

Available online at www.sciencedirect.com



Agricultural and Forest Meteorology xxx (2004) xxx-xxx



www.elsevier.com/locate/agrformet

Using a general circulation model to forecast regional wheat yields in northeast Australia

James W. Hansen^{a,*}, Andries Potgieter^{b,1}, Michael K. Tippett^{a,2}

^aInternational Research Institute for Climate Prediction, P.O. Box 1000, Palisades, NY 10964-8000, USA ^bQueensland Centre for Climate Applications, Department of Primary Industries, P.O. Box 102, Toowoomba, Qld 4350, Australia

Received 13 April 2004; accepted 26 July 2004

Abstract

Forecasting regional crop yields and aggregate production is of interest to grain markets and drought policy response. We demonstrate a method for using GCM-based seasonal rainfall forecasts with a wheat simulation model for forecasting district and state aggregate yields in Queensland, Australia, and compare it with predictions based on climatology alone, phases of the El Niño-Southern Oscillation (ENSO), and Southern Oscillation Index (SOI) phases. We predicted yields by linear regression of simulated yields, transformed to correct departures from normality, against GCM predictors optimized by a linear transformation. Regression residuals provided estimates of the forecast distribution. Cross-validation of predictor selection and regression ensured conservative assessment of prediction accuracy. Statistical transformation of GCM output improved average gridded rainfall predictions and expanded the area over northeast Australia with significant predictability. Yield forecasts made 1 May, prior to planting, accounted for a significant portion of the variability of simulated yields averaged across the state (r = 0.518) and in most wheat-producing districts ($\bar{r} = 0.497$, area-weighted average among districts). Correlations were higher with observed detrended yields for the state (r = 0.706) and districts ($\bar{r} = 0.543$). Uncertainty of predicted yields diminished with successive monthly updates. Correlations of district-scale predictions with detrended observed yields showed greater heterogeneity in space and less consistency in time than correlations with simulated yields. For every forecast date, the GCM predicted state average yields simulated with observed weather more accurately than the other methods. The most accurate predictions of detrended observed state average yields came from the GCM for May, July and August, and from ENSO phases in June. The advantage of the GCM-based forecasts was greatest at the longest lead time. The improvement of accuracy at a long lead time has the potential to benefit the grain marketing industry by supporting proactive bulk handling and trading. © 2004 Elsevier B.V. All rights reserved.

Keywords: Global climate models (GCM); Crop simulation models; ENSO; Queensland

* Corresponding author. Tel.: +1 845 680 4410; fax: +1 845 680 4864.

E-mail address: jhansen@iri.columbia.edu (J.W. Hansen), potgieter@dpi.qld.gov.au (A. Potgieter), tippett@iri.columbia.edu (M.K. Tippett).

¹ Tel.: +61 7 46881417; fax: +61 7 46881193.

² Tel.: +1 845 680 4420; fax: +1 845 680 4865.

0168-1923/\$ – see front matter \odot 2004 Elsevier B.V. All rights reserved. doi:10.1016/j.agrformet.2004.07.005

2

J.W. Hansen et al./Agricultural and Forest Meteorology xxx (2004) xxx-xxx

1. Introduction

Australia is a major exporter of wheat and coarse grains. Wheat production alone is worth in excess of AUS\$ 3400 million annually on average (ABARE, 2002), but this varies as a result of one of the most variable climates in the world (Russell, 1988). Large year-to-year fluctuations in yields (and production) are of major concern to marketing agencies that sell this grain on a volatile world market. Financial hardship for producers due to recurring drought led to a national drought policy that provides financial assistance in "exceptional drought circumstances" (White et al., 1998). To address the concerns of commodity markets and drought relief programs, several grain exporting countries, including Australia, have developed operational methods to forecast regional crop yields and aggregate production (Motha and Heddinghaus, 1986; Stephens, 1988; Walker, 1989; Genovese and Terres, 1999; ABARE, 2004).

The El Niño-Southern Oscillation (ENSO) accounts for a substantial portion of the year-to-year variability of rainfall (McBride and Nicholls, 1983; Stone and Auliciems, 1992) and crop yields in northeastern Australia (Rimmington and Nicholls, 1993; Potgieter et al., 2002). Stone et al. (1996) developed a system for probabilistic prediction of rainfall at a seasonal lead time based on five discrete categories or "phases" (i.e., positive, rapidly rising, negative, rapidly falling and neutral) of the Southern Oscillation Index (SOI). The set of past years falling within a given category serve as equally probable analogs for predicting a distribution of rainfall outcomes conditioned on the observed SOI phase.

The Queensland Department of Primary Industries has developed an operational wheat forecasting model, using a simple agro-climatic model (Stephens et al., 1989) linked with a seasonal climate forecasts based on SOI phases (http://www.dpi.qld.gov.au/ fieldcrops). For each wheat-producing district in Australia, wheat yield predictions are generated through the growing season, and updated each month based on actual weather up to the forecast date and future weather scenarios sampled from historic analogs with the SOI phase observed at the forecast date (Potgieter et al., 2003). Advance information on likely production and its geographical distribution is useful to bulk handling and marketing agencies that manage storage and transport logistics and export sales in the recently deregulated marketing environment, and to government in relation to policy interventions triggered by the degree of exceptional drought circumstances (Hammer et al., 2001).

Statistical seasonal climate forecasts based on historic analogs have successfully supported agricultural applications in several contexts around the world (e.g., Hammer et al., 2001; De Jager and Potgieter, 1998). Yet the prospects for improving such forecasts are somewhat constrained since the accuracy of statistical models is primarily limited by the length and quality of the historical observational record, and by assumptions such as the stationarity of the climate system. Dynamic climate models, on the other hand, are based on physical laws, but are unable to resolved all temporal and spatial scales. Enhanced descriptions of physical processes in dynamical model offer the potential for future improvements in climate prediction (Cane, 2001; Goddard et al., 2001).

