Theor. Appl. Genet. 59, 317-325 (1981)

THEORETICAL AND APPLIED GENETICS

Analysing Environments for Plant Breeding Purposes as Exemplified by Multivariate Analyses of Long Term Wheat Yields

J. Hill and N.A. Goodchild
Biometrics Unit, Institute of Agriculture, University of Western Australia, Nedlands, (W. Australia)

Summary. A previous attempt to relate long term yields from the Western Australian wheat belt to climatic factors proved only partially successful. Here, principal component analysis has been used to examine the patterns of variability created by those socio-economic factors which may have obscured any underlying relationship which existed between yield and climate. In fact, these analyses revealed the existence of variation peculiar to particular groups of years, a result which could explain why many attempts to relate crop yields directly to climatic factors have proved unsuccessful. The plant breeding implications of these genotype × environment interactions are considered.

Key words: Environmental evaluation – Plant breeding – Multivariate analysis – Wheat – Climatic variable

Introduction

In the breeding of new crop varieties a plant breeder has to assess all environmental variables which may affect performance over the target area. Generally these variables are based on seasonal and locational effects, and the way these interact. Clearly, therefore, the breeder must be able to estimate their effects if due cognizance is to be taken of them in a breeding programme.

Official records of wheat yields for the shires of the West Australian wheat belt have been published in their present form for over 50 years. A previous analysis of a 40 year run of these data by Goodchild and Boyd (1975) showed that variation within seasons could be related to shire mean rainfall. Major differences existed between years, however, which could not be related to any climatic variable or combination of these variables. A detailed investigation of this latter source of variation was therefore undertaken on a more extensive set of data.

Methods of Analysis

The basic data for this investigation comprised the mean annual wheat yield (t/ha) of 59 shires in the Western Australian wheat belt (Fig. 1) for the 47 successive planting seasons from 1929 to 1975 inclusive. Yields were based on the amount of grain delivered to the official collection depots, whilst the acreage of wheat planted in a particular shire was obtained from census returns. Concealed within these shire means therefore is an unknown level of variation caused by such factors as soil type and management. Consequently these data cannot be used to study the finer details of wheat yield in Western Australia. They can, however, be used to examine the broad patterns of variability occurring over this period of time.

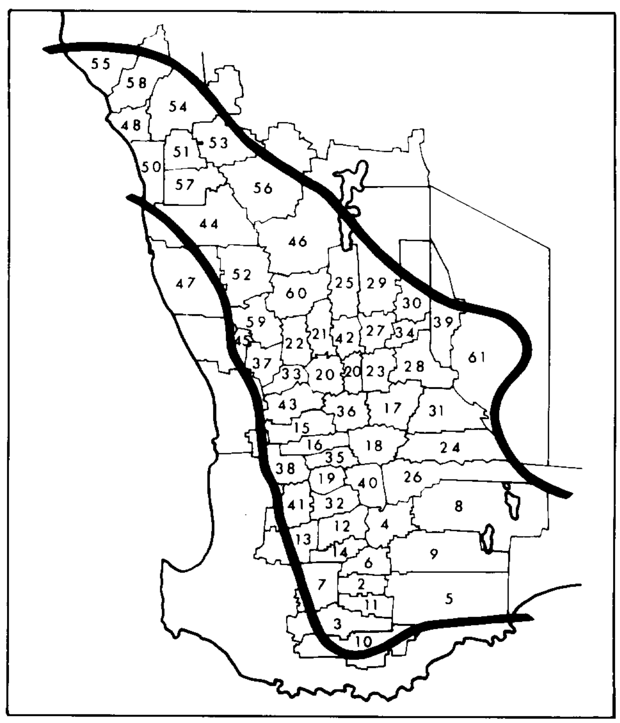

Fig. 1. Boundaries and code numbers of the shires in the Western Australian wheat belt

These data are summarized in Table 1, whilst some idea of the yield changes which occurred within selected shires during this period of time may be gained from Figure 2. Clearly this data set has not been obtained from a conventionally designed experiment. Consequently, conventional univariate methods of analysis would not be expected to tell the whole story as they would not be able, for instance, to display the interrelationships among all 59 shires simultaneously (Goodchild and Boyd 1975). Recourse was therefore made to multivariate techniques in the belief that, by reducing the number of sources of variation which have to be examined, they would reveal the major patterns of variability more clearly.

Broadly speaking, multivariate techniques fall into one of two categories: investigatory or discriminatory. In this particular study we are attempting to investigate variability in wheat yield over a period of time. Moreover, since the units of measurement were the same throughout, it was decided that a principal component analysis, based on covariance matrices, was the most appropriate method for analysing these data. One then has to decide whether shires or seasons are to be the variates. From a plant breeding viewpoint it is important to be able to recognize which shires respond alike to those identifiable factors that are generating variability in wheat yield over years. Logically, therefore, shires should be the variates, and, although this may be considered the more unusual choice, it does have added advantages as we shall see later.

