arbitrarily adjust the interannual variance in yields in each simulation to match that of NASS. [For comparison of variability in modeled and NASS yields, see Table 3,

column 1; for variance-adjustment methodology, see Eq. (1).] This manipulation is purely exploratory: we do not evaluate different variance-adjustment methodologies, nor do we seek to make recommendations. It is informative to note, however, that applying a simple variance adjustment improves yield estimates in extreme years (Figs. 7 and S3b). Variance adjustment roughly halves differences in 30-yr average simulated national yields relative to NASS yields (Table 3, column 2). For the AgCFSR simulation, variance correction lowers mean yield error to ~0.4 t ha^{-1}, or ~5% of mean normalized NASS yields. Variance adjustment also reduces yield differences between model simulations. After variance adjustment, the intermodel spread of simulated yield averages is only ~0.1 t ha^{-1}. Therefore, for the 1988 drought, variance adjustment almost entirely eliminates errors by all simulations in the magnitude of national yield reduction. Regardless of the absolute size of yield variance, results stand that observed precipitation inputs improve correlations between simulated and NASS yields (Table 3, column 3).
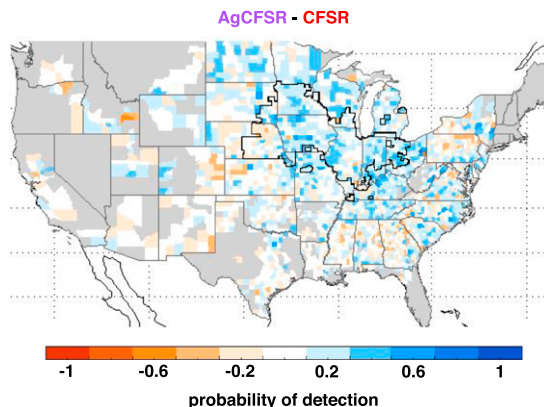
4. Discussion

a. Evaluating reanalysis

Our results suggest that driving crop models with non-bias-corrected reanalyses cannot reliably reproduce historical observational yields, confirming findings from previous studies (van Wart et al. 2013; Berg et al. 2010;

a) probability of detection aggregate
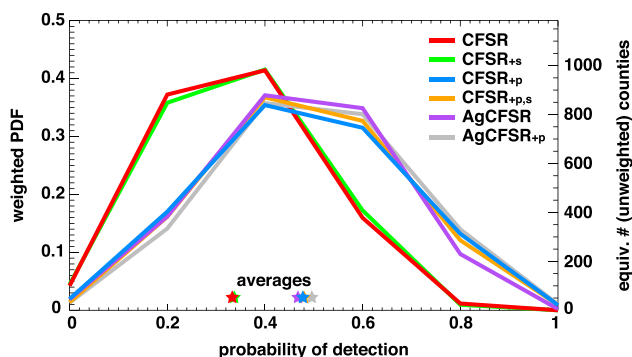
b) county difference in probability of detection
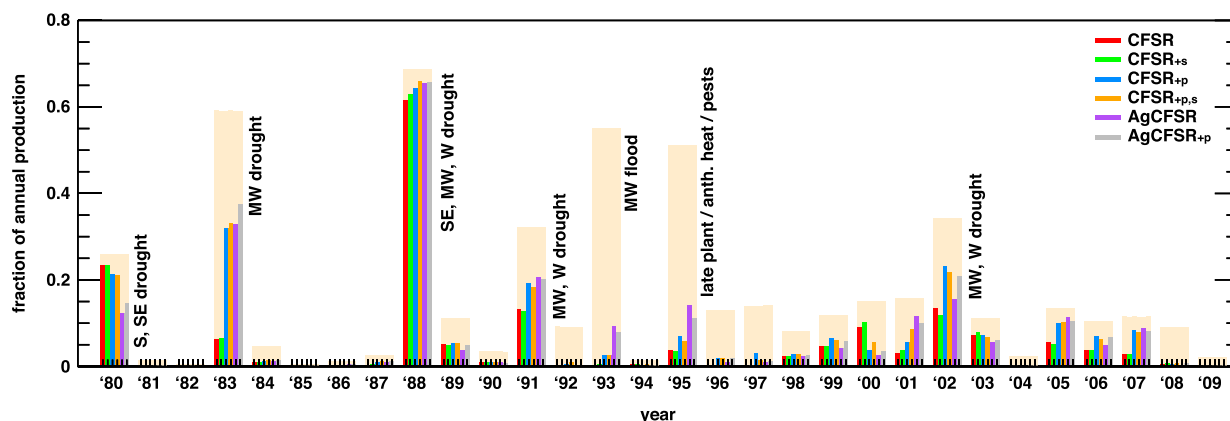

