Aust. J. Agric. Res., 1993, 44, 609–24

Rainfall Patterns in a Major Wheat-Growing Region of Australia

Rizaldi Boer,[A] *David J. Fletcher*[B] *and Lindsay C. Campbell*[A]

[A] Department of Crop Sciences, University of Sydney, N.S.W. 2006.
[B] Department of Mathematics and Statistics, University of Otago, Dunedin, New Zealand.

Abstract

Rainfall is an important variable in the wheat production areas of Australia. This analysis examines, firstly, the pattern of rainfall over 2·3 million ha of a high-quality wheat-producing region, and secondly, develops regression equations for rainfall prediction over this region.

Most of the variation in rainfall pattern across the region is accounted for by differences in October-to-March (summer) rainfall and in April-to-September (winter) rainfall. The summer rainfall differences account for over two thirds of the variation. Based on these two rainfall periods, a partitioning of the study area reveals five distinct regions.

The second part of the analysis uses multiple regression to provide a set of equations for rainfall prediction at any location in the region, for a number of rainfall periods. These equations use altitude, longitude and latitude as predictors. Nearly all of the equations explain between 80% and 94% of the variation in rainfall. Differences between regions are accounted for in the analysis, making the equations widely applicable. The validity of the mean rainfall equations was tested on three further sites: the mean prediction error was 6·9%. This approach may be applicable where large land masses with similar geographical features occur.

Keywords: rainfall pattern, principal components analysis, cluster analysis, weighted regression, mallows rainfall prediction.

Introduction

Rainfall is one of the leading climatic factors which influences the growth of wheat. In Australia, wheat yields are strongly dependent on winter rainfall (Millington 1961). In South Australia, the yields of wheat are strongly related to April-to-October, April-to-May, June-to-August and September-to-October rainfall (Cornish 1950; French 1989). The optimal amount of rainfall during the last three periods varies between soil types. In the eastern and north-eastern sectors of the wheat belt of Australia (Fig. 1), summer rainfall also contributes to the total water available during the crop cycle. In general, about 10 to 30% of summer fallow rainfall is stored in the soil (Waring *et al.* 1958; Fitzpatrick and Nix 1969).

The amount of stored soil water extracted by a crop growing in a Vertisol on the Darling Downs in Queensland was linearly related to the amount of rainfall that fell in the growing season. An increase in growing season rainfall of 100 mm decreased the amount of soil water storage extracted by the crop by 84 mm (Marley and Littler 1989). A study of the rainfall pattern in a wheat region would therefore be useful for agronomic decision support systems and econometric

0004-9409/93/040609$05.00

models; i.e. the selection of fallowing technology, the selection of site for variety trials, and the assessment of potential yield in the region.

The pattern of rainfall in Australia is related to altitude (Linacre and Hobbs 1977; Hutchinson and Bischof 1983; Adomeit *et al.* 1987), latitude and longitude (Hutchinson and Bischof 1983), and distance from the coast (Linacre and Hobbs 1977; Bureau of Meteorology 1988). We have developed equations which can be used to predict rainfall using geographical information in a major wheat-growing area (Fig. 1).

Methods

Area of Study

The region selected for this study was the wheat-belt in the north and centre of New South Wales (NATMAP 1982; Nix 1987). The area of this region is 2·3 million ha, comprising about 20% of the Australian wheat-belt (vide inset A, Fig. 1) or 65% of the NSW wheat-belt

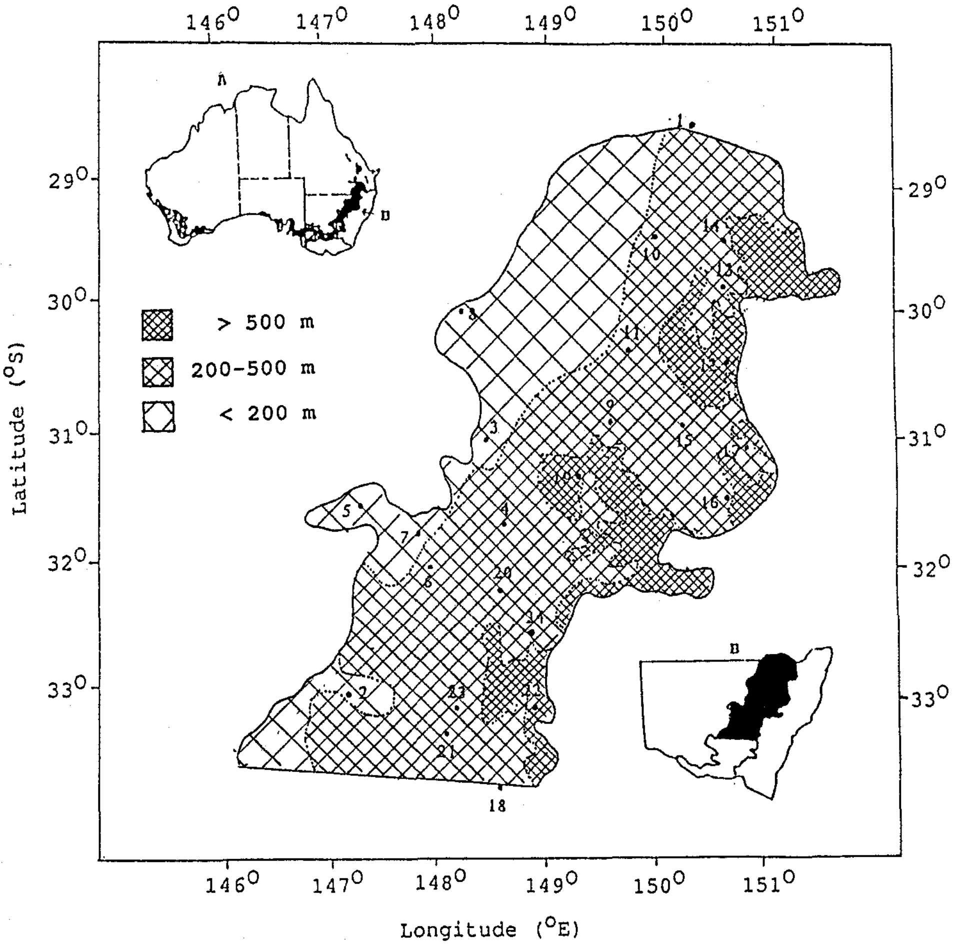

Fig. 1. The topography and geographical location of the meteorological stations. Inset A shows the Australian wheat-belt. Inset B delineates the study area in relation to the New South Wales wheat-growing regions. Numbers indicate the locations of the stations.