These data can now be viewed in a somewhat different light. Effectively we have a set of 59 measurements of wheat yield (shires), repeated on 47 occasions (years), which from previous work were known to be associated to a greater or lesser degree (Goodchild and Boyd 1975). A principal component analysis transforms the original correlated measurements into uncorrelated linear combinations of these variables. Each combination consists of a set of variate weightings or coefficients known collectively as a principal component. These in turn are the eigenvectors of the covariance matrix of the original data. Successive components, which are usually standardized, are so ordered that they account for a decreasing proportion of the total variability present amongst the original data. Furthermore, the relative contribution of a variate (shire) to a particular component can be determined from its standardized vector coefficient, since these measure the cosine of the angle subtended by that variate with the overall direction of the component (see later).

For the technically minded, the starting point for a principal component analysis is either a correlation matrix, a covariance matrix – or its equivalent – a dispersion matrix. Here a dispersion matrix was used. Simple descriptions of this technique have been given by Pearce (1969) and Paterson et al. (1978), but briefly it requires the extraction of the eigenroots and vectors from whichever matrix was used. In the present data set each eigenvector will comprise 59 elements, each element representing the standardized coefficient or 'loading' of a particular variate (shire) for that vector. Those shires most closely aligned with the direction of, and hence contributing most to the variation displayed by a particular vector, will have coefficients (cosines) approaching ± unity. Conversely, those shires which are independent of the vector concerned will have coefficients close to zero.

Eigenvectors may be conveniently classified as either a size or shape vector (Mosimann 1970). On the one hand the coefficients of a size vector all have the same sign. Consequently the variates of such a vector will tend to react alike, though not necessarily to the same extent, to those factors responsible for generating the variability represented by that vector. Shape vectors, on the other hand, contain both positive and negative coefficients, indicating that the variates are now behaving in quite different ways.

Before those factors affecting that aspect of wheat yield inclu-

Table 1. Summary of yield data of the 59 shires (t/ha)

Shire	Mean	Minimum	Maximum	S.D.
2	1.01	0.47	2.08	0.335
3	0.86	0.24	1.78	0.338
4	1.02	0.59	1.65	0.237
5	1.04	0.50	1.51	0.246
6	0.93	0.40	2.04	0.322
7	0.86	0.34	1.81	0.378
8	0.82	0.46	1.20	0.168
9	0.81	0.50	1.42	0.209
10	0.76	0.21	1.67	0.321
11	0.97	0.47	2.02	0.333
12	0.99	0.48	1.88	0.319
13	0.83	0.34	1.65	0.342
14	0.91	0.40	2.16	0.367
15	1.03	0.46	2.08	0.310
16	0.96	0.44	2.11	0.331
17	1.00	0.48	1.64	0.235
18	0.92	0.49	1.78	0.258
19	0.87	0.35	2.12	0.376
20	1.02	0.60	2.02	0.307
21	0.98	0.57	1.93	0.254
22	1.03	0.57	2.06	0.294
23	0.88	0.38	1.52	0.221
24	0.84	0.31	1.33	0.219
25	0.85	0.23	1.37	0.268
26	0.84	0.31	1.29	0.178
27	0.89	0.26	1.42	0.263
28	0.81	0.14	1.20	0.242
29	0.78	0.15	1.33	0.294
30	0.73	0.09	1.43	0.335
31	0.92	0.36	1.47	0.216
32	0.94	0.40	2.01	0.348
33	1.04	0.52	1.97	0.308
34	0.82	0.09	1.44	0.328
35	0.86	0.43	1.96	0.322
36	1.02	0.53	1.83	0.252
37	1.08	0.55	1.83	0.299
38	0.98	0.48	2.13	0.376
39	0.73	0.05	1.40	0.345
40	0.92	0.44	1.99	0.305
41	0.94	0.45	1.67	0.308
42	0.99	0.46	1.66	0.238
43	0.99	0.50	1.84	0.273
44	1.01	0.44	1.59	0.243
45	0.93	0.44	1.72	0.315
46	0.91	0.38	1.37	0.234
47	0.69	0.30	1.18	0.203
48	0.84	0.30	1.26	0.218
50	0.91	0.48	1.49	0.221
51	0.92	0.38	1.39	0.268
52	1.01	0.57	1.60	0.237
53	0.88	0.35	1.39	0.248
54	0.81	0.15	1.24	0.245
55	0.87	0.30	1.32	0.217
56	0.87	0.32	1.27	0.221
57	1.00	0.50	1.62	0.223
58	0.85	0.28	1.32	0.206
59	1.00	0.50	1.99	0.310
60	0.98	0.62	1.78	0.249
61	0.69	0.03	1.44	0.340