c) probability of detecting extreme years

FIG. 6. Probability of detection of bottom-five yielding years over the 30-yr period for each county in the NASS observational record. For comparison to NASS, all yields are detrended. Here we identify the number of years in each county that a yield simulation correctly identifies as a bottom-five yielding year (where bottom-five year order is unimportant). (a) PDF of all counties weighted by 30-yr average NASS production, and equivalent number of unweighted counties on right axis. (b) AgCFSR accurately identifies 1–2 more bottom-five yielding years than CFSR in the Corn Belt region (outlined in black), where most maize is grown. (c) Bottom-five yielding years for all counties, production weighted and binned by year. Tan boxes denote the total counties identified as a bottom-five yield year in the NASS record, and the non-tan colored boxes denote correct detections for each model simulation. Perfect detection would mean tan and colored boxes are the same. Simulations driven with improved precipitation estimates are better able to detect yield reductions caused by a Midwest drought.

Challinor et al. 2005). At the county level, year-to-year variations in reanalysis-driven maize yield simulations are largely uncorrelated with variations in surveyed yields (Fig. 2). At the national level, using reanalysis produces errors in production estimates of up to ~30% of the 2010 maize production value (Fig. 1). These errors are likely a lower limit in the context of a global assessment. Reanalyses and bias-correction datasets may have lower accuracy in the developing world where constraining observational datasets are less available (Ruane et al. 2014a).

While reanalyses poorly capture interannual variations in patterns of both precipitation and insolation (Figs. 1a and S2), only improvements in precipitation appear to be significant for yield estimates (Figs. 1b, S3,

S4, and Table 3). CFSR insolation is biased ~10% high (for national average weighted by maize production), but substituting observation-based values in crop simulations negligibly alters maize yields. CFSR precipitation is biased on average ~14% low, and substituting observation-based values in crop simulations increases yields significantly. [Note, however, that other reanalyses may better constrain precipitation (e.g., Reichle et al. 2011).] Simulating yields with observation-based rather than reanalysis precipitation also improves the ability to detect extreme years with high crop losses (Fig. 6). Extreme droughts are linked to some of the most severe large-scale losses in U.S. maize since 1980 (1983, 1988, and the recent drought of 2012, not simulated here). The fidelity of models in capturing these anomalous

TABLE 3. Variance comparison between modeled and NASS survey national yields for years 1980–2009. CV ratio = $(\sigma_{model}/\sigma_{NASS}) \times (\mu_{NASS}/\mu_{model})$, where σ_{NASS} is defined as the variance of the detrended sample. After variance adjustment, all modeled yields have similar average errors from NASS. Explained variance R^2 metrics (which do not change with variance adjustment) show large improvements in yield estimates when substituting reanalysis with observation-based precipitation.

Simulation	CV ratio	Avg|error|[a]	R^2
CFSR	2.63	0.55	0.18
CFSR$_{+s}$	2.60	0.54	0.20
CFSR$_{+p}$	1.88	0.45	0.38
CFSR$_{+p,s}$	1.85	0.45	0.39
AgCFSR	1.36	0.43	0.47
AgCFSR$_{+p}$	1.35	0.42	0.49

[a] Units: metric tons per hectare.

historical events may be a guide to their reliability in projecting yield changes in future scenario simulations under changing climate. Note that the quantitative benefit of improved precipitation shown here is robust despite the oversensitivity of the crop model to interannual climate variations (Table 3, column 3).