(vide inset B, Fig. 1); it produces 70% of Australia's prime hard wheat. A total of 24 weather stations were used in the analysis. Daily rainfall data at each station were obtained from the Australian Bureau of Meteorology. The locations of the stations and the topography of the region are shown in Table 1 and Fig. 1.

Table 1. Altitude, location, and mean annual rainfall (AR) for meteorological stations used in this study

Note: The records are for 1916–1990, except those from stations 1, 10 and 11, which are for 1907–1983, 1916–1965 and 1935–1976 respectively

No	Station	Alt. (m)	Lat. (°S.)	Long. (°E.)	AR (mm)
1	Goondiwindi	216·0	28·53	150·30	529
2	Condobolin	199·0	33·08	147·15	447
3	Coonamble	180·0	30·97	148·38	505
4	Gilgandra	278·0	31·72	148·67	563
5	Nyngan	177·0	31·56	147·20	440
6	Trangie	219·0	32·03	147·99	518
7	Warren	197·0	31·70	147·83	492
8	Walgett	131·0	30·02	148·12	483
9	Baradine	366·0	31·00	149·00	585
10	Moree	207·3	29·50	149·90	569
11	Narrabri	201·0	30·22	149·60	545
12	Barraba	500·0	30·38	150·62	694
13	Bingara	296·0	29·87	150·57	727
14	Warialda	320·0	29·55	150·58	675
15	Gunnedah	306·0	31·03	150·27	606
16	Quirindi	390·0	31·52	150·68	675
17	Tamworth	404·0	31·08	150·85	662
18	Cowra	360·0	33·80	148·70	622
19	Coonabarabran	509·0	31·28	149·28	740
20	Dubbo	275·0	32·22	148·57	597
21	Forbes	237·0	33·38	148·02	538
22	Molong	529·0	33·10	148·87	701
23	Parkes	339·0	33·13	148·18	603
24	Wellington	304·0	32·57	148·91	629

Analyses

The analysis involved several steps (Fig. 2). Principal Components Analysis (PCA) was used to identify the aspects of the rainfall which varied most across sites. The data used in the PCA consisted of a monthly rainfall summary for each of the 24 sites; i.e. a 24 sites (Units)×12 months (Variables) array. The arrangement of the data was perfomed using the TAMSIM program (McCaskill 1990). Two types of summary were chosen: mean (MnR) and median (MdR). The covariance matrix was used in each of the PCAs, since the 12 variables were all of the same kind.

Interpretation of the principal components (PCs) was aided by inspection of the factor loading matrix. This is a matrix of correlations between the PCs and the original variables. After varimax rotation of this matrix (Richman 1986), the correlations between some of the variables and the components generally became high or low, with a small number remaining at an intermediate level; the interpretation of the PCs was then easier.

In order to partition the region according to rainfall pattern, Cluster Analysis was carried out on the first two PCs, which accounted for most of the variance. The Furthest Neighbour algorithm was used to construct a dendogram. Determination of the optimal number of clusters was made by plotting the number of groups suggested by the dendogram versus the clustering level, i.e. the stage at which the current grouping merged. The optimal number of groups was chosen to be at a point just before the curve began to plateau (Aldenderfer and Blashfield 1984). A detailed explanation of PCA and Cluster Analysis is available in Chatfield and Collins (1980) and Krzanowski (1988).

Weighted Regression Analysis was used to develop equations for predicting rainfall in the region. The use of weighting was motivated by the need to allow for the different levels of precision for the two summary statistics (MnR and MdR) across the sites. These were caused by the differences in year-to-year variation and by differences in length of record. The theory underlying weighted regression is described by Myers (1990). The weights used were based on the reciprocal of the variance of the MnR or MdR for the site concerned. The variance for MnR was estimated by s^2/n, where s^2 is the sample variance and n the length of record from that site. The variance for MdR was estimated using the bootstrap method (Efron 1982). The predictor variables choosen were altitude, latitude and longitude; prior to analysis, these were centred to avoid potential numerical problems.

In order to allow the form of the equation to differ across the study area, the sites were grouped according to the partitioning described above. The use of indicator variables then allowed for different regression coefficients in the different groups. In assessing the homogeneity of these coefficients across the groups, a simple F test of the appropriate variable-by-group interaction could have been used. This would have had the disadvantage of not being able to indicate whether two of the groups had equivalent coefficients, with the third group being significantly different. Such an assessment could have been made using a procedure equivalent to any of the multiple comparison methods familiar in analysis of variance (O'Neill and Wetherill 1971). In order to avoid some of the problems associated with such methods, we applied a simple graphical technique analogous to that used by Perry (1986) for grouping treatment means. Full details of this procedure are given in Fletcher and Boer (1992). The final selection from the best candidate models was performed using Mallows' C_p criterion (Mallows 1973; Myers 1990).

To summarize the fit of each equation, a weighted least squares version of R^2 (denoted R^2_{wls}) was calculated (Willet and Singer 1988), as well as the square root of the mean of the

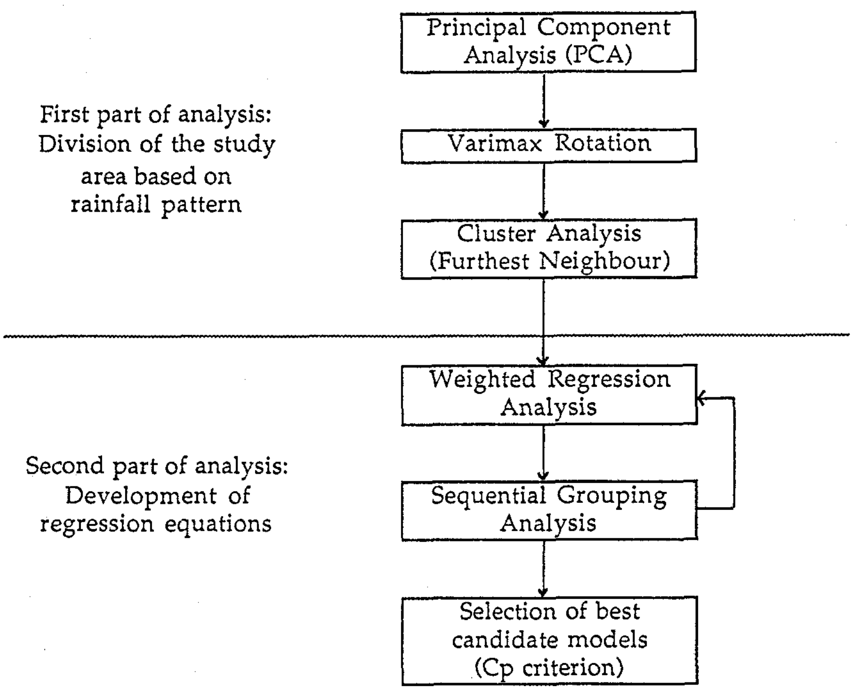

Fig. 2. A diagrammatic representation of the steps in the statistical analysis to produce regression equations for rainfall as a function of geographical variables.

squared residuals (RMS) for each region. The latter indicate the relative fit of the equation in the three regions. Visual examination of plots of residuals versus each predictor were used to detect any possible quadratic effects.

