
C S I R O P U B L I S H I N G

Australian Journal of Agricultural Research

Volume 51, 2000
© CSIRO 2000



A journal for the publication of original contributions
towards the understanding of an agricultural system

www.publish.csiro.au/journals/ajar

All enquiries and manuscripts should be directed to

Australian Journal of Agricultural Research

CSIRO PUBLISHING

PO Box 1139 (150 Oxford St)

Collingwood

Vic. 3066

Australia

Telephone: 61 3 9662 7628

Facsimile: 61 3 9662 7611

Email: jenny.fegent@publish.csiro.au



Published by **CSIRO PUBLISHING**
for CSIRO and
the Australian Academy of Science



Analysis of line \times environment interactions for yield in navy beans. 2. Pattern analysis of lines and environment within years

R. J. Redden^A, I. H. DeLacy^B, D. G. Butler^C, and T. Usher^A

^A Queensland Department of Primary Industries, Hermitage Research Station, Warwick, Qld 4370, Australia.

^B School of Land and Food Sciences, The University of Queensland, St Lucia, Qld 4072, Australia.

^C Biometry Branch, Queensland Department of Primary Industries, Tor Street, Toowoomba, Qld 4350, Australia.

Abstract. Seven years of multi-environment yield trials of navy bean (*Phaseolus vulgaris* L.) grown in Queensland were examined. As is common with plant breeding evaluation trials, test entries and locations varied between years. Grain yield data were analysed for each year using cluster and ordination analyses (pattern analyses). These methods facilitate descriptions of genotype performance across environments and the discrimination among genotypes provided by the environments. The observed trends for genotypic yield performance across environments were partly consistent with agronomic and disease reactions at specific environments and also partly explainable by breeding and selection history. In some cases, similarities in discrimination among environments were related to geographic proximity, in others management practices, and in others similarities occurred between geographically widely separated environments which differed in management practices. One location was identified as having atypical line discrimination. The analysis indicated that the number of test locations was below requirements for adequate representation of line \times environment interaction. The pattern analyses methods used were an effective aid in describing the patterns in data for each year and illustrated the variations in adaptive patterns from year to year. The study has implications for assessing the number and location of test sites for plant breeding multi-environment trials, and for the understanding of genetic traits contributing to line \times environment interactions.

Additional keywords: *Phaseolus vulgaris*, yield, cluster analysis, ordination analysis, environments.

Introduction

Pattern analysis (Williams 1976; DeLacy *et al.* 1996a) is the joint use of clustering and ordination analyses to investigate the information in data. All of the genotype \times environment information from plant improvement multi-environment trials (METs) grown over diverse environments can be investigated with clustering and ordination procedures (DeLacy *et al.* 1996a, 1996b). Clustering summarises data by grouping genotypes with similar performance across environments and grouping those environments that produce similar discrimination among the performance of genotypes. Thus, clustering summarises complexity in the data with retention of the majority of the information by describing performance with relatively few genotype groups or relatively few environmental groups or both (Mungomery *et al.* 1974; Cooper and DeLacy 1994; DeLacy *et al.* 1996a). With ordination, relationships among genotype performance and environment discrimination are represented in a low dimensional representation of the original data in fewer axes (Gabriel 1971;

Gauch and Zobel 1996). In these representations genotypes with similar performances and environments that produce similar discrimination among genotypes will be placed close together on the biplot.

The theoretical basis of complementary cluster and ordination analyses has been described for genotype by environment data sets both within and over years (Basford *et al.* 1991; Cooper and DeLacy 1994). These approaches are applicable to regional variety trials, where a range of crop varieties grown over a range of geographic environments are assessed for grain yield or other attributes (Byth *et al.* 1976; DeLacy *et al.* 1996b).

The entries and environments of the regional navy bean variety trials, with changes in each factor over years, have been described in a previous paper (Butler *et al.* 2000, tables 1 and 2). This paper describes how analysis of data might affect a plant breeding strategy for choice of entries and locations. The previous report on the regional variety trials analysed sources of variation across years, environ-

ments, and entries (Butler *et al.* 2000). This paper reports on pattern analyses within years, with key questions being (1) the level of either redundancy or inadequacy of the numbers of test locations, and (2) evaluation of any differences in adaptation to target environments between traditional and newly imported lines, and of whether these differences were reflected in new breeding lines derived from both genetic groups. Critically this affects choice of test environments and of appropriate levels of genetic diversity for the breeding program.