Although there is growing interest in linking seasonal forecasts based on dynamic general circulation models (GCMs) with biological simulation models to improve predictability of crop response, the difference in the spatial and temporal scales of GCMs and crop models complicates the task. Appropriate methodology for linking GCM output with crop models needs to be addressed before any gain in forecast quality can be realized (Meinke and Stone, in press; Hansen and Indeje, 2004). In this paper, we describe a method for combining GCMbased seasonal rainfall forecasts with a wheat simulation model for probabilistic regional yield forecasting, and demonstrate its application at a district scale in Queensland, Australia. We compare the GCM-based wheat forecasting system with predictions based on climatology alone, perfect knowledge of phases of the El Niño-Southern Oscillation (ENSO), and forecasts based on SOI phases.

2. Methods

We applied the following procedure to produce and evaluate GCM-based regional wheat yield forecasts. Using historic district rainfall data, for each district and forecast date, we simulated wheat yields with

every combination of observed antecedent rainfall up to the forecast date, and rainfall from the forecast date through harvest. A linear optimizing transformation of GCM seasonal forecast output fields provided predictor time-series. We transformed simulated yields to correct any departure from normality. For a given district, forecast date and year, we predicted yields by linear regression of transformed yields simulated with observed weather, as a function of GCM predictors, from all other years. We derived a probability distribution around each forecast, still in transformed space, from the distribution of regression residuals. Applying the inverse of the normalizing transformation produced the final forecast and its distribution. Consistent use of cross-validation for GCM predictor selection and yield prediction ensured conservative assessment of prediction accuracy. Details follow.

2.1. Wheat yield simulation

Stephens et al. (1989) developed a stress index model, STIN, for yield forecasting by joining the dynamic tipping-bucket soil water balance model implemented in the CERES models (Ritchie, 1972, 1998) to the FAO crop monitoring method (Frere and Popov, 1979). STIN calculates a stress index, SI, as cumulative function of water demand and plantextractable soil water simulated dynamically, using daily rainfall, and average weekly temperatures and solar irradiance required to calculate potential evapotranspiration. SI is sensitive to soil hydrological properties, and to changing crop water requirements as a function of phenology. The model treats phenology and its influence on water requirements (French and Schultz, 1984) as a fixed function of sowing date.

Developing the current operational forecasting system involved optimizing STIN to forecast wheat yields at a district scale (Stephens, 1995; Hammer et al., 1996). The model uses daily weather series averaged among available long-term stations within each district, weighted by areas of Thiessen polygons around each station. Soil parameters used for the soil water balance are from the single dominant soil series in each district. The soil water balance is particularly sensitive to available water holding capacity (AWHC). Aggregation error results when average or representative inputs (i.e., soil properties, weather, cultivar traits, management) are used to simulate crop yields at a spatial scale that encompasses heterogeneity of those inputs. To correct for aggregation error within each district, AWHC was calibrated to minimize error in predicted district aggregate yields (Hansen and Jones, 2000), using 19 years of available observed yields (1975–1993, Australian Bureau of Statistics). Final yields are estimated as linear regression functions of SI and year, accounting for a linear trend associated with changing technology. The direction and magnitude of the trend varied among districts. For state average yields, it was near zero (slope = -0.0007) and non-significant. Aggregated to the state level, correlations between observed and simulated wheat yields over the same period ranged from r = 0.87 to 0.95 among Australia's wheat-producing states (Potgieter et al., 2002).

For each district and forecast date (1 May, 1 June, 1 July or 1 August), STIN simulated 34 cropping seasons from 1968 to 2001, with year 2001 technology trends applied to all yield simulations. On 1 October of the year prior to harvest, STIN was initialized with an empty soil profile and then simulated with observed weather data up to the forecast date. After the forecast date, STIN was simulated through harvest (October-November, depending on location) with weather data from one of the 34 years. Using all available years of weather data after the forecast date for each cropping season results in a 34×34 matrix of simulated yields. The element y_{ii} of the crop yield matrix corresponds to the simulated crop yield obtained by driving STIN with antecedent weather data from year *i* up to the forecast date, and with within-season weather data from year j after the forecast date. The diagonal elements y_{ii} are yields simulated with year *i* weather through harvest, and represent crop model estimates of yields in the absence of climate uncertainty.

2.2. GCM predictor selection

We used output fields of the atmospheric general circulation model (GCM), ECHAM 4.5 (Roeckner et al., 1996; Goddard and Mason, 2002), run at approximately 280 km \times 280 km resolution in a seasonal hindcast mode. A 12-member ensemble of GCM model runs was forced with observed sea surface temperature (SST) boundary conditions up to the forecast start time. Persisted SSTs, obtained by

4

J.W. Hansen et al. / Agricultural and Forest Meteorology xxx (2004) xxx-xxx

adding SST anomalies observed during the month prior to the forecast to the historic SST climatology during each month of the forecast period, provided conservative predictions of SSTs to drive the GCM through the forecast period. For a given forecast year and date, each GCM ensemble member used the same SST boundary conditions, but different atmospheric initial conditions.

When selecting predictors based on GCM output fields, we took precautions to avoid model selection bias and over-fitting. We restricted our search to GCM precipitation in the region surrounding Queensland, 140.625°E to 157.5°E and 32.1°S to 15.3°S. Selection bias occurs when the same data are used to select predictors and estimate predictive ability, in which case the selected predictors appear to provide more predictive ability than they actually do (Zucchini, 2000). Although our goal was to find effective GCMbased predictors of wheat yields, we selected predictors based on observed seasonal rainfall rather than wheat yield to avoid selection bias. Predictor selection was based on seasonally-averaged rainfall data from the extended gridded data set of monthly precipitation (New et al., 2000) for 1968–2001. This data set, based on station observations interpolated to a 0.5° latitude by 0.5° longitude grid, was interpolated to the grid resolution of the GCM.