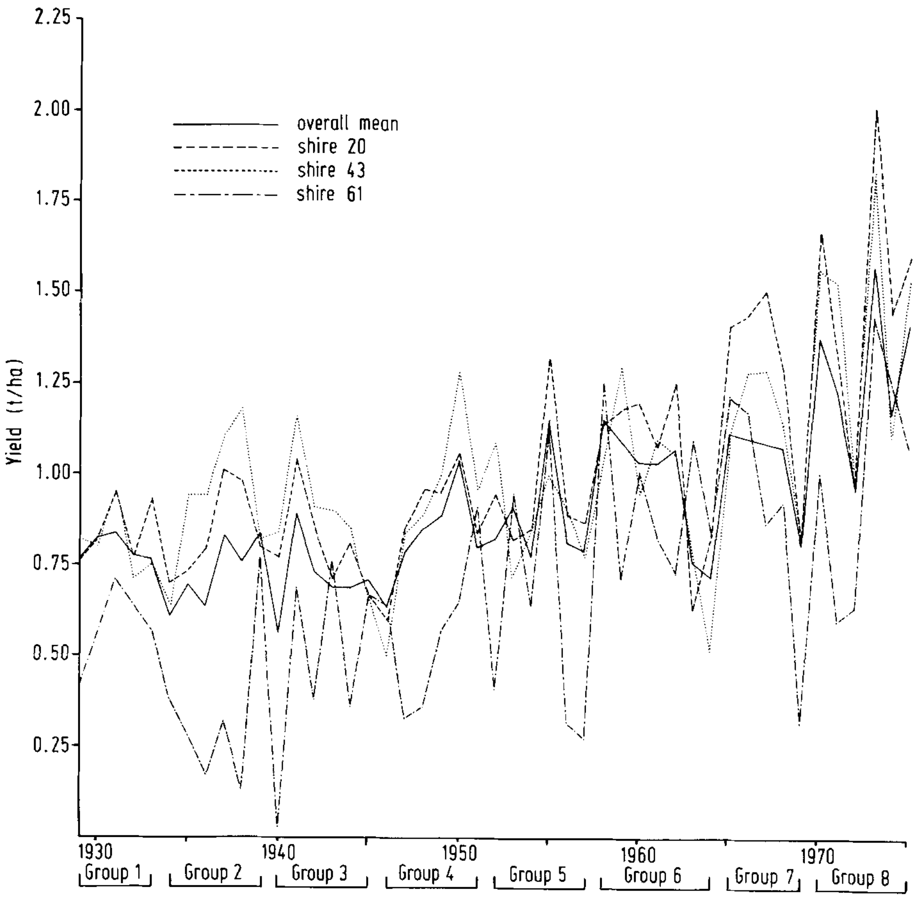

Fig. 2. Mean annual yields in the Western Australian wheat belt 1929-1975 inclusive

ded in the previously defined between season component can be examined in detail, the overall pattern of variability must be dissected into its component parts. This will require subdivision of the data set into smaller groups of years. Initially discriminatory multivariate techniques, such as various cluster analysis models, were applied. These proved unsatisfactory, however, because the resulting clusters were both inconsistent and unrelated to the major cultural, agronomic and environmental factors known to be operating at the time. Accordingly, a chronological grouping of the data, based on information made available by a number of authorities with extensive local knowledge, was decided upon. This provided eight groups of years (Table 2). Subsequently, a principal component analysis was conducted within each of these eight groups in turn. It must be emphasized that the vectors extracted from individual group matrices will be subject to greater errors than those computed from the overall data set if only because each group is based on fewer years.

Despite the arbitrary nature of this grouping, it is nevertheless of interest to assess the relative importance of the between groups and within groups between years components of variation at different times over the 47 year period. Two distinct, yet complementary, methods are available for this assessment, both of which rest upon the obvious fact that the complete data contains both components, whereas the between group component cannot contribute to variation within the individual groups.

The first method seeks to remove from the overall dispersion matrix that portion of the variation attributable to a vector of a particular group of years. This is accomplished by forming a covariance matrix from the coefficients of the vector concerned, as described by Holland (1969). This matrix is then subtracted from the overall dispersion matrix to give a reduced matrix. Subsequently, the first two eigenvectors were extracted from this reduced matrix and compared with their counterparts in the original dispersion matrix. If the overall and individual group vectors are independent of each other, then the former will be unaffected by the reduction process and will reappear in the reduced matrix. Hence the patterns of variability depicted by the vectors of the overall and reduced matrices will coincide. By contrast, a significant alteration in the pattern of variability displayed by the vector of a reduced matrix indicates that the particular group vector describes variation shared in common with the overall vector. Consequently the direction of the group and overall vectors will now be closely aligned. The second method measures this alignment directly by determining the cosine of the angle which a particular group vector subtends with each of the two overall vectors in turn. Values close to $\pm$ 1 indicate alignment. As the cosine approaches zero, however, so the vectors diverge and become progressively independent of each other.