The validity of the equation for mean seasonal rainfall was examined using data from sites other than those used to develop the equations (Bureau of Meteorology 1988). The stations chosen were Quambone (30·93 °S., 147·86 °E., 154 m), and Peak Hill (32·70 °S., 148·18 °E., 267 m) and Pindari Dam (29·40 °S., 151·23 °E., 472 m).

Results

Principal Components Analysis

From the PCA it was considered that the first two principal components were sufficient to summarize the variation of the original data between sites. They accounted for 88·7% and 89·5% of the variance for MnR and MdR respectively.

In the rotated factor loading matrix, the first component (PC1) was highly correlated with the October-to-March rainfall, and the second (PC2) with the April-to-September rainfall (Table 2). These suggest that PC1 appears to be a summer rainfall component and PC2 a winter rainfall component. As PC1 accounts for 67·9% and 67·5% of the variance for MnR and MdR respectively, the variation in monthly rainfall is mainly determined by the rainfall during the summer months.

Table 2. Correlations with the monthly rainfall for the first two components (PC1 and PC2)

Month	MnR		MdR	
	PC1	PC2	PC1	PC2
January	−0·943	−0·009	−0·960	0·117
February	−0·952	−0·037	−0·914	−0·083
March	−0·638	−0·436	−0·594	0·557
April	0·003	−0·896	−0·274	0·891
May	−0·140	−0·935	−0·019	0·868
June	−0·264	−0·894	−0·186	0·933
July	−0·430	−0·868	−0·156	0·957
August	−0·114	−0·957	−0·062	0·984
September	−0·438	−0·866	−0·502	0·803
October	−0·671	−0·672	−0·770	0·523
November	−0·886	−0·386	−0·927	0·279
December	−0·895	−0·257	−0·940	0·173

Rainfall Pattern

When the first two PCs account for a high percentage of the variation, a plot of PC1 *v.* PC2 can be a useful way of looking for clusters (e.g. Gadgil and Joshi 1978). Initial examination of these plots did not suggest an obvious pattern and therefore Cluster Analysis was performed on these two components (Fig. 3*a*).

The Furthest Neighbour algorithm was used to produce dendograms. The choice of threshold level used to determine clusters and subclusters was facilitated by plotting the number of clusters versus the level of clustering (Aldenderfer and Blashfield 1984). The graph of the number of clusters versus the level of clustering suggested that the study area could be divided into three main regions (Fig. 3*b*). Regions 2 and 3 were further divided into two subregions each (Figs 4*a* and 4*b*).

