Spatial and temporal patterns in Australian wheat yield and their relationship with ENSO

A. B. PotgieterA, G. L. HammerB, and D. ButlerC

A°DPI/DNR Queensland Centre for Climate Applications (QCCA), PO Box 102, Toowoomba, Qld 4350, Australia.
B°DPI/DNR/CSIRO Agricultural Production Systems Research Unit (APSRU), PO Box 102, Toowoomba, Qld 4350, Australia.
C°DPI Farming Systems Institute, PO Box 102, Toowoomba, Qld 4350, Australia.

Abstract. Spatial and temporal variability in wheat production in Australia is dominated by rainfall occurrence. The length of historical production records is inadequate, however, to analyse spatial and temporal patterns conclusively. In this study we used modelling and simulation to identify key spatial patterns in Australian wheat yield, identify groups of years in the historical record in which spatial patterns were similar, and examine association of those wheat yield year groups with indicators of the El Niño Southern Oscillation (ENSO). A simple stress index model was trained on 19 years of Australian Bureau of Statistics shire yield data (1975–93). The model was then used to simulate shire yield from 1901 to 1999 for all wheat-producing shires. Principal components analysis was used to determine the dominating spatial relationships in wheat yield among shires. Six major components of spatial variability were found. Five of these represented near spatially independent zones across the Australian wheatbelt that demonstrated coherent temporal (annual) variability in wheat yield. A second orthogonal component was required to explain the temporal variation in New South Wales. The principal component scores were used to identify high- and low-yielding years in each zone. Year type groupings identified in this way were tested for association with indicators of ENSO. Significant associations were found for all zones in the Australian wheatbelt. Associations were as strong or stronger when ENSO indicators preceding the wheat season (April–May phases of the Southern Oscillation Index) were used rather than indicators based on classification during the wheat season. Although this association suggests an obvious role for seasonal climate forecasting in national wheat crop forecasting, the discriminatory power of the ENSO indicators, although significant, was not strong. By examining the historical years forming the wheat yield analog sets within each zone, it may be possible to identify novel climate system or ocean–atmosphere features that may be causal and, hence, most useful in improving seasonal forecasting schemes.

Additional keywords: SOI phases, agro-climatic models, homologous wheat zones.

Introduction

Australian wheat yield is extremely variable from year to year and among key production regions (Figs 1 and 2). Although some positive technology trends have been identified (Hamblin and Kyneur 1993), the dominating effect of rainfall variability and water limitation is the major cause of variation in Australian wheat yield (Nix 1975; Cornish et al. 1998). This enables robust prediction of wheat yield at shire scale using simple agro-climatic models, such as those of Stephens et al. (1989).

Rainfall variability is high throughout most of the Australian wheatbelt and this variability has been related to the El Niño Southern Oscillation (ENSO) phenomenon (McBride and Nicholls 1983; Stone and Auliciems 1992). The associations between ENSO and Australian rainfall have provided a basis for prediction of seasonal rainfall likelihood using the Southern Oscillation Index (SOI) (Stone et al. 1996), which is an indicator of the state of ENSO based on the normalised pressure difference between Tahiti and Darwin. Using these associations, Rimmington and Nicholls (1993) related wheat yield in Australia directly to the state of ENSO prior to the cropping season. Subsequently, studies have shown some potential for use of ENSO-based seasonal climate forecasts in commodity forecasting (Stephens 1995). However, the length of historical production records is inadequate to test this conclusively using hindcasting. It has been shown, however, that a seasonal climate forecasting system based on the SOI phases of Stone et al. (1996) can be used effectively to assist in crop management at field and farm scale (Hammer et al. 1996a; Meinke and Hammer 1997).
Spatial and temporal variability in Australian wheat production is of major concern to bulk handling and marketing agencies that need to manage storage and transport logistics and export sales. Historical production records, however, are inadequate for detailed study of spatial and temporal yield patterns due to issues with technology effects and shifts in production areas. Nonetheless, the available historical records have been invaluable in the development and testing of shire wheat yield models (Stephens 1995; Hammer et al. 1996b). These models have been used with long-term historical rainfall data to simulate a historical time series of shire yields based on current technology and production districts (e.g. Stephens et al. 2000). This has shown a significant effect of ENSO by quantifying shifts in simulated yield likelihood associated with SOI phases at State and national scales. The long-term historical time series of simulated shire wheat yields enable examination of spatial and temporal patterns of wheat yield in Australia at a level of detail not previously possible.