Evaluating the results from a principal component analysis can pose problems, since the axes along which the variates are measured after transformation do not always lend themselves readily to biological interpretation. However, the choice of shires as variates enables the patterns of variability displayed by the various vectors to be mapped for ease of presentation. Accordingly, for both the overall and individual group analyses, shires have been ranked in descending order of magnitude of the vector coefficients and then grouped into 3 sets of 15 and one of 14, the latter being those shires with the lowest coefficients.

Analytical Results and their Geographic Interpretation

Variation within Shires

We may conveniently begin the analysis of these data by examining briefly the level of within shire variability. From Table 1 we note that the standard deviation of shire means varies between 0.168 (shire 8) and 0.378 (shire 7), indicating a five-fold range in the within shire variability. Interestingly, geographic trends are discernible among them, with two areas of relatively low variability occurring in the south-east and north-west, whilst two areas of relatively high variability exist in the west and east. Central shires are intermediate. Not all shires can be so grouped, but such trends may nevertheless provide an indication of those patterns of variability that may emerge from the principal component analyses, when a more efficient partitioning of the overall variation should be forthcoming.

Principal Component Analyses

The results of the principal component analysis for the complete data and each group of years are summarized in Table 2, from which it is apparent that the first two vectors account for at least 75% of the total variation in any particular set of years. Because each of the remaining vectors accounts for but a small percentage of this variation they have been regarded as reflecting random variation.

Since the first principal component overall is a size vector, it may be concluded that shires are responding in a similar way to those factors chiefly responsible for variation in wheat yield over this particular period of time. Vector 2, on the other hand, is a shape vector which exhibits a pattern of variability characteristic of the within season variation in rainfall distribution (Boyd et al. 1976). Greatest variability is found among the more marginal eastern and western shires of the wheat belt, that is those shires with the highest negative and positive coefficients respectively, whereas central shires, which have coefficients close to zero, are the least variable (Fig. 3).

Table 2. Chronology of groups and percentage variation accounted for by the principal vectors for each group and the overall data

Group	Years	Vector 1	Vector 2
1	1929-33	61 (shape)	18 (shape)
2	1934-39	51 (shape)	35 (shape)
3	1940-45	57 (essentially size)	27 (shape)
4	1946-51	59 (essentially size)	18 (shape)
5	1952-57	60 (essentially size)	17 (shape)
6	1958-64	68 (essentially size)	14 (shape)
7	1965-69	71 (essentially size)	20 (shape)
8	1970-75	76 (size)	12 (shape)
Overall	1929-75	67 (size)	13 (shape)

As a first step in the more detailed examination of vector 1, the annual mean wheat yields of each shire were correlated with the vector scores. Those shires with the highest coefficients, and which are therefore most closely aligned with the overall direction of the first vector, generally have the strongest associations (Table 3).

From the individual group analyses it is evident that, while vector 1 is not invariably a size vector, it becomes progressively more so with time (Table 2). Shires are, therefore, responding increasingly in unison to those factors affecting seasonal variation in wheat yield within a particular group of years. The pattern of variability portrayed by the first vector of each of these 8 groups is pre-

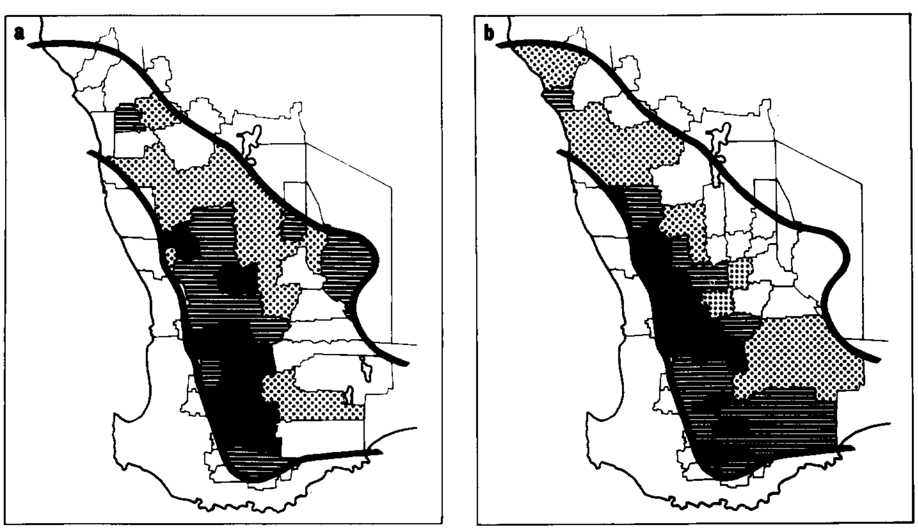

Fig. 3a and b. Mapping of overall vectors; a Vector 1, b Vector 2. Solid, first 15 shires; lines, second 15 shires; dots, third 15 shires; blank, last 14 shires