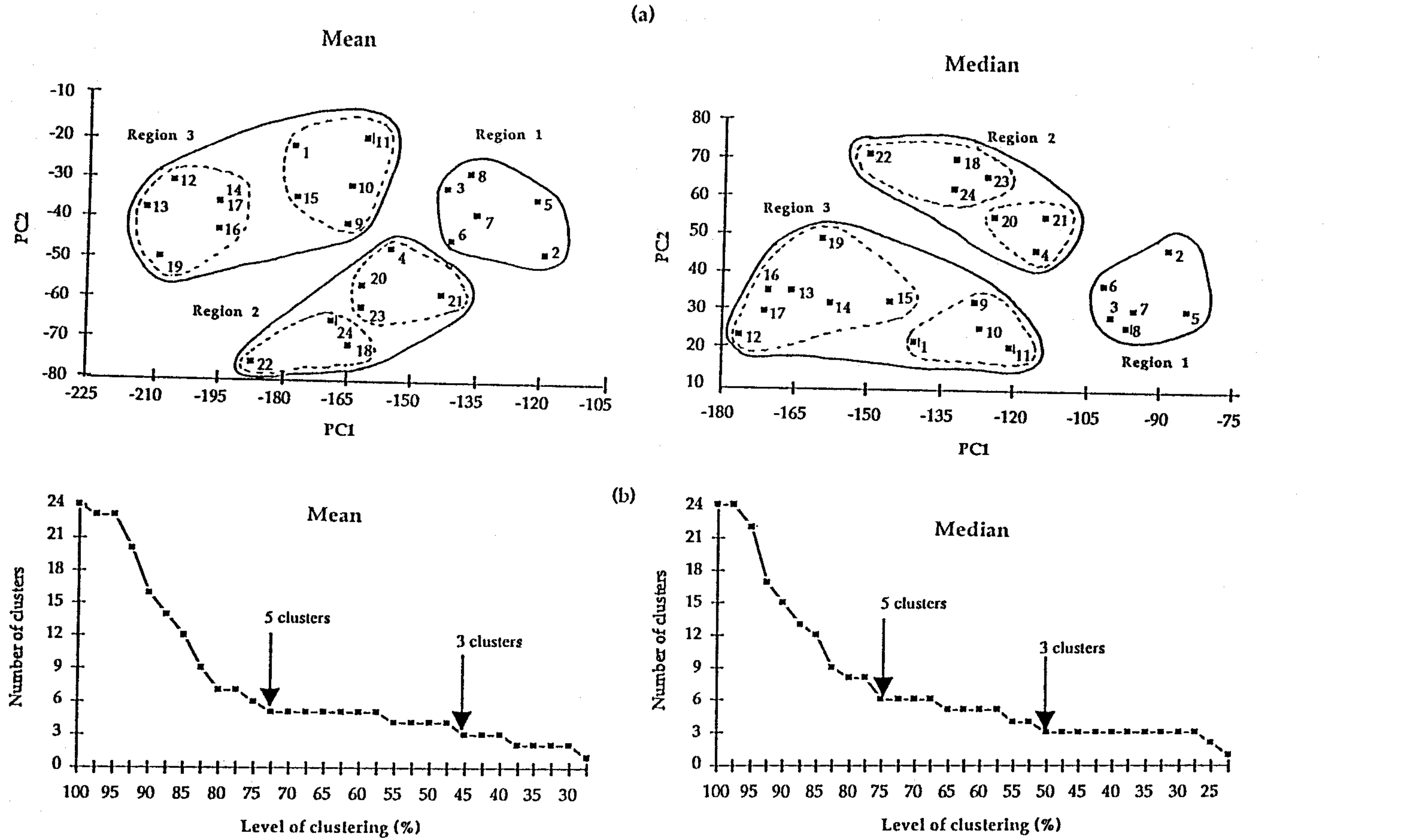

Fig. 3. (a) Plots of first principal component (PC1) versus second principal component (PC2), showing station numbers, clusters (solid lines), and subclusters (dotted lines), for mean and median rainfall. (b) Plots of the number of clusters versus level of clustering for mean and median rainfall. Arrows indicate points just before and at flattening of the curve.

Winter and summer rainfall are not very different in regions 1, 2a, 2b for both MnR and MdR. Region 1 is inland, while regions 2a and 2b are all south of 31° S. There is an increase in rainfall as altitude increases from west to east (Fig. 4; cf. Fig. 1). On the other hand, there is a marked difference between summer and winter rainfall in regions 3a and 3b (both of which are north of 32° S.). The rainfall again increases with altitude (Fig. 4; cf. Fig. 1). The region between 31° S. and 32° S. (excluding region 1) is therefore considered a transitional zone.

The ratios of summer to winter rainfall for regions 3a and 3b are more than 1·3, while in the westerly and southerly regions (regions 1, 2a and 2b) they are less than 1·3 (Table 3; cf. Fig. 4). Hoy (1978) has defined a location which has a ratio of summer to winter rainfall in excess of 1·33 as a summer rainfall maximum area, one for which it is less than 0·77 as a winter rainfall maximum area, and one for which it is between 0·77 and 1·33 as a uniform distribution area. Using this definition, the inland and the region south of 31·5° S. is therefore a uniform

Table 3. Mean (MnR) and median (MdR) rainfall in winter and the ratio summer: winter rainfall in the regions

Region	MnR (mm) Winter	Ratio	MdR (mm) Winter	Ratio
1	210	1·30	159	1·20
2a	267	1·16	220	1·05
2b	312	1·09	265	0·99
3a	224	1·59	171	1·61
3b	274	1·54	218	1·57

distribution area and that north of 31·5° S. a summer rainfall maximum area (Table 3).

Regression Equations for Rainfall

The variables analysed in the development of equations for the prediction of rainfall (mm) based on geographical information were:

(1) Annual rainfall (R1–12)
(2) Summer rainfall (R10–3)
(3) Winter rainfall (R4–9)
(4) April-to-May rainfall (R4–5)
(5) June-to-August rainfall (R6–8)
(6) September-to-October rainfall (R9–10)
(7) April-to-October rainfall (R4–10).

These periods were chosen because of their relevance to key agronomic events, namely, length of growing season, sowing time, vegetative and reproductive growth, and possible storage of water in the soil profile.

Inspection of the correlation coefficients showed that the geographical variable which was highly correlated with most of the rainfall variables was altitude. Longitude tended to be highly correlated with summer rainfall variables, as was latitude with winter rainfall variables (Table 4).

Eklundh and Pilesjo (1990) introduced two hybrid variables, namely, (longitude+latitude)/2 and (longitude−latitude)/2. In our analysis these two predictors were not used. Their use implies that the effects of longitude and latitude on

Table 4. Correlation coefficients for the rainfall and geographical variables in the wheat-belt of north and centre of New South Wales

	MnR			MdR		
	Alt.	Lat.	Long.	Alt.	Lat.	Long.
R1–12	0·839	−0·132	0·749	0·867	−0·112	0·732
R10–3	0·681	−0·469	0·908	0·685	−0·414	0·908
R4–9	0·826	0·444	0·264	0·809	0·479	0·239
R4–5	0·680	0·575	0·036	0·625	0·474	0·109
R6–8	0·830	0·411	0·290	0·811	0·494	0·235
R9–10	0·860	0·156	0·613	0·789	0·081	0·705
R4–10	0·846	0·380	0·351	0·841	0·417	0·307
Alt.	1·000					
Lat.	0·197	1·000				
Long.	0·460	−0·597	1·000			