We also show that results are insensitive to the representation of modeled evapotranspiration. Use of either the Penman–Monteith method or the Priestley–Taylor method produces similar results. (See Fig. 1 for results using the Penman–Monteith method and Fig. S3 for results using the Priestley–Taylor method.) Priestley–Taylor simplifies the Penman–Monteith representation by negating the need for wind observations yet produces yields that are nearly as reliable in reproducing interannual variations. This consistency suggests that accurate wind measurements are likely a second-order concern for U.S. agricultural assessments.

b. Incorporating climate observations

One alternative to reanalysis is bias-corrected reanalysis, which improves yield estimates significantly (Figs. 1 and 2). Previous studies also find improvements when bias correcting reanalysis (Challinor et al. 2005; Berg et al. 2010). [Similarly, de Wit et al. (2010) find no significant difference in yield estimates driven by partially bias-corrected reanalysis and interpolated station data but do not test the case of bias correcting precipitation variables or compare with yields driven by non-bias-corrected reanalysis.] While bias correction does rely on observational data, it can make use of monthly rather than daily data. For the purpose of agricultural applications, broader coverage of incomplete but accessible stations is likely more useful than a limited network of complete daily stations, as this would allow sufficient overlap to enable gap-filling approaches

using the types of blended model and observational datasets that have proven successful here.

Translating station observations to inputs suitable for gridded crop models can also introduce interpolation errors (Lobell 2013), but we find that yield estimates are mostly insensitive to these errors. We show that CPC rainfall frequency errors (caused by distance-weighting interpolation) only marginally affect U.S. maize yield estimates (Figs. 4 and S6). [See also Ensor and Robeson (2008) for assessment of CPC rainfall frequency errors.] Berg et al. (2010), however, find that rainfall frequency errors are important for estimating western Africa millet yields. We expand on these results and suggest that rainfall frequency errors may be important only in arid regions or seasons with severe drought (Fig. 5). For assessments in arid regions or seasons with severe drought, precipitation products like CPC may therefore benefit from an additional step in their interpolation schemes to adjust rainfall frequencies to better match observations. In fact, a more recent version of CPC uses the updated optimal interpolation algorithm to improve estimates from using the Cressman interpolation methodology (Xie et al. 2007).

c. Crop model limitations

Unlike studies that do not compare crop model output to observed yields (e.g., van Wart et al. 2013), we can distinguish errors related to the model representation of yield response from those related to climate inputs. Our results indicate that yield estimates suffer from problems related to the crop models themselves. In our experiments, pDSSAT is oversensitive to interannual changes in growing conditions (Fig. 7 and Table 3), likely meaning that pDSSAT exaggerates yield responses to precipitation changes in rainfed regions. (Both simulated and observed interannual yield variability are small in highly irrigated regions.) This response may be due to structural limitations but more likely results from simplified management inputs. In our experiment, planting dates are fixed within a narrow time window. In reality, farmers may respond to year-to-year weather or may minimize risk by diversifying planting dates and seed types at subfield scales to mitigate harms from extreme events. Note that we do not believe the oversensitivity in our pDSSAT output is due to excessively high row densities, which have been shown to increase crop sensitivity to drought (Lobell et al. 2014). First, we assume a relatively low crop density (5 plants per meter squared); second, crop yield responses to both high and low precipitation are amplified. Understanding crop responses to management changes, as well as methods to represent management must become a research priority.