In this study we sought to: (i) identify key spatial patterns in Australian wheat yield; (ii) identify groups of years in the historical record in which spatial patterns were similar; and (iii) examine association of wheat yield year groups with indicators of ENSO, both before the wheat season (predictive) and during the season (descriptive).

**Methods**

*Simulation of historical time series of wheat yield*

A historical time series of shire wheat yields was generated for the 99-year period from 1901 to 1999 for 284 wheat-growing shires in Australia using the shire wheat yield model of Stephens (1995) with historical climate data (http://www.bom.gov.au/silo/, 2000) in the manner described by Hammer et al. (1996b) and Stephens et al. (2000). The simulation was conducted at 284 locations throughout the Australian wheatbelt. At each location the agro-climatic stress index model described by Stephens et al. (1989) was used to generate simulated wheat yields. Basically, the model simulates an annual wheat water stress index by utilising daily rainfall with average weekly radiation, and maximum and minimum temperatures to drive a simple water balance algorithm. The model uses location-specific soil data and crop-specific water requirements. Accumulating the relative water stress of the crop throughout the growing season derives the annual stress index (SI). A weighted SI was calculated for each wheat-producing shire from the values for points falling inside that shire. Point values were weighted by the relative area of the shire they represented by constructing Thiessen polygons around each point. The shire SI was then related to historical actual shire wheat yields for 1975–93.
(Australian Bureau of Statistics 1975–93) by linear regression analysis. A term was included in the linear regression to account for any technology trend over time in the shire yield data. The technology trend was included to account for any time trend in the data that was generated by factors other than the shire SI (e.g. varieties, management). The resultant technology trend was examined by Cornish et al. (1998) and is not considered in this study.

The variance of actual wheat yield explained by the regression analysis ranged from 82 to 91% at State level and was 90% at national level (Hammer et al. 1996b). These regression models were then used in the long-term simulation to create a simulated historical shire wheat yield time series from 1901 to 1999, using 2000 as the base technology year. This extrapolation of wheat yield in time by simulation was valid due to the high level of correlation in the regression models (Fiering 1963). The simulation was conducted using historical long-term daily rainfall data and long-term average temperature and radiation data. Hence, the simulated yields reflected only year-to-year variation in the seasonal rainfall conditions. The simulated yields were tested qualitatively by comparison of patterns of spatial averages with previous studies and quantitatively by spatial correlation analysis for each of the 19 years where historical shire yield data were available, as described below.

**Analysis of spatial yield patterns**

Principal component analysis (Johnson and Wichern 1988; Wilks 1995) was conducted on the wheat yield data matrix to determine the major underlying orthogonal factors representing the spatial variability among shires. The wheat yield data matrix consisted of simulated yield values in each of the 99 years for each of the 284 shires. The incremental proportion of total variance explained was used to determine the number of orthogonal factors retained. To facilitate the spatial interpretability (Richman 1986) and to maximise the variance between components and minimise the within-component variance, these factors were rotated using the varimax normal rotation method (Kaiser 1958). The factor loading of each rotated principal component for each shire was mapped and key spatial associations among shires were examined.

**Analysis of year types**

The principal components represent the main determinants of spatial variability among the shires. The time series of factor scores represent the temporal variation associated with each principal component. The years in these time series were allocated into 3 equal groups depending on whether the factor score for that year was negative, near zero, or positive. Boundaries between groups were designed such that one-third
of all years was contained in each group. Spatial yield patterns associated with each group of years were examined by mapping the average deviation from the simulated long-term average wheat yield (from 1901 to 1999) by shire for that year group.