Methods

Navy bean METs were grown in Queensland for the 7 years 1983–1989. Descriptions of the 21 lines grown and the 16 locations used for these METs and the management and climatic descriptions for each year were given in a previous paper (Butler *et al.* 2000). The study of grain yield data for these trials utilised relevant agronomic and disease data (Table 1). The traditional lines from an earlier breeding program of limited genetic variance that were used in these trials (Redden *et al.* 1985) included Gallaroy, Kerman, Actolac, Revenue, Actosan, Selection 46, W1401 2GA, and W1885. All other lines used were variously imported from the USA (Banker, Campbell 11, Campbell 16, Campbell 15) or Colombia (NEP2, BAC134, BAC125), or were derived (CH series) breeding lines. The number of lines grown at all the locations used in each year varied from 8 to 14. Lines and environments changed among years, and trial entries grown in only 1 year, and locations used only once, were deleted (Butler *et al.* 2000) from the analy-

sis. The trial period spanned the initiation of a plant breeding program (Redden *et al.* 1985) through to the evaluation of the first cycle of derived selections.

Analytic procedures

Following Fox and Rosielle (1982), all analyses both for lines and environments were performed on environment standardised mean grain yield at each environment. Cooper and DeLacy (1994) and Cooper *et al.* (1996) showed that investigation of the phenotypic correlations amongst environments was important to understanding the discrimination among genotypes grown in them. Such an investigation is provided by pattern analysis of environment standardised data. For comparison purposes a pattern analysis was carried out on location centred data (DeLacy *et al.* 1996a).

Classification of lines and environments was performed using an agglomerative hierarchical clustering procedure (Williams 1976) with squared Euclidean distance (SED) as the dissimilarity measure and incremental sum of squares (ISS) or Ward's method as the clustering strategy (Ward 1963; Burr 1968, 1970; Wishart 1969). ISS is recommended as the preferred hierarchical clustering strategy for analysis of yield data from plant breeding METs (Byth *et al.* 1976; DeLacy 1996a, 1996b). The effectiveness of the classifications was investigated by examining the partitioning of the sums of squares among sources. The line classifications were truncated for summary purposes when 80% of the genotype \times environment (G \times E) was retained in the reduced matrix (Byth *et al.* 1976; Cooper and DeLacy 1994). Classifications were investigated using dendrograms and line group performance plots. Performance plots were constructed by plotting means of standardised yields for line groups against an environment index of unstandardised mean yields of environments.

Table 1. Mean deviations from trial means, averaged over environments and seasons, for agronomic and disease scores for the lines entered in 7 navy bean multi-environment yield trials grown in seasons 1983 to 1989

Positive deviations indicate above average levels: later flowering, taller canopy height, higher pods (1–9 index), increased lodging, more vining (1–4 scale of increased vining), greater grain weight, higher disease scores (rust, common bacterial blight and peanut mottle virus (PMV) on a 1–9 scale of increasing severity), and greater sensitivity to zinc deficiency (1–9 scale of increasing severity)