To select predictors for use in subsequent analyses, we first performed principle component analysis (PCA) on the GCM forecast precipitation (mean of the 12 GCM runs) and observed gridded precipitation, both normalized by their variance at each grid cell. Once the two data sets were decomposed into principle component (PC) timeseries, canonical correlation analysis (CCA) identified the linear combinations of model and observed PC time-series that were most highly correlated. We avoided over fitting by selecting the number of useful PCs and CCA modes based on cross-validated estimates of predictive ability. At each iteration of a Monte Carlo cross-validation procedure, 20 randomly-selected years were used to computed PCs and CCA modes and predict observed precipitation from GCM forecast precipitation for the remaining 14 independent years (Shao, 1993). Two hundred iterations of this method selected the first PC of GCM precipitation as the best predictor. Fig. 1 shows its spatial loading pattern. Applying this



Fig. 1. Map of EOF loading pattern from MOS transformation.

statistical correction expanded the portion of northeastern Australia with significant correlations between predicted and observed gridded rainfall (Fig. 2). Correlations between predicted and observed rainfall increased in the majority of GCM grid cells overlapping Queensland.

2.3. District rainfall and wheat yield prediction

We applied the same procedure to predict both district rainfall and district wheat yields. Ordinary least squares linear regression assumes that residuals are sequentially independent, normally distributed, and homoscedastic (i.e., constant in variance). Diagnostics of a subset of data showed some significant departures from normality by the Shapiro and Wilk (1965) test. Distributions of yields simulated with a given set of antecedent rainfall and withinseason rainfall sampled from all other years yields tended to be somewhat positively skewed, with mean g_1 ranging from 0.36 (antecedent rainfall through April) to 0.54 (antecedent rainfall through July) averaged across all districts and target years. Observed district rainfall generally showed stronger positive skewness ($g_1 = 1.05$, mean of May-August, all districts). The GCM-based predictors did not show significant departures from normality. The proportion of simulated series with skewness significantly different from zero (p < 0.05) ranged from 9.5%

J.W. Hansen et al. / Agricultural and Forest Meteorology xxx (2004) xxx-xxx



Fig. 2. Correlation between ECHAM 4.5 hindcast and gridded rainfall over Queensland (a) without and (b) with statistical transformation.

(antecedent rainfall through April) to 33.5% (antecedent rainfall through July).

We applied a Box and Cox (1964) transformation to each predictand time series to correct any departures from normality. The procedure finds the optimal transformation to normality within the family of power transformations:

$$y' = \begin{cases} \ln y, & \lambda = 0\\ \frac{y^{\lambda} - 1}{\lambda}, & \lambda \neq 0 \end{cases}$$
(1)

by selecting the value of λ that maximizes the loglikelihood function,

$$L = -\frac{n-1}{2}\ln S_{y'}^2 + (\lambda - 1)\frac{n-1}{n}\sum_{i=1}^n \ln y.$$
 (2)

Here y is an observation of a predictor (district rain or simulated yield), and y is its transformed value. In the case of simulated yields y_{ij} , we applied the transformation and subsequent regression to yields simulated with rainfall observed up to the forecast date in the given hindcast year *i*, and rainfall from the forecast date through harvest from each of the other years *j* in the 1968–2001 study period. Since the Box– Cox transformation is optimized independently for each sample distribution, it accommodates differences in predictand distributions among districts or forecast periods. In the case of rainfall, we applied the transformation to the distribution of years omitting the hindcast year. Fig. 3 illustrates the effect of the Box–Cox transformation on the distributions of rainfall and detrended yields observed at one district (Tara).

Hindcasts (i.e., forecasts for past periods) of *y* were obtained as a function of optimized GCM predictors *x* by cross-validated least-squares linear regression. As described earlier, GCM predictor selection was cross-validated. For each prediction \hat{y}_i (yield or rainfall), we first estimated \hat{y}' by linear regression from predictor $\mathbf{x}_i = x_{ij}, j = 1, ..., n, j \neq i$ and transformed predictand $\mathbf{y}'_i = y_{ij}, j = 1, ..., n, j \neq i$, then applied the inverse Box–Cox power transformation:

$$\hat{y}_i = \begin{cases} \exp \hat{y}'_i, & \lambda = 0\\ (\hat{y}'_i (\lambda + 1)^{(1/\lambda)} & \lambda \neq 0 \end{cases}$$
(3)

For the purpose of comparison, we also derived yield hindcasts based on the historical climatological distribution, and on categorical predictors (ENSO and SOI phases) associated with the El Niño-Southern Oscillation (ENSO). The cross-validated prediction of wheat yield in year i, belonging to a given predictor category (e.g., ENSO phase) k, is simply the mean of yields simulated with year i antecedent rainfall, and within-season rainfall from all other years within the category:

$$\hat{y}_i = (m_k - 1)^{-1} \sum_{\substack{j=1\\ j \in F_k, j \neq i}}^{m_k} y_{ij},$$
(4)

where m is the number of years within the set of years

J.W. Hansen et al. / Agricultural and Forest Meteorology xxx (2004) xxx-xxx



Fig. 3. Normal probability plots of (a, b) simulated yields and (c, d) observed rainfall (a, c) before and (b, d) after a Box–Cox transformation, Tara district.

 F_k belonging to predictor category k. This is analogous to the method that we used for cross-validated prediction by regression from GCM predictors.