**Association with ENSO indicators**

One indicator of ENSO that is coincident with the wheat growing season and one that precedes the growing season were used to examine associations with yield deviations in the spatial zones and year types identified in the principal component analysis. For the coincident indicator, the classification of years into El Niño, La Niña, or other (http://www.dnr.qld.gov.au/longpdk/hotline; R. C. Stone, pers. comm., 1999) was used. For the indicator preceding the season, 3 groups were derived from the 5 SOI phases (April/May) of Stone and Auliciems (1992). Combining SOI ‘consistently negative’ with SOI ‘rapidly falling’, SOI ‘consistently positive’ with ‘rapidly rising’, and SOI ‘near zero’ derived the 3 groups used. This grouping was required to ensure that an adequate number of occurrences occurred in each group for valid statistical testing of associations (De Groot 1986). Two-way tables were derived to examine association of years from the wheat analysis with years from the ENSO indicator analysis for each of the spatial zones. Statistical tests ($\chi^2$) of significance were conducted.

**Results and discussion**

**Simulated wheat yields**

At the national scale and assuming year 2000 technology, average simulated wheat yield for the period 1901–99 was 1.73 t/ha. Average simulated yields were greater in coastal areas, especially in the south-east, and decreased with distance from the coast (Fig. 2). This reflects experience and is consistent with previous agro-climatic analyses (Nix 1975). Annual average simulated yield for Australia ranged from 0.97 to 2.21 t/ha (Fig. 3) and had a coefficient of variation (CV) of 14%. The CV was greater in the northern wheatbelt and the inland sector of New South Wales (NSW) (Fig. 4). It was generally lower in Western Australia and the shires in the higher rainfall areas nearer to the coast of Victoria and South Australia. These findings were consistent with the recent study of Scoccimarro et al. (1994) on wheat yield variability in Australia. These observations provided a qualitative test of the model and indicate its satisfactory performance.

Spatial correlations between simulated and observed shire yields in any of the 19 years with actual data (1975–93) were high and averaged 0.86. Fig. 5 shows the deviation of the observed from the predicted shire yields for 1983—a year with spatial correlation close to the average. Although the simulation reflects the broad spatial pattern of shire yields well, there are a number of discrepancies in the detail. Nonetheless, the figure indicates the significant skill of the model in capturing the spatial pattern of wheat yield throughout the Australian wheatbelt. Examination of yield deviation maps for each of the 19 years showed that there was no systematic spatial bias (data not shown). The spatial correlation ranged from 0.71 to 0.91 and there were only 3 occasions when it was below 0.8, indicating a consistent predictive capability over diverse seasonal conditions. Hence, in addition to the high predictive capacity at aggregate State and national levels (Hammer et al. 1996b), the model maintains significant skill at shire scale. Testing on years that are independent of the training data set will be difficult, as the Australian Government has recently withdrawn funding for routine collection of shire production statistics.
Spatial yield patterns

In the analysis of variability among shires, the first 6 principal components (PCs) explained 69.3% of the total variance in wheat yield among shires (Table 1). Subsequent PCs contributed to the explained variance at a diminished marginal rate. As the first 6 PCs retained a large fraction of the variance in the original data, they were used to examine spatial associations among shires. Hence, this analysis reduced the spatial dimensionality from 284 shires to 6 spatially orthogonal modes of variance. The rotated factor loadings of each PC were mapped in order to examine the key spatial associations (Fig. 6a–f). PC1 explained 27.9% of the total variation in yield among shires and had strong positive loadings in South Australia, Victoria, and southern NSW (Fig. 6a). PC2 explained 11.9% of the total variance and had a strong positive loading in Queensland and northern NSW (Fig. 6b). PC3 and PC6 separated 2 distinct areas in Western Australia from the rest of Australia and explained 9.2% and 5.6% of the total variance, respectively (Fig. 6c and f). PC4 explained 8.7% of the total variance and had a strong positive loading in central and northern NSW (Fig. 6d). The shires nearer to the coast of southern Victoria were spatially delineated through PC5, which explained 6.1% of the total variance (Fig. 6e).

Each of the 6 PCs delineated a coherent regional area and 5 were spatially distinct. The PCs identify independent modes of variation of the total variation in the time series of simulated Australian shire wheat yields. The regional areas identify groups of shires that contribute similarly with respect to these independent components. As the first 5 regional areas delineated were spatially distinct, they can be considered to separate the wheatbelt into distinct homologous wheat zones (HWZ) that show similar yield variation over years. It is only through PC4 (Fig. 6d), which covers most of NSW, that a second dimension is required to
fully explain the underlying temporal variation. The southern half of NSW groups with areas in Victoria and South Australia in expressing variation identified via PC1 (Fig. 6a). The northern half of NSW, however, groups with areas in Queensland in expressing variation identified via PC2 (Fig. 6b). Both of these parts of NSW group together and separately from all other areas in expressing variation via PC4 (Fig. 6d). Hence, two distinct dimensions are needed to describe the temporal variation for shires in NSW.