a) detrended, normalized national yield
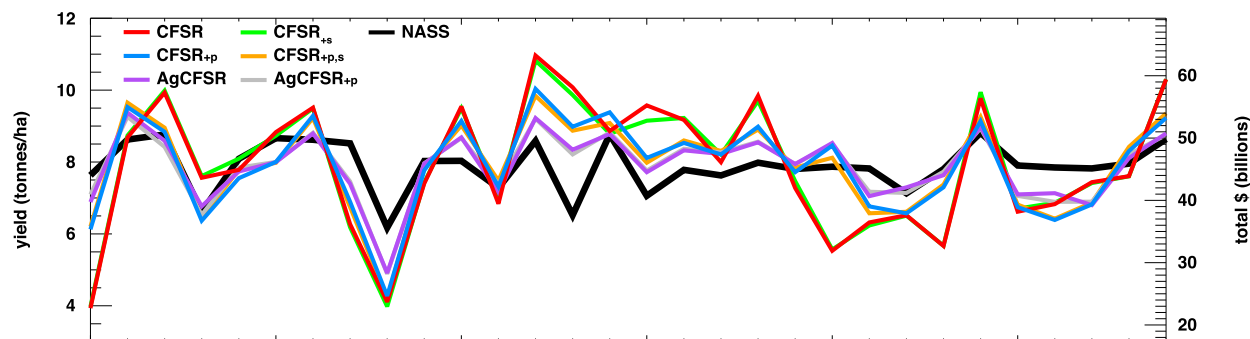
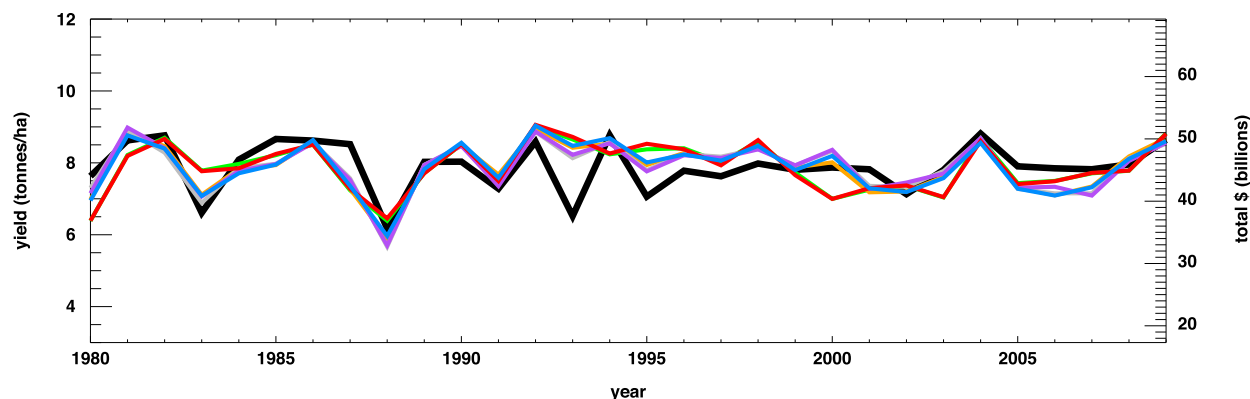### b) detrended, normalized, variance-adjusted national yield

FIG. 7. Times series plots of 1980–2009 U.S. modeled and NASS survey maize yields. (a) Yields are normalized and detrended to remove technological and management changes present in the survey data. (b) As in (a), plus yields are variance adjusted to remove errors in the modeled yields' sensitivity to changes in weather. [See Eq. (1) for variance-adjustment methodology.] Note that (a) is identical to Fig. 1b and is shown here only for comparison. Variance-adjusting model output significantly reduces differences between simulated and observed yields and between simulation scenarios.

Our simulations also point out several additional shortcomings in crop growth representation. All yield simulations fail to capture extreme events in 1993 and 1995 (Fig. 6c). In 1993, extreme precipitation throughout the Midwest led to floods and waterlogged soils, resulting in root death and low yields (Rosenzweig et al. 2002). Capturing flood effects in a point-based model is difficult in the absence of information about the wider basin hydrographs or damage from submergence and rushing water. It is therefore not surprising that all simulations overestimate 1993 yields. In 1995, a number of factors caused low yields: early spring flooding led to late planting and a shorter growing season, the timing of anthesis coincided with extreme heat, and corn borer pests were widespread (Elmore and Taylor 2013; Ostlie et al. 1997). Because we used fixed planting periods, we do not capture the shortened growing season and subsequently miss damages from the heat event near the crucial flowering stage. Resolving pest damages would have required a coupled or integrated pest model, which

was not included in this study. The 1993 and 1995 biases again point out the importance of using more realistic dynamic management inputs, and also the need to improve intrinsic issues in the representation of flood and pest damage.