ENSO phases are based on ± 0.4 °C anomalies of 5-month running averages of the NINO3.4 sea surface temperature index for at-least 6 months (Trenberth, 1997). The 1968-2001 study period includes 12 El Niño (1968, 69, 72, 76, 77, 82, 86, 87, 90, 91, 94 and 97) and 11 La Niña (1970, 71, 73, 74, 75, 84, 88, 95, 98, 99 and 2000) years. SOI phases are based on the SOI-standardized difference between Darwin and Tahiti surface air pressure-categorized by principal components analysis followed by cluster analysis into five phases: positive, rapidly rising, negative, rapidly falling, and near zero or neutral (Stone and Auliciems, 1992; available online at http://www.sci.usq.edu.au/staff/dunn/Datasets/applications/climatology/soiphases.html). Following Potgieter et al. (2002), to avoid categories with inadequate sample size we combined the positive and rapidly rising phases, and the negative and rapidly falling phases based on prior experience with wheat yield distributions among SOI phases. Combining SOI phases in this manner generally improved goodness of fit of cross-validated predictions with simulated and detrended observed yields. To simplify presentation of comparative results among forecast systems, we aggregated predicted, simulated, and detrended observed district wheat yields into Queensland state averages, weighted by recent estimates of the area under wheat in each district.

2.4. Wheat forecast distributions

For a continuous predictor (e.g., GCM output), the distribution of cross-validated regression residuals centered on the expected value of a given forecast provides a first approximation of the distribution of possible outcomes associated with that forecast. Antecedent rainfall influences the distribution of yields associated with the uncertainty of withinseason weather, and must therefore be held constant. For forecast year *i*, the appropriate forecast distribution of regression residuals, $\varepsilon_{ij} = y_{ij} - \hat{y}_{ij}, j \neq i$, centered on \hat{y}_i , where \hat{y}_{ij} is the regression prediction of simulated yield for within-season weather year *i*, but antecedent weather from hindcast year *i*. Because the regression model is developed for transformed

6

predictands y, we obtained forecast percentiles from the distribution of residuals, $\varepsilon'_{ij} = y'_{ij} - \hat{y}'_{ij}$, $j \neq i$, derived in transformed space, shifted by \hat{y}'_{i} , then backtransformed (Eq. (3)).

For yield forecasts based on climatology or analog years based on categorical predictors (e.g., ENSO phases), the forecast distribution is typically estimated as the distribution of results simulated with initial and antecedent conditions observed in the forecast year, and within-season weather sampled from other years within the predictor category. We employed this method for hindcasts based on climatology, ENSO phases and SOI phases. It is analogous to the use of regression residuals applied to GCM-based yield forecasts.



Fig. 4. Correlation between Queensland district wheat yields simulated with observed daily weather and GCM-based wheat yield hindcasts (a) 1 May; (b) 1 June; (c) 1 July; and (d) 1 August, 1968–2001, adjusted to 2001 technology trend.

8

J.W. Hansen et al. / Agricultural and Forest Meteorology xxx (2004) xxx-xxx

2.5. Analyses

We employed standard descriptive measures of goodness-of-fit to evaluate the accuracy of wheat yield predictions. Root-mean-squared error of prediction:

RMSE =
$$\sqrt{n^{-1} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2}$$
 (4)

represents overall error weighted by the square of deviations, where *n* is number of years *i*, *y* and \hat{y} are observed and predicted values. Mean absolute error:

$$MAE = n^{-1} \sum_{i=1}^{n} |\hat{y}_i - y_i|$$
(5)

also represents overall error, but is less sensitive than RMSE to errors in large predicted departures from the mean, and is therefore considered a more robust measure of accuracy. We also consider Pearson's coefficient of linear correlation, r, and mean bias error:

MBE =
$$n^{-1} \sum_{i=1}^{n} (\hat{y}_i - y_i).$$
 (6)

We evaluated the accuracy of wheat yield predictions both against yields simulated with observed daily weather to estimate error associated with climate prediction and the climate-crop model linkage, and against available (1975–1993) observed yields. To simplify presentation, we evaluated observed, simulated and predicted district wheat yields aggregated into Queensland state averages weighted by the area under wheat in each district.

3. Results and discussion

3.1. GCM-based wheat yield prediction

Fig. 4 shows the correlation between wheat yields predicted using the GCM and yields simulated with observed daily weather for each district. Evaluation against simulated yields captures random error (i.e., not correctable by linear calibration) associated with the seasonal climate forecast system and its link with the wheat simulation model. It does not incorporate crop model error. Forecasts made 1 May account for a significant portion of variability of simulated yields in most wheatproducing districts (Fig. 4a), even though planting has not yet started by that date in most of the state. Uncertainty of predicted district yields diminished with successive monthly updates (Fig. 4b-d, Table 1) as an increasing proportion of integrated water stress is due to rainfall that is observed rather than predicted. Correlations showed the greatest increase with the 1 August update. Simulated anthesis is complete in most districts by this time, and rainfall during the most critical growth stage for yield determination is now observed rather than predicted (Nix and Fitzpatrick, 1969).