Table 3. Shire coefficient for vector 1 overall (*i*) together with the correlation of shire mean annual yield with vector scores for vector 1 of the complete data (*ii*)

Shire	(*i*)	(*ii*)	Shire	(*i*)	(*ii*)
2	0.1541	0.8115	32	0.1829	0.9277
3	0.1353	0.6995	33	0.1526	0.8793
4	0.1115	0.8439	34	0.1091	0.6026
5	0.0586	0.4315	35	0.1622	0.8851
6	0.1666	0.9143	36	0.1238	0.8843
7	0.1907	0.8865	37	0.1377	0.8142
8	0.0699	0.7571	38	0.1967	0.9225
9	0.1022	0.8688	39	0.1128	0.5957
10	0.1150	0.6295	40	0.1601	0.9295
11	0.1686	0.8962	41	0.1527	0.8715
12	0.1617	0.8979	42	0.1028	0.7844
13	0.1719	0.8850	43	0.1311	0.8529
14	0.1923	0.9228	44	0.1133	0.8482
15	0.1518	0.8691	45	0.1139	0.6361
16	0.1717	0.9171	46	0.1084	0.8392
17	0.1057	0.8198	47	0.0407	0.3499
18	0.1345	0.9291	48	0.1009	0.8294
19	0.1956	0.9145	50	0.0793	0.6468
20	0.1627	0.9461	51	0.1174	0.7974
21	0.1321	0.9330	52	0.1142	0.8637
22	0.1520	0.9197	53	0.1045	0.7711
23	0.1047	0.8624	54	0.0900	0.6768
24	0.0992	0.8217	55	0.0951	0.7900
25	0.1061	0.7195	56	0.0994	0.8175
26	0.0803	0.8173	57	0.0851	0.7038
27	0.1107	0.7626	58	0.0890	0.7877
28	0.0809	0.6159	59	0.1540	0.8778
29	0.1111	0.6855	60	0.1331	0.9565
30	0.1169	0.6326	61	0.1290	0.6846
31	0.0826	0.7038			

sented in Figs. 4 and 5. As an aid to interpretation some background information is provided on the major factors operating at the time. Vector 2, which is a shape vector throughout, will be discussed only if it is interpretable.

Period 1929-33

For this period vector 1 is a shape vector, with the most variable shires occurring on the eastern and western margins of the wheat belt. Central shires are, by contrast, relatively invariate. This pattern typifies the east-west variability in seasonal rainfall distribution displayed by vector 2 overall. A similar intepretation has therefore been placed upon this group vector. Thus wet years would depress yields in western shires because of soil waterlogging, while eastern shires would experience good growing conditions. Conversely, dry years would favour western shires since eastern shires would suffer from drought.

Table 4. Total rainfall (mm) between April and October and annual mean yield (t/ha) for shire 61 during year group 2

Year	Rainfall	Yield
1934	289	0.38
1935	211	0.28
1936	124	0.17
1937	147	0.32
1938	110	0.13
1939	249	0.77
Mean 1929-1975	208	0.69

Period 1934-39

Again vector 1 depicts seasonal rainfall distribution, though the variance ascribable to this component is less than for 1929-33. This could well be due to the prolonged drought experienced by eastern shires during this period (Table 4).

Period 1940-45

Prior to the second war the individual group vectors suggest that climatic factors are primarily responsible for variability in wheat yield. The war years, however, mark a turning point because from now on vector 1 becomes progressively more of a size vector. Within this period vector 1 reflects a westerly decrease in variability arising from the loss of export markets, enlistment of farmers in the armed forces and the widespread drought of 1940, all of which led to a drastic reduction in wheat acreage, particularly in the marginal eastern shires. Variation associated with seasonal rainfall distribution has now been relegated to vector 2.

Period 1946-51

Vector 1 is again essentially a size vector, with western shires being the most variable. Although the imposition of credit restrictions and the consequent unavailability of equipment limited farming activity throughout the wheat belt during this period, nowhere were these effects more apparent than in the eastern shires.

Period 1952-57

For this particular group vector 1 indicates that eastern shires are the most variable, whilst southern and western shires are relatively invariate. At this time new, drought-resistant varieties of subterranean clover were being introduced into the drier eastern shires. This, combined with

Fig. 4a-d. Vector 1 for the first four groups of years; a 1929-33, b 1934-39, c 1940-45, d 1946-51. Key as for Fig. 3

better management and the easing of credit restrictions increased yields in the east in the more favourable years. In southern and western shires, however, these farming systems were already standard practice.