Line name	Abbrev.	Agronomic traits						Disease traits			
		Days to flower	Canopy height (cm)	Pod height	Lodging (%)	Vining	100-grain wt (g)	Rust	Bacterial blight	PMV	Zinc deficiency score
Gallaroy	Gy	-4.5	-6.8	-0.9	-2.7	-1.0	-0.6	3.1	0.7	-0.4	3.3
Kerman	Kerm	-1.6	-1.8	-0.4	3.4	-0.4	-1.3	1.9	0.4	0.7	0.6
Actolac	Actl	-3.1	-3.0	-0.3	-7.0	-0.7	1.0	0.1	0.7	-0.1	1.3
Banker	Bkr	-0.7	-1.3	-0.6	8.6	-0.1	-0.4	-0.6	-0.2	0.7	-1.0
Campbell 11	C11	0.9	4.7	1.0	-9.6	0.3	0	-0.6	0.5	0.3	-0.4
Revenue	Rev	-0.5	1.5	-0.5	16.0	0.6	-0.3	0.8	-0.4	0.6	-1.3
BAC125	B125	-3.5	-6.5	-0.9	-9.7	-0.1	1.5	0.3	-0.1	0.3	1.8
BAC134	B134	-1.5	-0.2	-0.9	8.2	2.2	4.6	0.7	0.1	0.1	
NEP2	Nep2	-0.3	1.5	0.8	3.0	-0.8	-0.8	0.7	0	1.1	
CH14-8D	8D	1.8	2.6	0.2	-6.4	-0.3	-0.3	0	-0.4	-0.9	
CH14-11D	11D	4.0	4.5	3.0	1.2	-0.3	-0.6	0.7	-0.7	-0.3	0.3
CH14-27D	27D	2.7	3.2	0.5	-1.9	-0.6	0.4	0.7	-0.5	-0.3	
CH14-28D	28D	1.2	5.9	1.2	0.1	0.4	1.2	-0.4	-0.3	-1.2	-0.1
CH9-4D	4D	-0.1	-0.8	-0.7	-1.3	0	-0.6	-0.3	0.1	-0.9	0.4
2GA	2GA	1.2	0.3	0.2	-1.7		0.6	-0.4	0.1	1.1	-1.0
W1401	W1401	-0.6	-2.8	-0.6	3.2	-0.3	-0.6	-0.3	0.1	1.3	0.3
Campbell 16	C16	1.4	4.5	1.4	-10.0		1.0	-0.5	1.0	1.0	-1.0
Actosan	Actn	-2.2	-5.7	-0.7	7.3		-0.3	-0.5	0.1	1.3	-0.9
Selection 46	S46	0.2	0.3	-0.6	1.6		0.7	-0.4	-0.6	0.6	-0.8
W1885	W1885	-1.5	0	0.3		0.7		0.4	0.4		
Campbell 15	C15	-10.6	-2.5	0.3		-1.0		0.3	0.6		

Table 2. Array size, proportion of sums of squares (SS) due to sources environment (E), lines (L), and L \times E interaction, truncation level, reduction in array size, proportion of SS retained in reduced array, and proportion of L \times E SS due to joint linear regression (JLR)

Season	Array size L \times E	Percentage of total SS due to sources:			Truncation level lines	Reduction in array size (%)	Percentage of source SS retained in reduced array		L \times E SS due to JLR (%)
		E	L	L \times E			L	L \times E	
1983	13 \times 4	79	6	15	4	69	74	86	33
1984	11 \times 6	95	1	4	5	55	64	80	27
1985	10 \times 4	90	4	6	5	50	96	86	38
1986	8 \times 5	89	3	8	3	63	73	88	78
1987	14 \times 10	84	3	13	5	64	73	80	26
1988	10 \times 8	89	3	8	6	40	89	80	17
1989	9 \times 8	70	12	18	5	44	83	84	32

Ordination was done on the environment standardised mean yield data using the singular value decomposition (SVD) algorithm (Eckhard and Young 1936) with the results represented by a bi-plot (Gabriel 1971; Bradu and Gabriel 1978; Gabriel 1981). The algorithm results in 2 sets of plotting points for each of the derived vectors, one set for the genotypes (lines or entries) and one set for the environments. The lines were plotted as points and the environments as direction vectors. Since the data have been environment standardised the total variance for each environment is 1 and the length of the vectors indicates how well the environments are modelled in the biplot. If an environment was modelled perfectly the length of its vector would be 1. The angles between the vectors represent the correlation structure among the environments. Together the line points and environment vectors enable a characterisation of regions of similar performance for particular environments. The data were modelled in 2 dimensions and the effectiveness of the model was determined by the proportion of the sum of squares (after transformation) accounted for by each vector. The squares of the singular values are measures of the sum of squares (after transformation) accounted for by each vector. Kempton (1984) illustrated the use of various forms of bi-plots for multi-environmental data. The AMMI model (Gauch 1988) uses SVD on the G \times E matrix and the AMMI plot is one of the many forms of bi-plot that can be used to investigate multi-environmental data. Mirzawan *et al.* (1994) successfully used the bi-plot on environment standardised data for displaying the performance of sugar cane clones. A joint linear regression (Finlay and Wilkinson 1963) was also carried out but as it did not model as much of the variation (measured by the proportion of the sum of squares explained by the regression) as the ordination it is not reported further.