Correlations of predictions with detrended observed yields (Fig. 5) showed less homogeneity in space and less consistency in time relative to correlations with simulated yields (Fig. 4). Correlations between predictions and observations decreased from 1 May to 1 June, then increased with subsequent updates (Table 1). The increase in forecast accuracy from July to August was much less

Table 1

Correlations of GCM-based wheat yield hindcasts with yields observed and simulated with observed weather, adjusted to 2001 technology trend, aggregated to state average yields and area-weighted mean of correlations among districts

Forecast date	State average y	ield		Mean correlation among districts			
	Simulated		Observed	Simulated		Observed	
	1968-2001	1975-1993	1975-1993	1968-2001	1975-1993	1975–1993	
1 May	0.518	0.588	0.706	0.497	0.537	0.543	
1 June	0.510	0.476	0.579	0.535	0.484	0.486	
1 July	0.600	0.640	0.781	0.633	0.636	0.669	
1 August	0.862	0.862	0.810	0.863	0.857	0.713	
Simulated			0.863			0.762	

J.W. Hansen et al. / Agricultural and Forest Meteorology xxx (2004) xxx-xxx



Fig. 5. Correlation between observed Queensland district wheat yields and GCM-based wheat yield hindcasts (a) 1 May; (b) 1 June; (c) 1 July; and (d) 1 August, 1975–1993, adjusted to 2001 technology trend.

pronounced for observed than for simulated yields (Fig. 6). This is likely due to the heterogeneity of phenology of actual wheat associated with varied planting dates, cultivars and local temperatures that the simulations do not account for. The reason for the apparent drop in GCM-based prediction accuracy of observed yields from 1 May to 1 June is not as clear. However, because it also appears for simulated yields for the 1975–1993 period when observed yields are available, but not for the entire 1968–2001

period used to fit the regressions, we speculate that it may be an artifact of the subsample of years for which observed data are available.

As expected (e.g., Hammer et al., 1996; Hansen and Jones, 2000), prediction accuracy was generally better at the state scale than at the smaller district scale. Although the wheat simulation model accounted for about 75% of the variance of detrended state average wheat yields, correlations for individual districts were generally lower, accounting for an average of 58% of

J.W. Hansen et al. / Agricultural and Forest Meteorology xxx (2004) xxx-xxx



Fig. 6. RMSE of hindcasts of (a) simulated (1968–2001) and (b) observed (1975–1991) Queensland average wheat yields, adjusted to 2001 technology trend, as a function of forecast date.

the variance (i.e., r^2 , Table 1), weighted by the area under wheat in each district. When aggregated to the state level, correlations with hindcasts were higher for detrended observed yields than for simulated yields for the May–July forecast dates, even when considering only the 1975–1993 period for which observations were available (Table 1). This was contrary to expectation, as predictions of simulated yields reflect

Table 2

Goodness of fit statistics for Queensland average wheat yield predictions by alternate methods vs. yields observed (1975–1993) and simulated with observed weather (1968–2001), adjusted to 2001 technology trend

Scenario	Simulated (1968–2001)				Observed (1975–1993)			
	RMSE (Mg ha ⁻¹)	MAE (Mg ha ⁻¹)	MBE (Mg ha ⁻¹)	r	RMSE (Mg ha ⁻¹)	MAE (Mg ha ⁻¹)	MBE (Mg ha ⁻¹)	r
Simulated					0.242	0.209	-0.002	0.863
1 May								
GCM-based	0.333	0.268	-0.029	0.518	0.368	0.296	-0.092	0.706
ENSO phases	0.351	0.287	0.011	0.435	0.415	0.342	-0.053	0.542
Marh–April SOI phases	0.386	0.309	0.013	0.208	0.440	0.378	-0.036	0.403
Climatology	0.378	0.301	0.009	0.200	0.460	0.403	-0.025	0.314
1 June								
GCM-based	0.336	0.250	-0.033	0.510	0.398	0.343	-0.079	0.579
ENSO phases	0.347	0.282	0.005	0.454	0.387	0.315	-0.067	0.673
April–May SOI phases	0.361	0.292	0.004	0.408	0.390	0.331	0.016	0.586
Climatology	0.383	0.312	0.005	0.155	0.445	0.385	-0.025	0.410
1 July								
GCM-based	0.315	0.240	-0.041	0.600	0.331	0.287	-0.089	0.780
ENSO phases	0.337	0.267	0.005	0.500	0.346	0.289	-0.058	0.769
May-June SOI phases	0.336	0.270	0.011	0.501	0.361	0.296	-0.058	0.694
Climatology	0.359	0.288	0.009	0.374	0.398	0.340	-0.021	0.651
1 August								
GCM-based	0.196	0.156	-0.016	0.862	0.292	0.260	-0.010	0.810
ENSO phases	0.211	0.169	0.010	0.844	0.305	0.266	-0.007	0.800
June–July SOI phases	0.231	0.183	0.009	0.804	0.333	0.278	0.018	0.736
Climatology	0.229	0.190	0.011	0.844	0.327	0.287	0.023	0.800

DTD 5

J.W. Hansen et al. / Agricultural and Forest Meteorology xxx (2004) xxx-xxx

only the uncertainty of rainfall while predictions of observed yields also incorporate crop model error. However, overall prediction error (Fig. 6, Table 2) was greater for observed than for simulated yields, reflecting the greater systematic error (i.e., MBE) in predicting observed yields. Correlations with predictions were more similar for simulated and observed yields when considering the area-weighted average correlation among districts (Table 1). Although the spatial distributions of predictability of rainfall (Fig. 7) and simulated yields (Fig. 4) are not identical, they do show some similarities. May yield forecasts show highest correlations in a region in the central and north-central part of the Queensland wheat belt, and insignificant correlations in the far northwest and a few districts in the east-central region (Fig. 4a). Rainfall hindcast correlations in May and June (Fig. 7a and c) are also weak in the



Fig. 7. Correlation between Queensland district rainfall totals and GCM-based hindcasts (a) June–September predicted 1 May; (b) June–September predicted 1 June; (c) July–September predicted 1 July; and (d) August–September predicted 1 August 1968–2001.

12

ARTICLE IN PRESS

northwest and in the same east-central region. A region of relatively high correlations in the central and north-central region apparent for the July and August forecasts (Fig. 7c and d) roughly coincides with the region with relative high correlations for May forecasts of simulated yields. It is more difficult to relate the spatial distribution of rainfall hindcast correlations to yield hindcast correlations with detrended observed yields (Fig. 5) without understanding how yield determinants other than water stress vary across the region. Correlations between predictions and observations are substantially higher for yields than for rainfall.