5. Conclusions

Our results suggest that care must be taken to avoid compromising historical agricultural validations by any of several factors: precipitation inputs, management inputs, and model limitations in representing effects from floods, pests, and extreme heat. We show that bias correcting precipitation is important, but the data needs for bias correction are not overly taxing, since correction methods can make use of monthly climatologies instead of daily observations. When our results are combined with previous studies, the need for bias correcting precipitation appears relatively robust across crop models and global regions

(Challinor et al. 2005; Berg et al. 2010). International collaborative efforts are ongoing to further evaluate available weather products and identify appropriate correction/aggregation techniques for this purpose (Elliott et al. 2015; Ewert et al. 2015). Improving management inputs and reducing model limitations is a longer-term project, but efforts are also ongoing here. Simulation of damages from anthesis heat waves and pests (such as those experienced in 1995) are areas of particular focus for model improvement in AgMIP (e.g., Asseng et al. 2015).

Historical validations suggest that model errors could compromise future projections. Multimodel means from phase 5 of the Coupled Model Intercomparison Project (CMIP5) archive forecast higher annual average precipitation in the U.S. Corn Belt and robustly predict increases in precipitation intensity (IPCC 2013). Making agricultural projections in such a wetter future using a model that is oversensitive to precipitation changes (and undersensitive to flooding damages) would lead to overestimates of future yields. DSSAT, the focus of the study presented here, is one of most widely used agricultural models, but many others are also commonly used to make future projections [e.g., Agricultural Production Systems Simulator (APSIM), Environmental Policy Integrated Climate (EPIC), or Lund-Potsdam-Jena Managed Land (LPJmL)]. It is critically important to similarly assess their performance. Some of this assessment is currently under way in AgMIP, which includes efforts to evaluate multimodel sensitivities and to improve the climatological responses of modeled maize (Bassu et al. 2014; Ruane et al. 2014b; McDermid et al. 2015). These continued efforts are especially timely because reliable crop growth representations are vital for decisions on policy and adaptive responses to future climate conditions.

REFERENCES

Asseng, S., and Coauthors, 2013: Uncertainty in simulating wheat yields under climate change. *Nat. Climate Change*, **3**, 827–832, doi:10.1038/nclimate1916.

——, and Coauthors, 2015: Rising temperatures reduce global wheat production. *Nat. Climate Change*, **5**, 143–147, doi:10.1038/nclimate2470.

Bassu, S., and Coauthors, 2014: How do various maize crop models vary in their responses to climate change factors? *Global Change Biol.*, **20**, 2301–2320, doi:10.1111/gcb.12520.

Berg, A., B. Sultan, and N. de Noblet-Ducoudré, 2010: What are the dominant features of rainfall leading to realistic large-scale crop yield simulations in West Africa? *Geophys. Res. Lett.*, **37**, L05405, doi:10.1029/2009GL041923.

Bosilovich, M. G., D. Mocko, J. O. Roads, and A. Ruane, 2009: A multimodel analysis for the Coordinated Enhanced Observing Period (CEOP). *J. Hydrometeor.*, **10**, 912–934, doi:10.1175/2009JHM1090.1.

Challinor, A., T. Wheeler, J. Slingo, P. Craufurd, and D. Grimes, 2005: Simulation of crop yields using ERA-40: Limits to skill and nonstationarity in weather-yield relationships. *J. Appl. Meteor.*, **44**, 516–531, doi:10.1175/JAM2212.1.

Cressman, G. P., 1959: An operational objective analysis system. *Mon. Wea. Rev.*, **87**, 367–374, doi:10.1175/1520-0493(1959)087<0367:AOOAS>2.0.CO;2.

de Wit, A., and Coauthors, 2010: Using ERA-INTERIM for regional crop yield forecasting in Europe. *Climate Res.*, **44**, 41–53, doi:10.3354/cr00872.