Results and discussion

General trends

The proportion of the total sum of squares (SS) due to environment (E), line (L), and L \times E associated with each source varied from 70% to 95% for environments, 1% to 12% for lines, and 4% to 18% for L \times E interaction (Table 2). These values reflect the proportions due to these sources obtained in other field crops (DeLacy *et al.* 1990). The major proportion of the information attributable to environments directly reflects a much wider range of environment means than of line means in these trials. L \times E interaction was larger than line main effects in all trials, with the ratio of L \times E to line SS varying from 1.5 to 4.33. It is preferable to examine the importance of sources of variation by generalised mixed

models. However, as original data were not available and the effectiveness of pattern analysis models is (still) commonly evaluated by proportion of total SS explained by the model (these are R^2 values or coefficients of determination) these statistics are given for comparison purposes.

Truncation levels where 80% or more of the L \times E interaction was retained led to line groups of 4, 5, or 6, giving 40–70% reduction in array size and retaining 64–96% of the line and 80–86% of the L \times E SS, respectively (Table 2). The percentage sums of squares accounted for by the sums of vectors 1 and 2 in the bi-plots for each year ranged from 56% in 1988 to 83% in 1986, with a mean of 72.6%. Thus, ordination provided a relevant tool for joint presentation of both line and environment relationships in describing the L \times E interaction.

Environment and line trends

Specific adaptation to low-yielding environments (Hermitage standard, Hermitage narrow, and Clifton) were shown by the lower yielding groups 24 and 11, corresponding to the main split in the 1983 and 1986 classifications (Fig. 1). In both cases these groups were lower yielding in irrigated environments with high yield potential (Fig. 2a), whereas the higher yielding groups were relatively stable in yield performance over environments. Lines in both lower yielding groups were predominantly traditional and early-maturing, selected for local adaptation to rainfed environments, but which displayed both greater lodging and susceptibility to rust and blight in the high-yielding environments. Their advantage in lower yielding environments was due to a combination of a terminal drought and early maturity. These trends were reinforced in the bi-plots where these traditional lines (Gallaroy, Kerman, Actolac, W1401, and C15 in 1983, and Gallaroy and Actolac only in 1986) occupied positions on the biplot on the negative projections of the vectors for the higher yielding environments. In contrast the mainly imported entries in the high-yielding groups occupied positions on the positive projections (higher than average yields) for the high-yielding environments.

More complex L \times E interactions occurred in other seasons. The first split in 1984 separated a low-yielding

group, 17, with only traditional entries (Fig. 1). The performance plot (Fig. 2) did not show any yield penalty for this group at 3 locations which ranged from low-yielding Clifton

to high-yielding Mareeba. However, notably, yield depressions occurred at Hermitage standard (with rust) and at Biloela (lodging, rust) and Inglewood (lodging and reaction