Period 1958-64

Here vector 1 indicates an easterly decrease in variability. Towards the end of the period the entire wheat belt was affected by a rust epidemic. Although yields generally suffered as a result, the most drastic reductions undoubtedly occurred in the wetter western shires. Moreover, this pattern would be reinforced by the gradual extension eastwards of power farming techniques.

Period 1965-69

Over this period vector 1 reveals a corridor of high variability extending eastwards from the central shires. This delimits the area most affected by the severe drought of 1969. By contrast shires in the north and south-west, which according to Fitzpatrick (1970) escaped the worst effects of the drought, are shown by vector 1 to be the least variable. Furthermore, by the end of this period the recently released, rust-resistant variety 'Gamenya' so dominated the wheat belt that it accounted for approximately 60% of the acreage planted.

Period 1970-75

By now vector 1 has become a size vector, reflecting in part the continued widespread use of 'Gamenya'. Central shires prove to be the most variable, due mainly to the record harvest of 1973, when the yield from several of these shires exceeded two t/ha. Vector 2 meanwhile reverts to a pattern characteristic of the seasonal rainfall distribution.

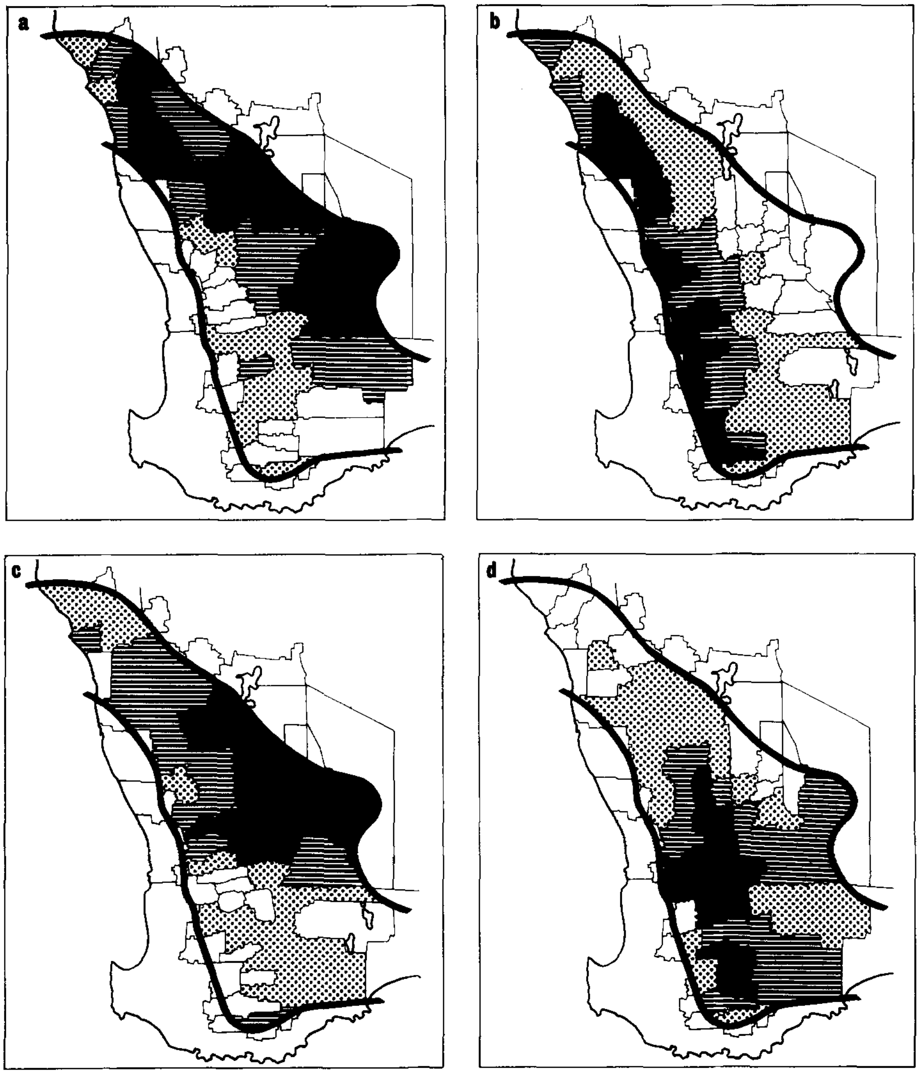

Fig. 5a-d. Vector 1 for the last four groups of years; a 1952-57, b 1958-64, c 1965-69, d 1970-75. Key as for Fig. 3

