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

James W. Hansena,*, Andries Potgieterb,1, Michael K. Tippettb,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 ($r = 0.497$, area-weighted average among districts). Correlations were higher with observed detrended yields for the state ($r = 0.706$) and districts ($r = 0.543$). 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
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 (Motta 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 GCM-based 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 plant-extractable 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 \( 34 \times 34 \) matrix of simulated yields. The element \( y_{ij} \) 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
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 GCM-based 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) time-series, 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 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 within-season 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%
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:

\[ y' = \begin{cases} \ln y, & \lambda = 0 \\ \frac{y^\lambda - 1}{\lambda}, & \lambda \neq 0 \end{cases} \tag{1} \]

by selecting the value of \( \lambda \) that maximizes the log-likelihood function,

\[ L = -\frac{n-1}{2} \ln S_y^2 + (\lambda - 1) \frac{n-1}{n} \sum_{i=1}^{n} \ln y. \tag{2} \]

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 \( \hat{y}'_i \) by linear regression from predictor \( x_{ij} \), \( j = 1, \ldots, n \), \( j \neq i \) and transformed predictand \( y'_i = y_{ij}, j = 1, \ldots, n, j \neq i \), then applied the inverse Box–Cox power transformation:

\[ \hat{y}_i = \begin{cases} \exp \hat{y}'_i, & \lambda = 0 \\ \left( \frac{\hat{y}'_i (\lambda + 1)}{\lambda} \right), & \lambda \neq 0 \end{cases} \tag{3} \]

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:

\[ \hat{y}_i = (m_k - 1)^{-1} \sum_{j=1}^{m_k} y_{ij}, \tag{4} \]

where \( m \) is the number of years within the set of years
$F_k$ 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 within-season weather, and must therefore be held constant. For forecast year $i$, the appropriate forecast distribution can be estimated from the empirical distribution of regression residuals, $v_{ij} = y_{ij} - \hat{y}_{ij}, j \neq i$, centered on $\hat{y}_i$, where $\hat{y}_{ij}$ is the regression prediction of simulated yield for within-season weather year $j$, but antecedent weather from hindcast year $i$. Because the regression model is developed for transformed

![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.](image-url)
predictands $y$, we obtained forecast percentiles from the distribution of residuals, $e'_{ij} = y'_{ij} - y'_0$, $j \neq i$, derived in transformed space, shifted by $y'_0$, then back-transformed (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.
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:

\[ \text{RMSE} = \sqrt{n^{-1} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2} \]  

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:

\[ \text{MAE} = n^{-1} \sum_{i=1}^{n} |\hat{y}_i - y_i| \]

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:

\[ \text{MBE} = n^{-1} \sum_{i=1}^{n} (\hat{y}_i - y_i). \]

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 wheat-producing 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>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1 May</td>
<td>0.518</td>
<td>0.588</td>
<td>0.706</td>
<td>0.497</td>
<td>0.537</td>
<td>0.543</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 June</td>
<td>0.510</td>
<td>0.476</td>
<td>0.579</td>
<td>0.535</td>
<td>0.484</td>
<td>0.486</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 July</td>
<td>0.600</td>
<td>0.640</td>
<td>0.781</td>
<td>0.633</td>
<td>0.636</td>
<td>0.669</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 August</td>
<td>0.862</td>
<td>0.862</td>
<td>0.810</td>
<td>0.863</td>
<td>0.857</td>
<td>0.713</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simulated</td>
<td></td>
<td>0.863</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.762</td>
</tr>
</tbody>
</table>
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
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

![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.](image)

### 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

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE (Mg ha$^{-1}$)</td>
<td>MAE (Mg ha$^{-1}$)</td>
</tr>
<tr>
<td>Simulated</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 May</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCM-based</td>
<td>0.333</td>
<td>0.268</td>
</tr>
<tr>
<td>ENSO phases</td>
<td>0.351</td>
<td>0.287</td>
</tr>
<tr>
<td>Marh–April SOI phases</td>
<td>0.386</td>
<td>0.309</td>
</tr>
<tr>
<td>Climatology</td>
<td>0.378</td>
<td>0.301</td>
</tr>
<tr>
<td>1 June</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCM-based</td>
<td>0.336</td>
<td>0.250</td>
</tr>
<tr>
<td>ENSO phases</td>
<td>0.347</td>
<td>0.282</td>
</tr>
<tr>
<td>April–May SOI phases</td>
<td>0.361</td>
<td>0.292</td>
</tr>
<tr>
<td>Climatology</td>
<td>0.383</td>
<td>0.312</td>
</tr>
<tr>
<td>1 July</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCM-based</td>
<td>0.315</td>
<td>0.240</td>
</tr>
<tr>
<td>ENSO phases</td>
<td>0.337</td>
<td>0.267</td>
</tr>
<tr>
<td>May–June SOI phases</td>
<td>0.336</td>
<td>0.270</td>
</tr>
<tr>
<td>Climatology</td>
<td>0.359</td>
<td>0.288</td>
</tr>
<tr>
<td>1 August</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCM-based</td>
<td>0.196</td>
<td>0.156</td>
</tr>
<tr>
<td>ENSO phases</td>
<td>0.211</td>
<td>0.169</td>
</tr>
<tr>
<td>June–July SOI phases</td>
<td>0.231</td>
<td>0.183</td>
</tr>
<tr>
<td>Climatology</td>
<td>0.229</td>
<td>0.190</td>
</tr>
</tbody>
</table>
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

![Maps showing correlation between Queensland district rainfall totals and GCM-based hindcasts](image)
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](image-url)
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 GCM-based 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)
the influence of differing antecedent rainfall; and (b) the cross-validation procedure that omits within-season 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 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 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 within-season 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 GCM-based 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


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.