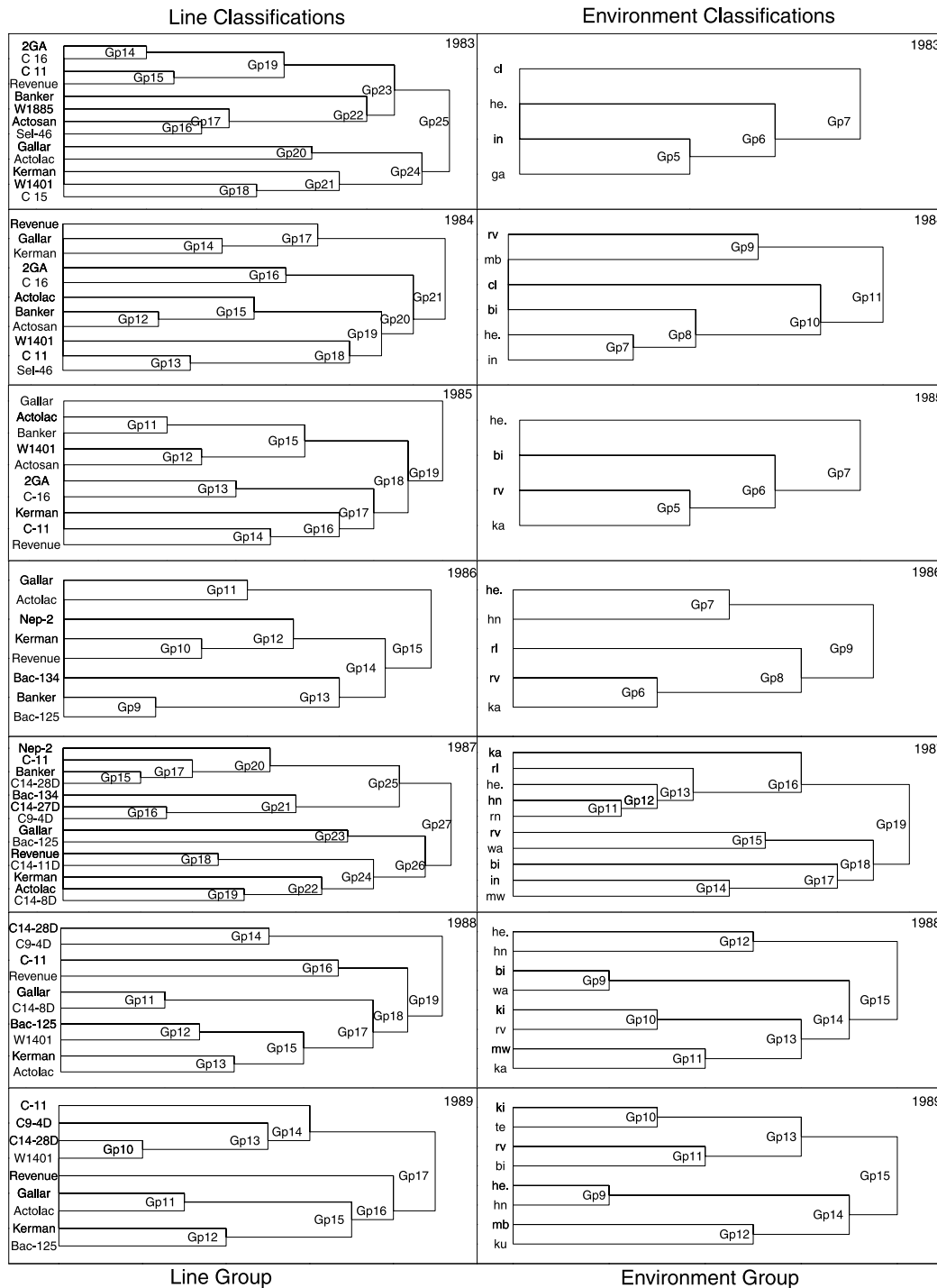


Fig. 1. Dendrograms from the classification of lines and environments using Ward's method on environment standardised mean yield data for navy bean METs for the 1983–1989 seasons (see Table 1 for line names). Environment abbreviations: Biloela, bi; Clifton, cl; Hermitage narrow row, hn; Hermitage standard, he; Inglewood, in; Kairi, ka; Kingaroy, ki; Kumbia, ku; Mareeba, ma; Mt Wooroolin, mw; Redvale, rv; Redvale narrow row, rn; Rocklea (Brisbane), rl; Teakle, te; Wallaville, wa.

to low zinc availability). The performance at the Redvale location is inconsistent despite noting the high disease infestation. In the biplots (Fig. 3) rainfed Redvale and irrigated Mareeba were similar in their discrimination (close together on the biplot) and unrelated to another group of environments, which were similar to each other in discrimination of lines: Hermitage standard (rainfed), Inglewood (irrigated), and Biloela (irrigated) environments. This suggests that water utilisation was not the major source producing different discrimination among environments in 1984. The performance plot for the higher yielding group 20 was stable across environments, reflecting less expression of disease and agronomic and edaphic stresses amongst its entries.

Separation of lines by genetic background was most evident in the first dendrogram split in 1989 (Fig. 1), where nearly all imported and newly bred lines (except W1401) were in group 16, while traditional lines were in group 15. The performance plot for the latter (Fig. 2) was consistently inferior over locations. The direction and range of vectors for environments indicated more uniformity for $L \times E$ interactions than in other years, associated both with less environmental variation and relatively higher genotypic expressions amongst lines.