Dunn, R. J. H., K. M. Willett, P. W. Thorne, E. V. Woolley, I. Durre, A. Dai, D. E. Parker, and R. E. Vose, 2012: HadISD: A quality controlled global synoptic report database for selected variables at long-term stations from 1973–2011. *Climate Past*, **8**, 1649–1679, doi:10.5194/cp-8-1649-2012.

Elliott, J., and Coauthors, 2013: Predicting agricultural impacts of large-scale drought: 2012 and the case for better modeling. RDCEP Working Paper 13-01, 8 pp., doi:10.2139/ssrn.2222269.

——, D. Kelly, J. Chryssanthacopoulos, M. Glotter, K. Jhunjhnuwala, N. Best, M. Wilde, and I. Foster, 2014: The parallel system for integrating impact models and sectors (pSIMS). *Environ. Modell. Software*, **62**, 509–516, doi:10.1016/j.envsoft.2014.04.008.

——, and Coauthors, 2015: The Global Gridded Crop Model Intercomparison (GGCMI): Data and modeling protocol for phase 1 (v1.0). *Geosci. Model Dev.*, **8**, 261–277, doi:10.5194/gmd-8-261-2015.

Elmore, R. W., and S. E. Taylor, 2013: Analog years for weather forecasting and correlating corn planting dates with yield in Iowa. Iowa State University Extension, Integrated Crop Management News. [Available online at http://crops.extension.iastate.edu/cropnews/2013/05/analog-years-weather-forecasting-and-correlating-corn-planting-dates-yield-iowa.]

Ensor, L. A., and S. M. Robeson, 2008: Statistical characteristics of daily precipitation: Comparisons of gridded and point data-sets. *J. Appl. Meteor. Climatol.*, **47**, 2468–2476, doi:10.1175/2008JAMC1757.1.

Ewert, F., and Coauthors, 2015: Uncertainties in scaling-up crop models for large-area climate change impact assessments. *Handbook of Climate Change and Agroecosystems: The Agricultural Model Intercomparison and Improvement Project (AgMIP) Integrated Crop and Economic Assessments*, C. Rosenzweig and D. Hillel, Eds., ICP Series on Climate Change Impacts, Adaptation, and Mitigation, Vol. 3, World Scientific, 261–277.

Feng, S., M. Oppenheimer, and W. Schlenker, 2012: Climate change, crop yields, and internal migration in the United States. National Bureau of Economic Research Working Paper 17734, 43 pp. [Available online at http://www.nber.org/papers/w17734.pdf.]

Harris, I., P. Jones, T. Osborn, and D. Lister, 2014: Updated high-resolution grids of monthly climatic observations—The CRU TS3.10 dataset. *Int. J. Climatol.*, **34**, 623–642, doi:10.1002/joc.3711.

Hatfield, J. L., K. J. Boote, B. Kimball, L. Ziska, R. C. Izaurralde, D. Ort, A. M. Thomson, and D. Wolfe, 2011: Climate impacts on agriculture: Implications for crop production. *Agron. J.*, **103**, 351–370, doi:10.2134/agronj2010.0303.

Higgins, R., V. Kousky, V. Silva, E. Becker, and P. Xie, 2010: Intercomparison of daily precipitation statistics over the United States in observations and in NCEP reanalysis products. *J. Climate*, **23**, 4637–4650, doi:10.1175/2010JCLI3638.1.

Hsu, K.-L., X. Gao, S. Sorooshian, and H. V. Gupta, 1997: Precipitation estimation from remotely sensed information using artificial neural networks. *J. Appl. Meteor.*, **36**, 1176–1190, doi:10.1175/1520-0450(1997)036<1176:PEFRSI>2.0.CO;2.

Huffman, G. J., and Coauthors, 2007: The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeor.*, **8**, 38–55, doi:10.1175/JHM560.1.

IPCC, 2013: *Climate Change 2013: The Physical Science Basis*. Cambridge University Press, 1535 pp., doi:10.1017/CBO9781107415324.

Jones, J., and Coauthors, 2003: The DSSAT cropping system model. *Eur. J. Agron.*, **18** (3–4), 235–265, doi:10.1016/S1161-0301(02)00107-7.