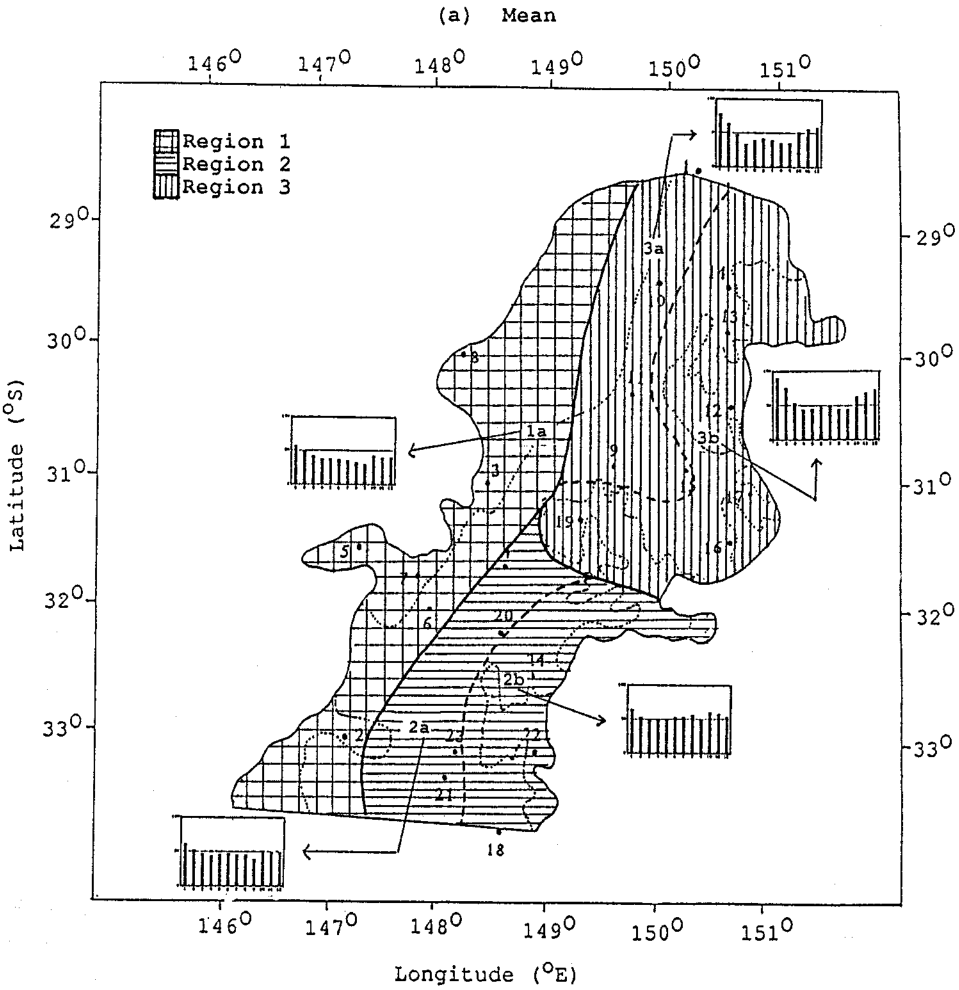

Fig. 4. Rainfall patterns in the wheat-belt of the north and centre of New South Wales for (*a*) mean rainfall and (*b*) median rainfall.

(b) Median

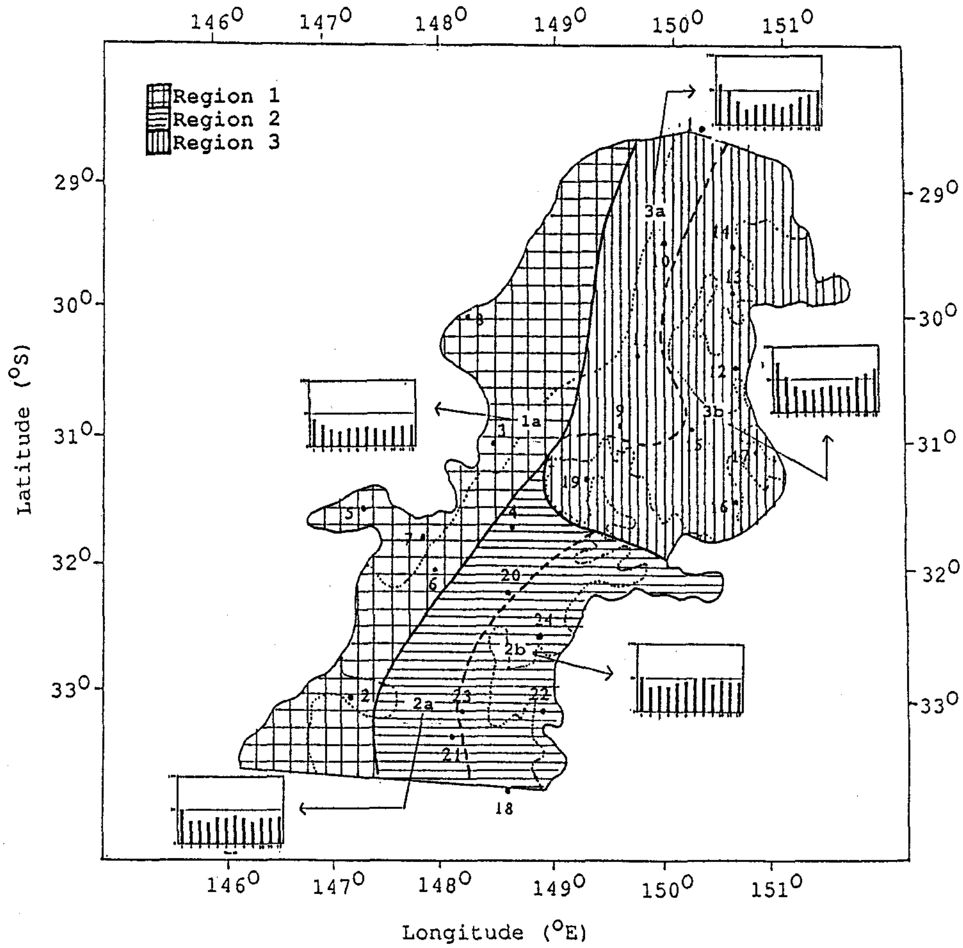

Fig. 4. (continued).

rainfall are equal. Keeping longitude and latitude separate in the equations avoids this restriction.

In the development of the regression equations (Fig. 2) the analysis suggested that the seasonal rainfall in each region could be predicted using one of a number of alternative models. This was due to the existence of more than one possible combination of the coefficients at a particular stage of the sequential process (Fig. 2). The model which gave the smallest value for Mallows' C_p was chosen from amongst these candidate models. Most of the error mean squares for the candidate models were smaller than the error mean squares for the full model (a model which used all 12 predictors).

The resulting equations explained up to 93% and 94% of the variation (R^2_{wls}) for MnR and MdR respectively (Tables 5 and 6). With two exceptions the altitude effect was significant for all seasonal rainfalls in each region, whereas latitude and longitude were not so consistently influential. In general, the effect of altitude on rainfall during winter was the same across all regions (Tables 5 and 6). For example, the regression equation for estimating mean winter rainfall (Table 5) is:

Table 5. Regression equations for mean rainfall (MnR) in the wheat-belt of north and centre of New South Wales

All coefficients of the equations were significant at 5% level. The predictors alt, lat and long were all centred before analysis; thus alt is altitude minus 300 m, lat is latitude minus 30° and long is longitude minus 150°. The numbers in in parentheses are the standard errors of the coefficients. The variable reg followed by number 1, 2, and/or 3 is the label for the intercept in regions 1, 2 and/or 3 respectively; e.g. in R1–12, 589·9reg12 means the intercept for R1–12 in regions 1 and 2 is 596·9