Analyses of Reduced Matrices

Considering the overall matrix reduced by the group vectors, all the first vectors derived from these reduced matrices are size vectors. The removal of vector 1 of year groups 1, 2, 3 and 7 leaves a pattern of variability which is substantially the same as that displayed by vector 1 overall. This contrasts markedly with the picture which emerges when vector 1 of groups 4, 5, 6 and 8 is removed. Each of these groups produces a pattern of variability in the resultant reduced matrix which differs to a varying extent from that portrayed by the first vector of the original dispersion matrix. Amongst these latter four groups, the removal of vector 1 from group 5 has only a minor effect, while the removal of vector 1 from group 8 completely disrupts the original pattern. The results of these group analyses are confirmed by the direct measurement of divergence, namely the cosine of the angle subtended by appropriate pairs of vectors (Table 5). Again vector 1 of group 8, which subtends an angle of only 21°24′ (i.e. cosine of 0.931), stands out as being closely aligned with vector 1 overall. Clearly this particular group vector is making a major contribution to the overall vector, due no doubt to the higher yields and increased seasonal variability which occurred between 1970 and 1975. Two further examples will suffice to illustrate the utility of this particular approach. It will be recalled that vector 1 of groups 1 and 2 were both shape vectors which exhibited patterns of variability similar to the seasonal rainfall distribution and hence not unlike vector 2 overall. Inspection of Table 5 confirms that these group vectors are indeed quite closely aligned with vector 2 overall, having cosines of −0.789 and −0.911 respectively.

When the variability due to the second principal component of the individual groups is removed from the overall matrix, vector 1 overall remains substantially the same.

Table 5. Cosine of the angle subtended by the individual group vectors with the first two vectors overall

Group	Vector	Vector 1 overall	Vector 2 overall
1	1	0.173	−0.789
	2	−0.323	0.111
2	1	0.222	−0.911
	2	0.768	0.280
3	1	0.649	−0.586
	2	0.393	0.618
4	1	0.896	0.295
	2	0.079	0.602
5	1	0.721	−0.592
	2	0.190	0.488
6	1	0.912	0.246
	2	0.098	−0.795
7	1	0.600	−0.652
	2	−0.060	0.169
8	1	0.931	0.090
	2	0.020	0.841

However, the removal of group 2 produces a size vector which differs to some extent from the overall first vector. This suggests that a within group size component exists which is different from the between group size component.

Dealing briefly with the second principal components resulting from the reduced matrices, these are all shape vectors, and, generally of a shape different from vector 2 overall. None of them are readily interpretable.

Discussion

Because of the unusual nature of these data they were not amenable to normal univariate methods of analysis. Indeed, even when multivariate techniques were applied, it was only after shires, as opposed to years, were used as variates that a meaningful picture of seasonal variability in wheat yield could be assembled. It is this picture and its implications for wheat breeding in Western Australia that is the basis of the ensuing discussion.

The results confirm that the total variability of wheat yields can be partitioned into at least three main sources, namely between and within groups of years and within years. The latter has been discussed and interpreted elsewhere (Goodchild and Boyd 1975; Boyd et al. 1976). Although the between and within groups of years variation must be artificial to some extent, since they will depend upon the way in which the overall data has been subdivided, it is evident both from the reduced matrices and the cosine measurements that the relative contributions of these two components fluctuate markedly over the period covered by these data. On the one hand, vector 1 of groups 1, 2, 3 and 7 exhibit varying degrees of divergence from vector 1 overall. Hence, within these particular periods of time vector 1 of the overall analysis represents mainly between rather than within group variation. Groups 4, 6 and 8, on the other hand, present a totally different picture, because the first vector of each of these groups is closely aligned with the first vector overall. Accordingly, this latter vector must now contain variability attributable to the within group component.

Besides emphasizing the importance of the various seasonal components of variation, these analyses disclose that sources of variation exist whose effects upon wheat yield are peculiar to a particular group of years. Furthermore, it is evident from the individual group analyses that different parts of the wheat belt, especially the more marginal eastern and western shires, respond differently to these transitory factors, to give, in the broadest meaning of the phrase, a genotype-environment interaction. Many of these factors are wholly outside the farmers' and plant breeders' control, such as those which occurred during the war years. Thus, to the more conventional between and within seasons components of environmental variability must be added a between groups of seasons component. Whilst all three components will impede the progress of a plant breeding programme, this additional component is likely to prove the most difficult to quantify.