3.2. Comparison of wheat prediction methods

Table 2 summarizes results of district yield hindcasts aggregated to the state level, based on the GCM, ENSO phases, SOI phases and climatology. For every forecast period, the GCM-based method gave better results than the other methods for state average yields simulated with observed weather. Mean bias (MBE) was generally low (<10% of RMSE), but greatest and consistently negative for GCM-based forecasts, and greater for detrended observed than for simulated yields. Bias tended to be greater for detrended observed yields. Yield hindcasts generally under-predicted observed yields because mean yields (both observed and simulated) during the 1975-1993 period for which observations are available were slightly higher (27 kg ha^{-1}) than for the entire 1968–2001 period used to fit the prediction models. When evaluated against observed state average yields, the best predictions resulted from the GCM in May, July and August, and ENSO phases in June. The advantage of the GCM over SOI phases was greatest at the longest lead time, prior to planting (Table 2). As the season progresses, the impact of remaining climatic uncertainty on yield diminishes. By 1 August, observed antecedent rainfall accounts for most of the uncertainty of yields, and climate forecasts provide little additional information beyond climatology. Although we included climatology as a naive forecast system, it does provide some predictability even before planting due to the influence of antecedent soil moisture, and the high

water-holding capacity of soils in the Queensland wheat belt.

ENSO phases, as we use them, are not a true forecast system, as they cannot generally be known with certainty until about the end of the wheat growing season. The state of ENSO is somewhat predictable at a lead time relevant to crop forecasting. A coupled ocean-atmosphere model that has provided operational SST forecasts since 1998, has demonstrated a high degree of predictability of the NINO3.4 SST index, used to classify ENSO years in



Fig. 8. Time series of observed Queensland average wheat yields and percentiles of GCM-based hindcast distributions (a) 1 May; (b) 1 June; (c) 1 July; and (d) 1 August 1968–2001, adjusted to 2001 technology trend.

our study, at one and four months lead time (Goddard et al., 2003). SOI phases at the time of the forecast are one method to anticipate the likely state of ENSO later in the year. However, since ENSO events tend to develop between April and June, predictions based on either SOI phases or the coupled model during this time of the year tend to be more uncertain than later in the year. For the 1968–2001 period for which ECHAM predictors were available, the GCM generally provided the most accurate wheat yield forecasts, particularly at long lead times. Longer time series of predictors available for ENSO or SOI phases than for the GCM can potentially improve forecasts, assuming stationarity of rainfall and ENSO teleconnections. We therefore cannot conclude from our analyses that the GCM will give better predictions than operational wheat forecasts based on SOI phases.

3.3. Wheat yield forecast distributions

Appropriate use of operational yield forecasts requires a clear understanding of the uncertainty associated with a given forecast. Fig. 8 shows percentiles of the distribution of each year's GCMbased Queensland wheat yield hindcast for each forecast period, along with observed, detrended yields. Although the procedure for estimating forecast distributions applies to each district, we show Queensland averages for simplicity of presentation. The forecast distribution's shape and dispersion changes from year to year due to (a)



Fig. 9. Queensland average observed wheat yield, and yield hindcast percentiles as a function of forecast date (1 May through the time of harvest), adjusted to 2001 technology trend, for an illustrative El Niño, La Niña and neutral year, based on the GCM, ENSO phases and climatology. The centerline (-..-) shows the median simulated yield (1.49 Mg ha⁻¹).

14

ARTICLE IN PRESS

the influence of differing antecedent rainfall; and (b) the cross-validation procedure that omits withinseason rainfall for a given hindcast year from estimation of that year's distribution. The forecast distributions tend to narrow and the magnitude of predicted shifts from the long-term mean $(1.51 \text{ Mg ha}^{-1})$ increase with successive updates as the season progresses (Fig. 8).

For an illustrative El Niño, La Niña and neutral year, Fig. 9 shows how the forecast distribution changes through the season for yield forecasts based on the GCM, ENSO phases and climatology. In 1988, a La Niña year, the dispersion (e.g., the dark shaded inter-quartile distance) about each updated forecast median was higher for the ENSO phase than for the climatology-based forecast (Fig. 9), even though the correlations are greater and prediction errors lower for forecasts based on ENSO phases than for climatology-based forecasts for each forecast date (Table 2). The uncertainty of simulated wheat yields appears to be greater in La Niña years than in El Niño or neutral years (Fig. 10). For the illustrative neutral year (1989, Fig. 9), inter-quartile distances were smaller for ENSO phase-based forecasts than for GCM-based forecasts for the 1 May, 1 July and 1 August forecast dates. This suggests that the



Fig. 10. Box plots showing 0, 25, 50, 75 and 100th percentiles of Queensland average wheat yields simulated with observed daily rainfall adjusted to 2001 technology trend, for all years (1975–1993) and by ENSO phase.

uncertainty of wheat yield forecasts may be substantially higher than climatology suggests in La Niña years, and that advance knowledge of neutral ENSO conditions may reduce the uncertainty of predicted wheat yields more than GCM forecasts. However, variances of wheat yields did not differ significantly among ENSO phases (p > 0.4) by the Levene (1960) and Brown and Forsythe (1974) tests.