Year types
The years associated with positive, near zero, and negative factor scores on each PC are shown in Table 2. Sample maps of average deviations from the simulated long-term average shire wheat yield for the 3 year groups derived from PC1 are presented in Fig. 7. As expected, the spatial yield anomaly patterns in the year groups closely mimic the spatial patterns from the principal component analysis (Fig. 6a–e). For example, the spatial extent of the anomalies for the negative and positive year types of PC1 (Fig. 7a and b, respectively) corresponds well with the spatial extent of the factor loadings for PC1 (Fig. 6a). The largest anomalies occurred mainly in the shires with the highest factor loadings. Furthermore, a negative average factor score is associated with a negative yield anomaly and vice versa. Similar results (data not shown) were found for the yield anomaly patterns associated with the remaining PCs.

### Table 1. Spatial representation and variance explained for the first 6 rotated principal components (PC1–PC6) from the analysis of simulated shire wheat yields for the period 1901–99

<table>
<thead>
<tr>
<th>Factor</th>
<th>Spatial representation</th>
<th>Variance explained (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PC1</td>
<td>South Australia, Victoria, southern New South Wales</td>
<td>27.9</td>
</tr>
<tr>
<td>PC2</td>
<td>Queensland, northern New South Wales</td>
<td>11.8</td>
</tr>
<tr>
<td>PC3</td>
<td>North-east inland Western Australia</td>
<td>9.2</td>
</tr>
<tr>
<td>PC4</td>
<td>Central, northern New South Wales</td>
<td>8.7</td>
</tr>
<tr>
<td>PC5</td>
<td>Coastal shires in Victoria</td>
<td>6.1</td>
</tr>
<tr>
<td>PC6</td>
<td>South-west coastal Western Australia</td>
<td>5.6</td>
</tr>
<tr>
<td>Total:</td>
<td>69.3</td>
<td></td>
</tr>
</tbody>
</table>

**Fig. 5.** Deviation of observed from predicted shire wheat yield (t/ha) across Australia for 1983.
Fig. 6. Spatial representation of significant factor loadings of shires across Australia for each principal component (PC) from the spatial analysis. The time series of the factor score for each PC is included with the map for that PC: (a) PC1, (b) PC2.
Fig. 6. (continued) (c) PC3, (d) PC4.
Fig. 6. (continued) (e) PC5, (f) PC6.
and consequent yield reduction with high rainfall. Much of
patterns is associated with regional differences in soil type,
time among the regions. Some of this variation in yield
of PCs 1, 4, and 5. It is in the near zero third for PC2, and
the same group level across all PCs. For example, the major

Inspection of Table 2 indicates that there are few years in
the same group level across all PCs. For example, the major
drought and El Niño year of 1982 is in the negative third only
for PCs 1, 4, and 5. It is in the near zero third for PC2, and
the positive third for PCs 3 and 6. Similarly, the generally wet
following year (1983) is in the positive third only for PCs 1,
2, and 4. Hence, although temporal variation in yield is
consistent within the regions defined by the PCs (Table 1,
Fig. 6), there are significant differences in yield patterns with
time among the regions. Some of this variation in yield
patterns is associated with regional differences in soil type,
which introduces differences in likelihood of waterlogging
and consequent yield reduction with high rainfall. Much of
the variation, though, is most likely associated with regional
differences in rainfall in any given year (Stephens and Lyons
1998).

Relationship of year types to ENSO and SOI phases

There was a strong and significant association of the year
groups within each PC with the concurrent ENSO year type
classification. Table 3 gives the $P$-values of the chi-square
test of non-randomness when examining distributions of
year types among ENSO categories. All PCs, except PC3
($P$-value: 0.47; north-east inland Western Australia), showed
highly significant associations. $P$-values ranged from 0.000
(PC2, Queensland) to 0.03 (PC6, coastal Western Australia).