Equations	RMS (mm) Region 1	2	3	R^2_{wls} (%)	Sites with high residuals
R1–12 = 589·9reg12+619·3reg3+0·5183alt2+0·3840alt3+49·8long1+37·9long3	14·0	19·0	38·5	88	13, 19
(12·5) (10·5) (0·1466) (0·0982) (7·3) (16·2)					
R10–3 = 385·1reg123+0·3324alt1+0·2332alt23–11·4lat123+24·7long123	3·2	9·6	20·5	93	13
(5·1) (0·1032) (0·0474) (4·3) (6·5)					
R4–9 = 238·2reg13+274·8reg2+0·2539alt123	9·8	10·2	18·4	85	12, 13
(3·6) (5·8) (0·0299)					
R4–5 = 79·2reg123+0·0551alt123+4·5lat2	5·5	4·4	6·9	70	1
(1·5) (0·0124) (1·0)					
R6–8 = 142·9reg12+123·5reg3+0·1416alt123+9·8long1	3·6	5·6	10·5	86	12, 13
(3·1) (2·7) (0·0215) (2·0)					
R9–10 = 92·4reg123+0·0588alt123+5·1lat123+8·6long123	2·1	3·6	6·3	88	13
(1·5) (0·0158) (1·3) (1·7)					
R4–10 = 297·0reg123+0·2197alt123+10·3lat123+17·2long1	8·5	15·2	22·2	84	13
(5·4) (0·0466) (2·8) (4·5)					

Table 6. Regression equations for median rainfall (MdR) in the wheat-belt of north and centre of New South Wales

All coefficients of the equations were significant at 5% level. The predictors alt, lat and long were all centred before analysis; thus alt is altitude minus 300 m, lat is latitude minus 30° and long is longitude minus 150°. The numbers in in parentheses are the standard errors of the coefficients. The variable reg followed by number 1, 2, and/or 3 is the label for the intercept in regions 1, 2 and/or 3 respectively; e.g. in R1–12, 648·3reg12 means the intercept for R1–12 in regions 1 and 2 is 648·3

Equations	RMS (mm) Region 1	2	3	R^2_{wls} (%)	Sites with high residuals
R1–12 = 648·3reg12+610·4reg3+0·5737alt123+45·3long123	17·4	16·6	35·8	89	13, 19
(18·2) (8·0) (0·0580) (9·7)					
R10–3 = 360·9reg13+296·0reg2+0·2745alt123–14·1lat3+33·6long13	9·8	16·3	19·3	91	—
(5·0) (6·5) (0·0512) (6·7) (4·0)					
R4–9 = 231·6reg123+0·2848alt12+0·1260alt3+12·4lat23	10·7	9·4	16·4	88	12, 13
(4·0) (0·049) (0·0415) (2·3)					
R4–5 = 98·9reg12+69·4reg3+0·0449alt3+14·5long12	5·9	5·9	5·0	60	13, 23
(7·1) (1·9) (0·0185) (3·7)					
R6–8 = 118·9reg13+141·3reg2+0·1393alt123	5·0	5·6	9·7	87	13
(1·9) (3·2) (0·0158)					
R9–10 = 37·2reg1+82·2reg23–0·2804alt1+0·0621alt2+16·8lat1+8·8lat23+17·0long123	2·8	2·9	5·1	94	—
(15·4) (1·6) (0·0987) (0·0234) (3·1) (1·2) (1·7)					
R4–10 = 284·1reg123+0·3636alt12+0·1940alt3+12·2lat23	9·9	10·9	21·0	86	13
(4·6) (0·0545) (0·0528) (2·5)					

$$\text{Ry} - 9 = 238 \cdot 2\text{reg13} + 274 \cdot 8\text{reg2} + 0 \cdot 2539 \text{ alt123}.$$

Thus the mean winter rainfall at any location in region 1 or 3 is estimated by:

$$\text{Ry} - 9 = 238 \cdot 2 + 0 \cdot 2539 \text{ alt},$$

while in region 2 it is estimated by

$$\text{Ry} - 9 = 274 \cdot 8 + 0 \cdot 2539 \text{ alt}.$$

In addition to R^2_{wls} values, Tables 5 and 6 show the values of the RMS summary for each region. It can be seen that these are generally higher in region 3, indicating that the equations tend to fit better in regions 1 and 2.

Validation

The prediction errors in the validation study ranged from 0·3% to 20·3%, with a mean of 6·9% (Table 7). With the exception of winter rainfall at Pindari Dam (region 3) all the prediction errors were 10% or less.

Table 7. Comparison between long-term mean (LTM) and estimated (EST) seasonal rainfall of Quambone, Peak Hill and Pindari Dam

	Quambone			Peak Hill			Pindari Dam		
	LTM (mm)	EST (mm)	Error (%)	LTM (mm)	EST (mm)	Error (%)	LTM (mm)	EST (mm)	Error (%)
R1–12	442	483	9·3	554	573	3·4	772	732	5·2
R10–3	249	273	9·6	301	302	0·3	488	463	5·1
R4–9	193	201	4·1	253	266	5·1	284	282	0·7
R4–5	66	71	7·6	86	90	4·6	106	89	16·0
R6–8	98	101	3·1	131	138	5·3	123	148	20·3
R9–10	66	70	6·1	85	88	3·5	132	110	16·7
R4–10	230	238	3·5	302	317	5·0	361	329	8·9

Discussion

Callaghan and Millington (1956), Hoy (1978) and Nix (1987) considered that May to October are winter rainfall months and November to April are summer rainfall months in Australia. However, in our study it appears more natural to define summer and winter rainfall months as those which correspond to PC1 and PC2 (Table 2), viz. October-to-March and April-to-September respectively.

In South Australia, the variation in wheat yields between seasons has been found to be closely related to rainfall variation (Millington 1961; French 1989). The variation of rainfall can be expressed as a Variability Index, i.e. the ratio of (90% percentile–10% percentile) to the 50% percentile. In our study, the Variability Index for winter rainfall in the north west was 1·1, while elsewhere it was lower (Fig. 5). Furthermore, Gregory and Cooke (1986) found that in the north-west, every two years the total rainfall for a 3, 4 and 6 month period would be less than 35, 40 and 55% of the long-term average respectively. Elsewhere these percentages were higher. Similarly, every 5 years the total rainfall for the corresponding periods would be less than 15, 25 and 35% of the long-term average respectively. These suggest that the variability in yield in our northern region would be higher than that expected in the southern region. However, as rainfall in the north is summer dominant (Fig. 4) and soils have greater water storage capacity than those in the south (Forrest *et al.* 1985), the difference in the variability in yields between the northern and southern regions may be partly offset by the stored soil water obtained during the summer.

Regions 1 and 2 (Fig. 4) are considered to be uniform distribution areas, while the neighbouring regions (3a and 3b) are considered to be summer maximum rainfall areas. Linacre and Hobbs (1977) noted that convectional uplift, tropical cyclones and monsoons are factors affecting summer rainfall. This suggests that the effect of those factors is consistent with the decreases in this rainfall from east to west and from north to south.