Mainly imported and newly bred genotypes were in the higher yielding group 25 in the first split in 1987. The groups had similar performance plots at Redvale narrow row, and Inglewood locations (Figs 2, 3), with group 25 superior at other locations except Wallaville, the only location in all seasons with high levels of sclerotinia, which was arrested with application of a fungicide. Group 26 with traditional lines was depressed at locations with high levels of rust and blight (Hermitage standard, Hermitage narrow) or zinc deficiency (Hermitage narrow). Within group 26, group 23 with early-maturing lines may have partially escaped sclerotinia at Wallaville, while late non-erect varieties in group 24 lodged very badly at the irrigated Biloela location.

An inconsistency was found in 1988 in the performance plot for the high-yielding group 14 of the first dendrogram split, which had relatively low yields at Hermitage standard (wide rows) only. This group comprised 2 newly bred selections and there was no obvious biological explanation for this reversal of trend. Consequently, the lower yielding group 18 with many entries had the most stable $L \times E$ performances to environment.

Erect plant architecture interacted differentially with narrow and wide rows for a management impact on $L \times E$ expression, at Hermitage (1986–1989) and Redvale (1987) (Figs 2, 3). In 1987 despite almost unrelated $L \times E$ interactions between the wide row Hermitage standard and Redvale locations displayed in the bi-plot, the corresponding narrow row trials (Hermitage narrow and Redvale narrow row) indicated near identical $L \times E$ expressions. The 4 locations had similar mean yields (Redvale narrow row 660 kg/ha, Redvale 940 kg/ha, Hermitage narrow 830 kg/ha, Hermitage standard 980 kg/ha, in contrast to Kairi with 3320 kg/ha). The third split of the 1987 performance plots illustrated contrasting

performances at these locations between group 20 (Nep2, C11, Banker, and CH14-28D—very erect lines, imported/bred) and group 21 (BAC134, CH9-4D, and CH14-27D—shorter and less erect). It was unexpected that the effects of plant architecture in narrow rows could transcend the location differences in $L \times E$ found in wide rows. (Disease stresses were low in both groups.) A related event was the marked depression of group 20 at Inglewood, a location heavily infested with broad leafed *Datura* (thorn apple) weed, which remained problematic despite repeated manual weeding. Here the erect genotypes appeared to be at a disadvantage with lower bush heights than group 21, which was reversed at Hermitage narrow locations. Although the error variance was high at Inglewood (Butler *et al.* 2000) and this location would normally be discarded as a failure, the interpretable information was agriculturally relevant. The group performance plots enabled these effects of plant architecture to be recognised.

Specific separations of groups of lines may reflect particular disease, agronomic, and edaphic stresses. Notably, Gallaroy was separated in the first dendrogram split in 1985, a season in which this line expressed high levels of blight (Kairi, Biloela locations), rust (Redvale), and lodging and zinc sensitivity (Biloela). It was also isolated and negative in yield performance at all locations except Hermitage standard on the bi-plot. No obvious factors were associated with the groupings and performance patterns for the second and third splits. In the second split in 1989 Revenue separated from other group 16 entries at Kingaroy and Teakle locations due to severe lodging.

However, in the third split for 1988 the traditional multi-stress sensitive Gallaroy was grouped (11) with CH14-28D, which is erect and disease resistant. Group 11 was lower yielding than group 15 (traditional/early lines) at all locations except Redvale where zinc was deficient despite Gallaroy being very sensitive, whereas Actolac and Kerman in group 15 are less sensitive. In this case discrimination of groups of lines was not associated with identifiable agronomic, pathologic, or edaphic expressions.

Trends over seasons

Over all 7 seasons, there was a strong tendency for North Queensland irrigated environments (Kairi, Mareeba) to be closely clustered with a rainfed environment near Kingaroy (Redvale, Kumbia, or Mt Wooroolin). This occurred in 1984, 1985, 1986, 1988, and 1989. This is quite remarkable given the environment differences in mean yield, the greater disease levels of virus, rust, and bacterial blight at Kingaroy environments, and the wide geographic separation with seasonal differences in sowing time. In contrast there was only weak association between the irrigated and rainfed locations at Kingaroy.