Joyce, R. J., J. E. Janowiak, P. A. Arkin, and P. Xie, 2004: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeor.*, **5**, 487–503, doi:10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2.

Lobell, D. B., 2013: Errors in climate datasets and their effects on statistical crop models. *Agric. For. Meteor.*, **170**, 58–66, doi:10.1016/j.agrformet.2012.05.013.

——, M. J. Roberts, W. Schlenker, N. Braun, B. B. Little, R. M. Rejesus, and G. L. Hammer, 2014: Greater sensitivity to drought accompanies maize yield increase in the US Midwest. *Science*, **344**, 516–519, doi:10.1126/science.1251423.

McDermid, S., and Coauthors, 2015: The AgMIP Coordinated Climate-Crop Modeling Project (C3MP): Methods and protocols. *Handbook of Climate Change and Agroecosystems: The Agricultural Model Intercomparison and Improvement Project (AgMIP) Integrated Crop and Economic Assessments*, C. Rosenzweig and D. Hillel, Eds., ICP Series on Climate Change Impacts, Adaptation, and Mitigation, Vol. 3, World Scientific, 191–220, doi:10.1142/9781783265640_0008.

Mo, K. C., L. N. Long, Y. Xia, S. Yang, J. E. Schemm, and M. Ek, 2011: Drought indices based on the Climate Forecast System Reanalysis and ensemble NLDAS. *J. Hydrometeor.*, **12**, 181–205, doi:10.1175/2010JHM1310.1.

Nachtergaele, F., and Coauthors, 2009: *Harmonized World Soil Database (version 1.1)*. FAO and IIASA, 43 pp. [Available online at http://www.fao.org/fileadmin/templates/nr/documents/HWSD/HWSD_Documentation.pdf.]

Nelson, G. C., and Coauthors, 2014: Agriculture and climate change in global scenarios: Why don't the models agree. *Agric. Econ.*, **45**, 85–101, doi:10.1111/agec.12091.

OECD/FAO, 2012: OECD-FAO agricultural outlook 2012. Organisation for Economic Co-operation and Development and U.N. Food and Agricultural Organization Rep., 286 pp., doi:10.1787/agr_outlook-2012-en.

Ostlie, K., W. Hutchison, and R. Hellmich, Eds., 1997: Bt corn and European corn borer. University of Minnesota Extension Service, NCR Publ. 602, 20 pp. [Available online at http://www.extension.umn.edu/agriculture/corn/pest-management/bt-corn-and-european-corn-borer/.]

Porter, J. R., and Coauthors, 2014: Food security and food production systems. *Climate Change 2014: Impacts, Adaptation, and Vulnerability, Part A: Global and Sectoral Aspects*, C. B. Field et al., Eds., Cambridge University Press, 485–533.

Reichle, R. H., R. D. Koster, G. J. De Lannoy, B. A. Forman, Q. Liu, S. P. Mahanama, and A. Touré, 2011: Assessment and enhancement of MERRA land surface hydrology estimates. *J. Climate*, **24**, 6322–6338, doi:10.1175/JCLI-D-10-05033.1.

Rodell, M., and Coauthors, 2004: The Global Land Data Assimilation System. *Bull. Amer. Meteor. Soc.*, **85**, 381–394, doi:10.1175/BAMS-85-3-381.

Rosenzweig, C., F. N. Tubiello, R. Goldberg, E. Mills, and J. Bloomfield, 2002: Increased crop damage in the US from excess precipitation under climate change. *Global Environ. Change*, **12**, 197–202, doi:10.1016/S0959-3780(02)00008-0.

——, and Coauthors, 2013: The Agricultural Model Intercomparison and Improvement Project (AgMIP): Protocols and pilot studies. *Agric. For. Meteor.*, **170**, 166–182, doi:10.1016/j.agrformet.2012.09.011.