4. Conclusions

In this study we demonstrated a regression approach for connecting GCM outputs with a crop simulation model for probabilistic prediction of district-scale wheat yields in Queensland, Australia. The GCM-based approach showed greater accuracy than obtained from perfect knowledge of ENSO phases or the SOI phase system, particularly at the longest lead time. We cannot conclude from this study that the GCM based system will give better predictions than the current operational wheat yield forecast based on the SOI phase system (due primarily to differences in the number of years of available predictor data for the two systems). Yet, the GCM-based system's ability to improve forecast accuracy during the pre-planting period (end of April) when ENSO seems to be less predictable, is encouraging. Improving the accuracy of forecasts issued prior to planting may substantially increase their value to farmers who may use crop forecast information for land allocation or forward price contract decisions. Further research is necessary to see if the lead time of forecasts can be increased even further. The proposed methodology is only one of several proposed approaches for connecting GCMs with biophysical models (e.g., Hansen and Indeje, 2004). Exploration of other approaches (e.g., dynamic downscaling of GCM outputs, stochastic weather generation, Markov methods based on synoptic weather types, analog approaches) are beyond the scope of this study.

The higher apparent predictability of yields than rainfall is contrary to the argument that predictability of crop response to rainfall must be less than that of seasonal rainfall totals, due to accumulation of the error in predicting local seasonal rainfall from climate predictors and the error in predicting yields from seasonal rainfall. This argument overlooks two

considerations. First, yields respond to additional information beyond seasonal rainfall-observed antecedent rainfall in this case-that contributes to predictability. In Queensland's wheat belt, the amount of rainfall during the spring-summer fallow period, and hence the amount of water stored in the soil profile at planting, can exert greater influence than withinseason rainfall on final yields (Meinke and Hochman, 2000). This is why climatology-based forecasts show some predictability of yields (Table 2). Second, we predict yields as a function of seasonal climate predictors, and not as a function of local seasonal rainfall totals. By bypassing rainfall as an intermediate prediction, we reduce the accumulation of errors, and potentially incorporate relevant information about the distribution of rainfall within the season or other relevant meteorological variables that are embedded in seasonal climatic predictors (i.e., the GCM and its SST boundary forcing), but that are lost when converting them into seasonal rainfall totals.

The question of whether apparent differences in the variability of yields in the different ENSO phases is a consequence of differing strengths of teleconnections, an artifact of the positive skewness of the yield distribution, or an effect of small sample size is still open to debate. It relates to the question of whether forecast distributions based on residuals about a continuous prediction system provide information that is different from distributions arising from categorical predictors and historic analogs. Answering the latter question would require considering other interrelated components of forecast quality such as reliability (i.e., the consistency, through time, between forecast distributions and the distributions of outcomes conditioned on the forecasts), resolution (i.e., the distribution shift from climatology) and sharpness (i.e., the dispersion of the forecast distribution). Characterizing and verifying these aspects of forecast quality beyond accuracy would enhance the assimilation of GCMbased probabilistic forecasts within the operational commodity-forecasting environment (Potgieter et al., 2003). However, methodology for assessing the quality of cumulative forecast distributions about continuous forecasts is not as well developed as for categorical probabilistic forecasts. Questions about the consistency and interpretation of forecast probabilities derived from categorical versus continuous predictors are a relevant topic of future research.

Acknowledgements

We gratefully acknowledge helpful discussions with H. Meinke and S. Mason. This work was supported in part by a grant/cooperative agreement number NA67GP0299 from the National Oceanic and Atmospheric Administration. The views expressed herein are those of the authors, and do not necessarily reflect the views of NOAA or any of its sub-agencies.

References

- ABARE, 2002. Australian Commodity Statistics. Australian Bureau of Agricultural and Resource Economics, Canberra, Australia.
- ABARE, 2004. Australian Crop Report No. 128. Australian Bureau of Agricultural and Resource Economics, Canberra, Australia.
- Box, G.E.P., Cox, D.R., 1964. An analysis of transformations. J. R. Stat. Soc. B 26, 211–243.
- Brown, M.B., Forsythe, A.B., 1974. Robust tests for the equality of variances. J. Am. Stat. Assoc. 69, 264–267.
- Cane, M.A., 2001. Understanding and predicting the world's climate system. In: Rosenzweig, C., Boote, K.J., Hollinger, S., Iglesias, A., Phillips, J. (Eds.), Impacts of El Niño and Climate Variability on Agriculture. ASA Special Publication No. 63. American Society of Agronomy, Madison, WI, USA, pp. 1–20.
- De Jager, J.M., Potgieter, A.B., Van Den Berg, W.J., 1998. Framework for forecasting the extent and severity of drought in maize in the Free State Province of South Africa. Agric. Syst. 57, 351– 365.
- French, R.J., Schultz, J.E., 1984. Water use efficiency of wheat in a Mediterranean-type environment: 1. The relation between yield, water use and climate. Aust. J. Agric. Res. 35, 743–764.
- Frere, M., Popov, G.F., 1979. Agrometeorological crop monitoring and forecasting. FAO Plant Prod. and Protect. Pap. No. 17, FAO, Rome.
- Genovese, G., Terres, J.M., 1999. Crop yield forecasts derived from agrometeorological models at European level: the MARS project. in: Proceedings of the Workshop on the Application of Seasonal-timescale Weather Forecasts, 14–15 June 1999. European Centre for Medium Range Weather Forecasts, Reading, UK, pp. 35–41.
- Goddard, L., Mason, S.J., 2002. Sensitivity of seasonal climate forecasts to persisted SST anomalies. Climate Dynamics 19, 619–632.
- Goddard, L., Mason, S.J., Zebiak, S.E., Ropelewski, C.F., Basher, R., Cane, M.A., 2001. Current approaches to seasonal to interannual climate predictions. Int. J. Climatol. 21, 1111–1152.
- Goddard, L., Barnston, A.G.A., Mason, S.J., 2003. Evaluation of the IRI's "net assessment" seasonal climate forecasts 2001. Bull. Am. Meteorol. Soc. 84, 1761–1781.
- Hammer, G.L., Stephens, D., Butler, D., 1996. Development of a national drought alert strategic information system. Wheat Modelling Sub-Project—Development of Predictive Models

J.W. Hansen et al. / Agricultural and Forest Meteorology xxx (2004) xxx-xxx

of Wheat Production. Report 6a:51, Land and Water Resources Research and Development Corporation, Canberra, Australia.