In an analysis of this data set with location centred data, there was a unique separation of Kairi or Mareeba from all other environments. As Kairi or Mareeba was usually the

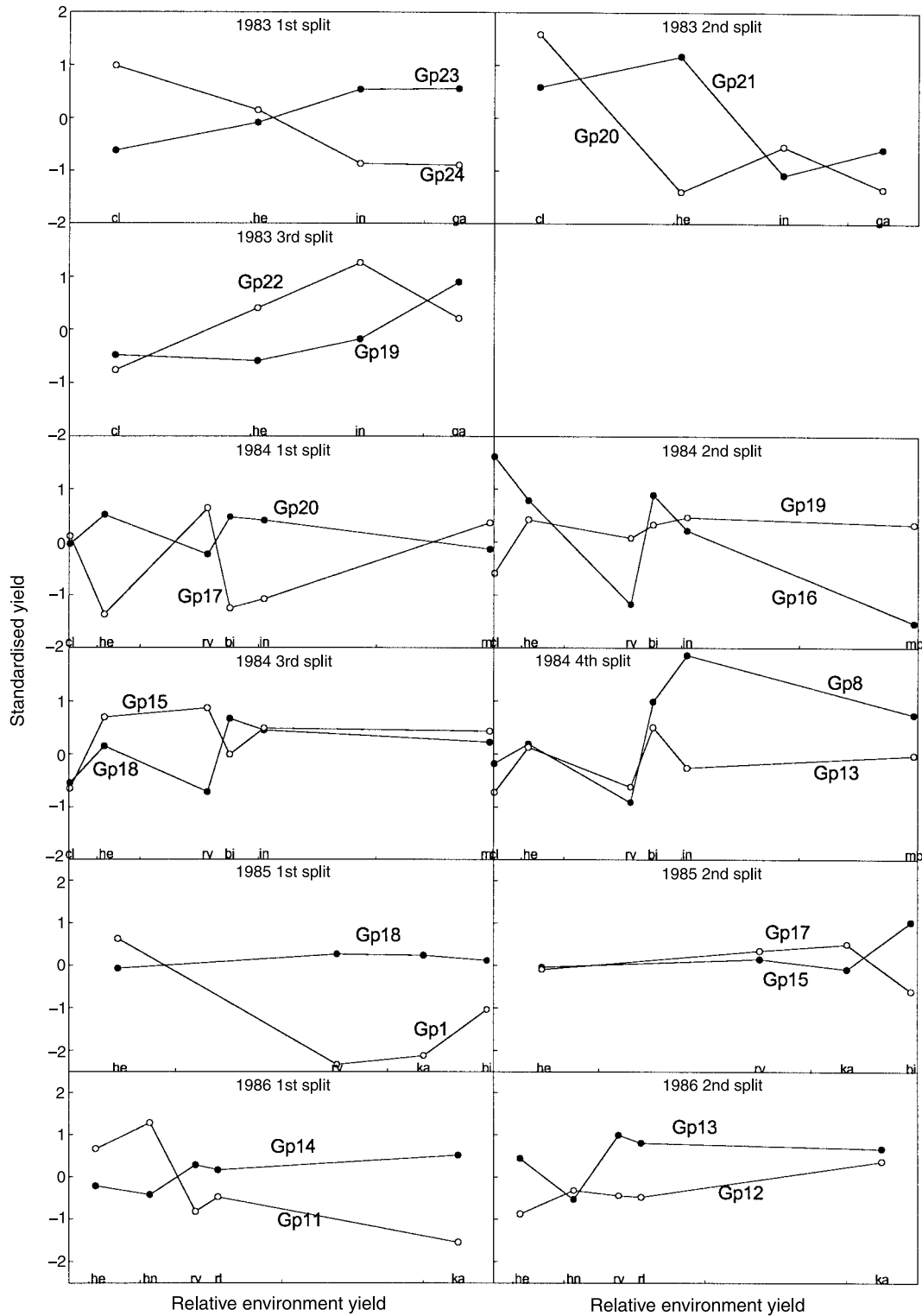


Fig. 2. Performance plots of navy bean line groups from the classifications using Ward's method on environment standardised mean yield data for the 1983–1989 seasons. Means of environment standardised grain yield of the 2 groups at each split of the hierarchy down to a specified truncation level for each year are plotted against environment mean yield (untransformed) from low (LHS) to high.