Strong or slightly stronger relationships were also found
between the year types and ENSO when the year type
indicator that preceded the growing season was used. By
forming 3 groups from the 5-group SOI phase system of
Stone et al. (1996), the chi-square test of non-randomness
revealed significant $P$-values, except for PC6 (0.27; coastal
Fig. 7. Average deviation of year types from long-term average shire yield using year groupings based on (a) lowest third, and (b) highest third of factor loadings for PC1. See Table 2 for actual years.
The analyses presented here identified homologous modes and zones of wheat yield variation, each containing specific analog years (i.e. year types) for Australian wheat. Although the predictive SOI phase system of Stone et al. (1996) showed significant skill in discriminating among most of these year types, considerable potential for improved discrimination remained. By examining the historical years forming the analog sets within each zone, it may be possible to identify climate system or ocean–atmosphere features that may be causal and would be most useful in improved predictive schemas. Alternately, examining associations of year groups with other climate forcing factors and mechanisms such as the Indian Ocean dipole (Nicholls 1989) and the Antarctic circular polar wave (White 2000) may be illuminating. Any improved system that is better able to discriminate among these year groupings for the homologous wheat zones would be of immediate relevance and utility.

Challenges for seasonal forecasting and climate research

The significant relationship between year types and pre-season ENSO indicators suggests an obvious role for seasonal forecasting. But, although significant, the discrimination remains weak. Table 4 shows the number of years in each PC3 (inland Western Australia) factor score group that is associated with SOI phase groups. Although the shifts from random are noticeable (and result in the significant chi-square statistic), the association is far from perfect. Can climate research do better? Can we identify seasonal forecasting systems or climate classification systems that are better able to discriminate among wheat year types? Or is the remaining noise unpredictable?

**Table 4.** Observed number of years from negative (–ve), near zero (nz), and positive (+ve) groups of PC3 occurring by groups of the SOI phase in April–May at the start of the wheat season in the relevant year

<table>
<thead>
<tr>
<th>SOI phase</th>
<th>–ve</th>
<th>PC3 group</th>
<th>+ve</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>I and III</td>
<td>15 (9)</td>
<td>5 (9)</td>
<td>7 (9)</td>
<td>27</td>
</tr>
<tr>
<td>II and IV</td>
<td>12 (15.67)</td>
<td>16 (15.67)</td>
<td>19 (15.67)</td>
<td>47</td>
</tr>
<tr>
<td>V</td>
<td>6 (8.33)</td>
<td>12 (8.33)</td>
<td>7 (8.33)</td>
<td>25</td>
</tr>
<tr>
<td>Total</td>
<td>33</td>
<td>33</td>
<td>33</td>
<td>99</td>
</tr>
</tbody>
</table>

**Western Australia.** P-values ranged from 0.002 (PC2, Queensland) to 0.09 (PC1, south-eastern Australia). This can be partly explained by the strong association of the SOI phase system with ENSO. The SOI phase system is a skilful predictor of ensuing ocean/atmosphere dynamics when extreme states of the ENSO system (e.g. El Niño, La Niña, etc.) are likely (Stone et al. 1996).

In general, a significant relationship exists between ENSO and the wheat year types within each PC across the whole of Australia. The April/May SOI phase system was as good an indicator of the subsequent wheat year types as the system based on concurrent classification of ENSO during the growing season. This strengthens the case for use of the SOI phase system as a seasonal indicator in commodity forecasting systems. This is consistent with findings at field and farm scale (Hammer 2000) where the SOI phase system has proved effective and valuable in relation to farmer decision-making when implemented in a risk management context. Although the homologous wheat zones defined by the PCs demonstrated independently varying temporal patterns of simulated wheat yield (Table 2), they nonetheless all related significantly to year classifications based on ENSO indicators. This suggests that although the effect of ENSO varies among regions in any one year, there is an underpinning coherence (i.e. relationship to ENSO) over many years. This is consistent with findings of previous studies of association with wheat yields at broad scale (Rimmington and Nicholls 1993; Stephens et al. 2000).

Acknowledgments

We thank Peter deVoil at Queensland Department of Primary Industries for his programming assistance during this study. We also thank the Australian Bureau of Statistics in Brisbane for supplying the shire wheat yield census data.

References


Manuscript received 2 January 2001, accepted 26 July 2001