- Hammer, G.L., Hansen, J.W., Phillips, J., Mjelde, J.W., Hill, H.S.J., Love, A., Potgieter, A., 2001. Advances in application of climate prediction in agriculture. Agric. Syst. 70, 515–553.
- Hansen, J.W., Jones, J.W., 2000. Scaling–up crop models for climate variability applications. Agric. Syst. 65, 43–72.
- Hansen, J.W., Indeje, M., 2004. Linking dynamic seasonal climate forecasts with crop simulation for maize yield prediction in semi-arid Kenya. Agric. Forest Meteorol. 125, 143–157.
- Levene, H., 1960. Robust tests for equality of variance. In: Olkin, I. (Ed.), Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling. Stanford University Press, Palo Alto, California, pp. 278–292.
- McBride, J.L., Nicholls, N., 1983. Seasonal relationships between Australian rainfall and the Southern Oscillation. Monthly Weather Rev. 111, 1998–2004.
- Meinke, H., Hochman, Z., 2000. Using seasonal climate forecasts to manage dryland crops in northern Australia—experiences from the 1997/98 seasons. In: Hammer, G.L., Nicholls, N., Mitchell, C. (Eds.), Applications of Seasonal Climate Forecasting in Agricultural and Natural Ecosystems. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 149–165.
- Meinke, H., Stone, R.C., in press. Seasonal and inter-annual climate forecasting: the new tool for increasing preparedness to climate variability and change in agricultural planning and operations. Climatic Change.
- Motha, R.P., Heddinghaus, T.R., 1986. The joint agricultural weather facility's operational assessment program. Bull. Am. Meteorol. Soc. 67, 1114–1122.
- New, M.G., Hulme, M., Jones, P.D., 2000. Representing 20th century space-time climate variability. II. Development of 1901–1996 monthly terrestrial climate fields. J. Climate 13, 2217–2238.
- Nix, H.A., Fitzpatrick, E.A., 1969. An index of crop water stress related to wheat and grain sorghum yields. Agric. Meteorol. 6, 321–337.
- Potgieter, A.B., Hammer, G.L., Butler, D., 2002. Spatial and temporal patterns in Australian wheat yield and their relationship with ENSO. Aust. J. Agric. Res. 53, 77–89.
- Potgieter, A.B., Everingham, Y.L., Hammer, G.L., 2003. On measuring quality of a probabilistic commodity forecast for a system that incorporates seasonal climate forecasts. Int. J. Climatol. 23, 1195–1210.
- Rimmington, G.M., Nicholls, N., 1993. Forecasting wheat yields in Australia with the Southern Oscillation Index. Aust. J. Agric. Res. 44, 625–632.

- Ritchie, J.T., 1972. Model for predicting evaporation from a row crop with incomplete cover. Water Resource Res. 8, 1204–1213.
- Ritchie, J.T., 1998. Soil water balance and plant water stress. In: Tsuji, G.Y., Hoogenboom, G., Thornton, P.K. (Eds.), Understanding Options for Agricultural Production, vols. 41–54. Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Roeckner, E., Arpe, K., Bengtsson, L., Christoph, M., Claussen, M., Dümenil, L., Esch, M., Giorgetta, M., Schlese, U., Schulzweida, U., 1996. The atmospheric general circulation model ECHAM-4: model description and simulation of present-day climate. Max-Planck-Institut fuer Meteorologie Report No. 218, Hamburg, p. 90.
- Russell, J.S., 1988. The effect of climatic change on the productivity of Australian agroecosystems. In: Pearman, G.I. (Ed.), Greenhouse—Planning for Climate Change. CSIRO, Australia, Melbourne, pp. 491–505.
- Shao, J., 1993. Linear model selection by cross-validation. J. Am. Stat. Assoc. 88, 486–494.
- Shapiro, S.S., Wilk, M.B., 1965. An analysis of variance test for normality (complete samples). Biometrika 52, 591–611.
- Stephens, D.J., 1995. Crop yield forecasting over large areas in Australia. Unpublished Ph.D. thesis. Murdoch University, Perth, Western Australia, p. 305.
- Stephens, D.J., 1988. The investigation of crop yield forecasting in North America. Fin. Rep. to Queen Elizabeth 2 Silver Jubilee Trust for Young Australians, Murdoch University, Perth, Western Australia, p. 9.
- Stephens, D.J., Lyons, T.J., Lamond, M.H., 1989. A simple model to forecast wheat yield in Western Australia. J. R. Soc. West. Aust. 71, 77–81.
- Stone, R.C., Auliciems, A., 1992. SOI phase relationships with rainfall in eastern Australia. Int. J. Climatol. 12, 625– 636.
- Stone, R.C., Hammer, G.L., Marcussen, T., 1996. Prediction of global rainfall probabilities using phases of the Southern Oscillation Index. Nature 384, 252–255.
- Trenberth, K.E., 1997. The Definition of El Niño. Bull. Am. Meteorol. Soc. 78, 2771–2777.
- Walker, G.K., 1989. Model for operational forecasting of western Canadian wheat yield. Agric. For. Meteorol. 44, 339–351.
- White, D.H., Howden, S.M., Walcott, J.J., Cannon, R.M., 1998. A framework for estimating the extent and severity of drought, based on a grazing system in south-eastern Australia. Agric. Syst. 57, 259–270.
- Zucchini, W., 2000. An introduction to model selection. J. Math. Psychol. 44, 41–61.

16