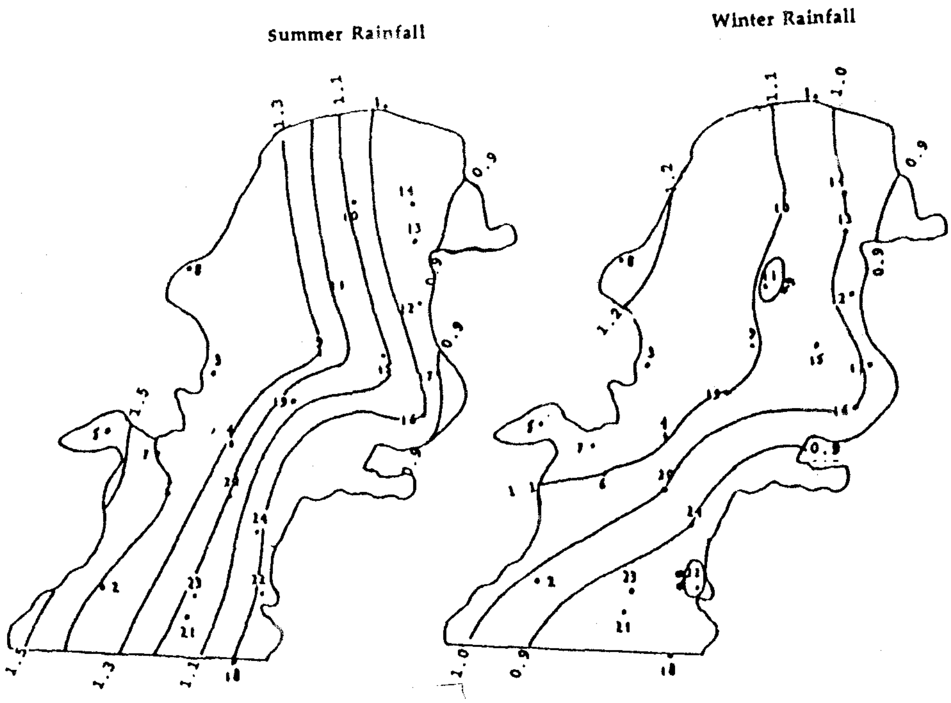

Fig. 5. Variability indices of summer and winter rainfall.

From the regression equations, it appears that altitude is the dominant factor affecting rainfall level. The effect of longitude appears to be stronger during summer, while that of latitude is stronger during winter (Table 4; cf. Tables 5 and 6). The equations for summer rainfall variables have higher R^2_{wls} values than those for winter rainfall variables; this suggests that the geographical variables considered are more strongly related to the former.

In general, the Root Residual Mean Squares (RMS) for region 3 were higher than those for regions 1 and 2. In particular, site 13 (Bingara) appeared to be an outlier for nearly all the equations. High residuals were also observed for sites 12 and 19 which are both at an altitude of greater than 500 m (Tables 5 and 6; cf. Fig. 1). The residual plots indicated that this lack of fit could not be explained using quadratic effects. This suggests that the equations fit less well in region 3b (Fig. 4), particularly at an altitude of more than 500 m. High residuals consistently occurred at Bingara (site 13), which has an altitude of 296 m. This was because it had higher rainfall than sites of similar altitude in region 3. A possible explanation for this is that Bingara is surrounded by areas with altitudes greater than 500 m (Fig. 1).

Application of the equations for predicting the mean seasonal rainfall at Quambone, Peak Hill and Pindari Dam, situated in regions 1, 2 and 3 respectively, shows that the equations perform particularly well in region 2. They perform less well for the prediction of mean winter rainfall at Pindari Dam, a high elevation site in region 3b (Table 7; cf. Fig. 4).

There are other techniques available for the spatial interpolation of weather parameters, e.g. kriging (Seaman and Hutchinson 1985), thin plate smoothing splines (Wahba and Wendelberger 1980), and the thiessen polygon weighting technique (De Jong *et al.* 1992). A comparative study of several interpolation methods was described by Laslett *et al.* (1987) and Hutchinson (1991). A recent computationally intensive technique, used for spatial interpolation of weather parameters across the Australian continent, involves thin plate smoothing splines (Hutchinson and Bischof 1983; Hutchinson *et al.* 1984; Hutchinson 1989, 1991). In applying this technique to mean monthly rainfall, the standard error was between 10% and 15%. The standard errors for our equations range from 1·2% to 7·2%, with a mean of 2·4%.

As Sumner (1988) argued, median rainfall is a more reasonable summary to use if the data are not normally distributed. However, mean rainfall is also commonly used in agricultural studies: e.g. as an indicator of the productivity of a region (Freebairn *et al.* 1991) or the prediction of mean wheat yield of a region (Nix 1987). Knowledge of both mean and median is preferable for any study involving the assessment of risk to the crop from extreme rainfall periods.

Since the equations explain such a high degree of the variation in rainfall over a major wheat-growing region, this type of approach may be used for developing rainfall estimation equations for other parts of Australia and for other countries with large land masses and similar geographical features.

Acknowledgments

We thank Malcolm R. McCaskill for supplying the TAMSIM program.

References

Adomeit, E. M., Austin, M. P., Hutchinson, M. F., and Stein, J. L. (1987). Annual mean rainfall and temperature surface contour maps of south-eastern Australia. CSIRO Aust. Div. Water Resources Res., Tech. Mem. 87/15.

Aldenderfer, M. S., and Blasfield, R. K. (1984). 'Cluster analysis. Series: Quantitative Applications in the Social Sciences.' No. 07–044. pp. 53–62. (Sage Publications: London.)

Bureau of Meteorology. (1988). 'Climatic averages of Australia.' (Department of Administrative Services, Bureau of Meteorology, Australian Government Publishing Service: Canberra.)

Callaghan, A. R., and Millington, A. J. (1956). 'The Wheat Industry in Australia.' 1st Edn. (Angus and Robertson: Melbourne.)

Chatfield, C., and Collins, A. (1980). 'Introduction to Multivariate Analysis.' pp. 57–81. (Chapman and Hall: New York.)

Cornish, E. A. (1950). The influence of rainfall on the yield of wheat in South Australia. *Aust. J. Sci. Res.* B**3**, 178–218.

De Jong, R., Dumanski, J., and Bootsma, A. (1992). Implication of spatial averaging weather and soil moisture data for broad scale modelling activities. *Soil Use and Management* **8**, 74–9.