The picture conveyed by the overall analysis of these data suggests a relatively uniform response by the shires to those socio-economic factors primarily responsible for generating variability in wheat yield, with climatic factors playing a secondary role. However, this probably reflects the weighting given to the later years of this period when yields were both higher and more variable, a suggestion which is confirmed by the similarity of vector 1 overall with vector 1 for group 8 (see Figs. 3a and 5d). The true picture, as revealed by the individual group analyses, is one of increasing uniformity with climatic factors being the prime cause of variability prior to the second world war and then being relegated to a secondary role as other factors assumed increasing importance. In practice, wheat growing in Western Australia before the second world war was very much at the mercy of the elements, particularly at sowing time which was then a protracted operation. Obviously wheat growing still depends on the weather, but this dependence has been lessened to some extent by increased mechanization and the advent of new varieties. If to these factors is added the economic pressures that have forced small farms to amalgamate, then the reasons for this increased uniformity are not hard to find. Moreover, the pace of this trend has accelerated recently with the planting of large tracts of the wheat belt to the variety

'Gamenya'. This trend is mirrored faithfully by vector 1 of the individual group analyses, which has become progressively more of a size vector over this period (Table 2). Thus, since the second world war each group of years has been characterized by its own series of events which have overridden the effects of climate. This may explain why many of those models based on climatic factors have failed to provide a satisfactory explanation of the variation in crop yields with time. A long run of seasons is likely to include unquantifiable historical events whose effects upon yield, though possibly large, are nevertheless statistically intractable. Unless the plant breeder is satisfied either that such unique events will not distort his data, or that their effects can be removed, he would be ill-advised to incorporate extensive sets of historical data in his predictive models. However, Fisher (1924) and Cornish (1950) were able to relate long term wheat yield to rainfall. In part, Fisher's success can be attributed to the use of data derived from experiments with a known history of continuous treatments and cropping. The monoculture system and relatively consistent climatic and cultural patterns in South Australia may account for Cornish's results.

The analyses presented here highlight the problems associated with interpreting trends in long term commercial wheat yields, where the material is not subject to proper experimental control. Climatic variation between seasons will often be confounded both with differences in management and changes in variety. Similar considerations will also enter into any discussion of the way in which shires differentially respond to those factors that affect wheat yield within a particular growing season. Nevertheless these analyses have enabled the effects of these interacting factors to be expressed in a manner which can be used as a basis for the development of a coherent plant breeding programme.

Clearly the central shires are the least variable, possibly because they are less liable to climatic extremes. Consequently, they have been influenced less by socio-economic pressures. Furthermore, these shires are able to incorporate new management techniques and adopt new varieties more readily. Within this area it would seem that the plant breeder can concentrate on improving yield per se. By contrast, in the more variable eastern and western shires yields are unreliable because they can be limited by factors such as drought. Here the plant breeder will need to scrutinize the possibilities of introducing tolerance or resistance to these factors from a wider gene pool.

Acknowledgement

The authors wish to express their appreciation to Mr. R.G. Martin, Chairman of the Wheat Quota Committee of Western Australia, and Dr. J. Millington for the comments and information freely given. We are grateful to Mrs. Christine Robertson for preparing some computer programs and for carrying out the many computations, and to Dr. E.L. Breese for reading and commenting upon the manuscript. Thanks are due to the Agricultural Research Council of the United Kingdom and the State Wheat Committee of Western Australia for financial support and to the Director of the Welsh Plant Breeding Station for granting leave of absence to one of us (J.H.).

Literature

Boyd, W.J.R.; Goodchild, N.A.; Waterhouse, W.K.; Singh, B.B. (1976): An analysis of climatic environments for plant-breeding purposes. Austr. J. Agric. Res. 27, 19-33

Cornish, E.A. (1950): The influence of rainfall on yield of wheat in South Australia. Austr. J. Sci. Res. Ser. B 3, 178-218

Fisher, R.A. (1924): The influence of rainfall on the yield of wheat at Rothamsted. Philos. Trans. Roy. Soc. (Lond.) Ser. B 213, 89-142

Fitzpatrick, E.A. (1970): The expectancy of deficient winter rainfall and the potential for severe drought in the south west of Western Australia. Misc. Publ. Agron. Dept. Univ. West. Austr. No. 70/1

Goodchild, N.A.; Boyd, W.J.R. (1975): Regional and temporal variations in wheat yield in Western Australia and their implications in plant breeding. Austr. J. Agric. Res. 26, 209-217

Holland, D.A. (1969): Component analysis: an aid to the interpretation of data. Expl. Agric. 5, 151-164

Mosimann, J.E. (1970): Size allometry: size and shape variables with characterizations of the log-normal and generalized gamma distributions. J. Amer. Statist. Ass. 65, 930-945

Paterson, J.G.; Goodchild, N.A.; Boyd, W.J.R. (1978): Classifying environments for sampling purposes using a principal component analysis of climatic data. Agric. Meteorol. 19, 349-362

Pearce, S.C. (1969): Multivariate techniques of use in biological research. Expl. Agric. 5, 67-77

Received October 15, 1980
Communicated by R. Riley

Dr. J. Hill
Welsh Plant Breeding Station
Aberystwyth (Wales)

Dr. N.A. Goodchild
Biometrics Unit
Institute of Agriculture
University of Western Australia
Nedlands (W. Australia 6009)