Ruane, A. C., R. Goldberg, and J. Chryssanthacopoulos, 2014a: Climate forcing datasets for agricultural modeling: Merged products for gap-filling and historical climate series estimation. *Agric. For. Meteor.*, **200**, 233–248, doi:10.1016/j.agrformet.2014.09.016.

——, S. McDermid, C. Rosenzweig, G. A. Baigorria, J. W. Jones, C. C. Romero, and L. DeWayne Cecil, 2014b: Carbon–temperature–water change analysis for peanut production under climate change: A prototype for the AgMIP Coordinated Climate-Crop Modeling Project (C3MP). *Global Change Biol.*, **20**, 394–407, doi:10.1111/gcb.12412.

Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015–1057, doi:10.1175/2010BAMS3001.1.

Schneider, U., A. Becker, P. Finger, A. Meyer-Christoffer, B. Rudolf, and M. Ziese, 2011: GPCC full data reanalysis version 6.0 at 0.5°: Monthly land-surface precipitation from rain-gauges built on GTS-based and historic data. Global Precipitation Climatology Centre, accessed 31 May 2012, doi:10.5676/DWD_GPCC/FD_M_V6_050.

Stackhouse, P., Jr., S. Gupta, S. Cox, T. Zhang, J. C. Mikovitz, and L. Hinkelman, 2011: The NASA/GEWEX surface radiation budget release 3.0: 24.5-year dataset. *GEWEX News*, No. 21, International GEWEX Project Office, Silver Spring, MD, 10–12.

Statistical Methods Branch, 2012: The yield forecasting program of NASS. U.S. Department of Agriculture, National Agricultural Statistics Service, NASS Staff Rep. SMB 12-01, 104 pp. [Available online at http://www.nass.usda.gov/Publications/ Methodology_and_Data_Quality/Advanced_Topics/Yield% 20Forecasting%20Program%20of%20NASS.pdf.]

Stensrud, D. J., 2007: *Parameterization Schemes: Keys to Understanding Numerical Weather Prediction Models*. Cambridge University Press, 459 pp.

van Wart, J., P. Grassini, and K. G. Cassman, 2013: Impact of derived global weather data on simulated crop yields. *Global Change Biol.*, **19**, 3822–3834, doi:10.1111/gcb.12302.

Wang, W., P. Xie, S.-H. Yoo, Y. Xue, A. Kumar, and X. Wu, 2011: An assessment of the surface climate in the NCEP climate forecast system reanalysis. *Climate Dyn.*, **37** (7–8), 1601–1620, doi:10.1007/s00382-010-0935-7.

Watson, J., and A. Challinor, 2013: The relative importance of rainfall, temperature and yield data for a regional-scale crop model. *Agric. For. Meteor.*, **170**, 47–57, doi:10.1016/j.agrformet.2012.08.001.

White, J., G. Hoogenboom, P. Stackhouse, P. Wilkens, and J. Hoel, 2011: Evaluation of satellite-based, modeled-derived daily solar radiation data for the continental United States. *Agron. J.*, **103**, 1242–1251, doi:10.2134/agronj2011.0038.

Willmott, C. J., and K. Matsuura, 1995: Smart interpolation of annually averaged air temperature in the United States. *J. Appl. Meteor.*, **34**, 2577–2586, doi:10.1175/1520-0450(1995)034<2577: SIOAAA>2.0.CO;2.

Xia, Y., and Coauthors, 2012: Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. *J. Geophys. Res.*, **117**, D03109, doi:10.1029/2011JD016048.

Xie, P., M. Chen, S. Yang, A. Yatagai, T. Hayasaka, Y. Fukushima, and C. Liu, 2007: A gauge-based analysis of daily precipitation over East Asia. *J. Hydrometeor.*, **8**, 607–626, doi:10.1175/JHM583.1.

You, L., and Coauthors, 2010: Spatial Production Allocation Model (SPAM), version 3, release 2. International Food Policy Research Institute. [Available online at http://MapSPAM. info.]

Zhao, Y., and Coauthors, 2007: Swift: Fast, reliable, loosely coupled parallel computation. *Proc. 2007 IEEE Congress on Services*, Salt Lake City, UT, IEEE, 199–206.