Efron, B. (1982). 'The Jackknife, the Bootstrap and other Resampling Plans.' pp. 27–36. (Society for Industrial and Applied Mathematics: Philadelphia, U.S.A.)

Eklundh, L., and Pilesjo, P. (1990). Regionalization and spatial estimation of Ethopian mean annual rainfall. *Int. J. Climatol.* **10**, 473–94.

Fitzpatrick, E. A., and Nix, H. A. (1969). A model for simulating soil water regimes in alternating fallow-crop system. *Agric. Meteorol.* **6**, 303–19.

Fletcher, D. J., and Boer, R. (1992). A graphical technique for assessing the homogeneity of regression coefficients. Research Report. (Dept. of Mathematics and Statistics, University of Otago: Dunedin.)

Forrest, J. A., Beatty, J., Hignett, C. T., Pickering, J., and Williams, R. G. P. (1985). A survey of the physical properties of wheatland soils in eastern Australia. CSIRO Aust. Div. Soils Rep. No. 78.

Freebairn, D. M., Littleboy, M., Smith, G. D., and Coughlan, K. J. (1991). Optimising soil surface management in response to climatic risk. In 'Climatic Risk in Crop Production: Models and Management for the Semi-arid Tropics'. (Eds R. C. Muchow and J. A. Bellamy.) pp. 238–305. (CAB International: Wallingford, Oxon.)

French, R. J. (1989). Simple models for use by advisory officers to define crop strategies and to predict farm yields. In 'Need for Climatic and Hydrologic Data in Agriculture in Southeast Asia'. CSIRO Aust. Tech. Mem. 89/5, pp. 185–208.

Gadgil, S., and Joshi, N. V. (1980). Use of principal component analysis in rational classification of climates. In 'Climatic Classification: a Consultant's Meeting'. pp. 17–26. (ICRISAT: Patancheru, A.P.)

Gregory, S., and Cooke, T. (1986). Extreme rainfall deficits—a New South Wales case study. *Aust. Met. Mag.* **34**, 13–25.

Hoy, R. D. (1978). The relationship between annual rainfall and pan evaporation in Australia. *Aust. Met. Mag.* **26**, 45–54.

Hutchinson, M. F. (1989). A new objective method for spatial interpolation of meteorological variables from irregular networks applied to the estimation of monthly mean solar radiation, temperature, precipitation, and windrun. In 'Need for climatic and hydrologic data in agriculture in southeast Asia'. CSIRO Aust. Tech. Mem. 89/5.

Hutchinson, M. F. (1991). The application of thin plate smoothing splines to continent-wide data assimilation. In 'Data Assimilation Systems'. (Ed. D. Jasper.) (BMRC Research Report Series, Bureau of Meteorology: Melbourne.)

Hutchinson, M. F., and Bischof, R. J. (1983). A new method for estimating the spatial distribution of mean seasonal and annual rainfall applied to the Hunter Valley, New South Wales. *Aust. Met. Mag.* **31**, 179–84.

Hutchinson, M. F., Booth, T. H., McMahon, J. P., and Nix, H. A. (1984). Estimating monthly mean values of daily total solar radiation for Australia. *Solar Energy* **32**, 277–90.

Krzanowski, W. J. (1988). 'Principles of Multivariate Analysis. A User's Perspective.' pp. 53–83. (Clarendon Press: Oxford.)

Laslett, G. M., McBratney, A. B., Pahl, P. J., and Hutchinson, M. F. (1987). Comparison of several spatial prediction method for soil pH. *J Soil Sci.* **38**, 325–41.

Linacre, E., and Hobbs, J. (1977). 'The Australian Climatic Environment.' pp. 90–100. (Wiley: Brisbane.)

Mallows, C. L. (1973). Some comments on C_p. *Technometrics* **15**, 661–75.

Marley, J. M., and Littler, J. W. (1989). Winter cereal production on the Darling Downs—an 11 year study of fallowing practices. *Aust. J. Exp. Agric.* **29**, 807–27.

McCaskill, M. R. (1990). TAMSIM—a program for preparing meteorological records for weather-driven models. CSIRO Aust. Div. Tropical Crops and Pastures, Trop. Agron. Tech. Mem. No. 65.

Millington, R. J. (1961). Relations between yield of wheat, soil factors and rainfall. *Aust. J. Agric. Res.* **12**, 397–408.

Myers, R. H. (1990). 'Classical and Modern Regression with Applications.' 2nd Edn. Duxbury Advanced Series in Statistics and Decision Sciences. (PWS Kent Publ. Co.: Boston.)
NATMAP. (1982). 'Agriculture. Atlas of Australian Resources.' Vol. 3. 3rd Edn. (Division of National Mapping: Canberra.)
Nix, H. A. (1987). The Australian climate and its effects on grain yield and quality. In 'Australian field crops. Vol. I. Wheat and other Temperate Cereals'. 2nd Edn. (Eds A. Lazenby and E. M. Matheson.) pp. 183–225. (Angus and Robertson: Sydney.)
O'Neill, R., and Wetherill, G. B. (1971). The present state of multiple comparison methods. *J. Roy. Stat. Soc.* B **36**, 242–4.
Perry, J. N. (1986). Multiple-comparison procedures: a dissenting view. *J. Economic Entomology* **79**, 1149–55.
Richman, M. B. (1986). Rotation of principal components. *J. Climatol.* **6**, 193–335.
Seaman, R. S., and Hutchinson, M. F., (1985). Comparative real data tests of some objective analysis using splines and cross validation. *Aust. Met. Mag.* **33**, 37–46.
Sumner, G. (1988). Precipitation analysis in time. In 'Precipitation Process and Analysis'. pp. 337–95. (Wiley: New York.)
Wahba, G., and Wendelberger, J. (1980). Some new mathematical methods for variational objective analysis using splines and cross-validation. *Monthly Weather Rev.* **108**, 1122–43.
Waring, S. A., Fox, W. E., and Teakle, L. J. H. (1958). Fertility investigations on the black earth wheatlands of the Darling Downs, Queensland. II. Moisture and evapotranspiration in relation to the wheat crop. *Aust. J. Agric. Res.* **9**, 717–29.
Willet, J. B., and Singer, J. B. (1988). Another cautionary note about R^2: its use in weighted least square regression analysis. *Am. Statist.* **42**, 236–8.

Manuscript received 11 May 1992, accepted 13 November 1992