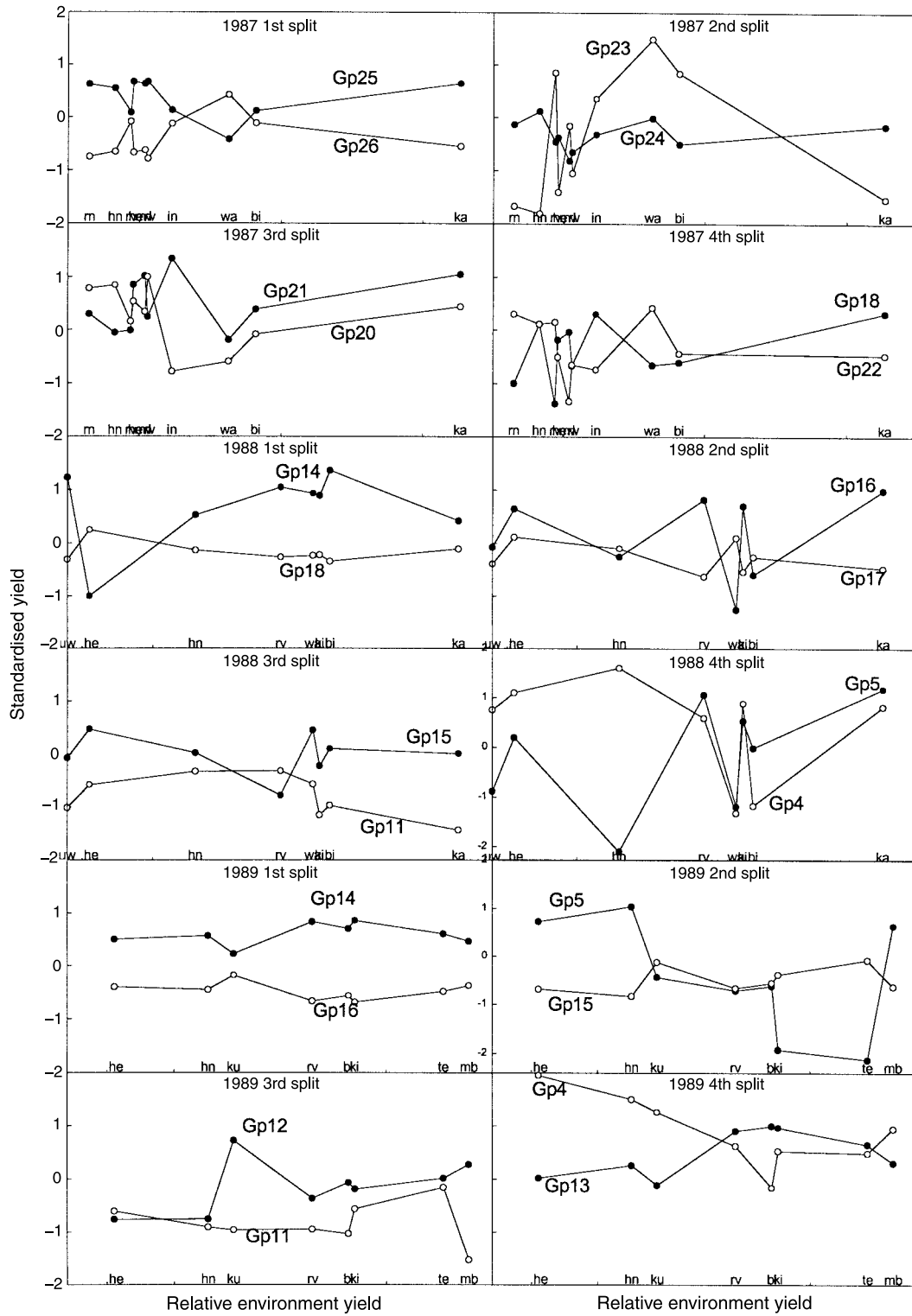


Fig. 2. (continued)

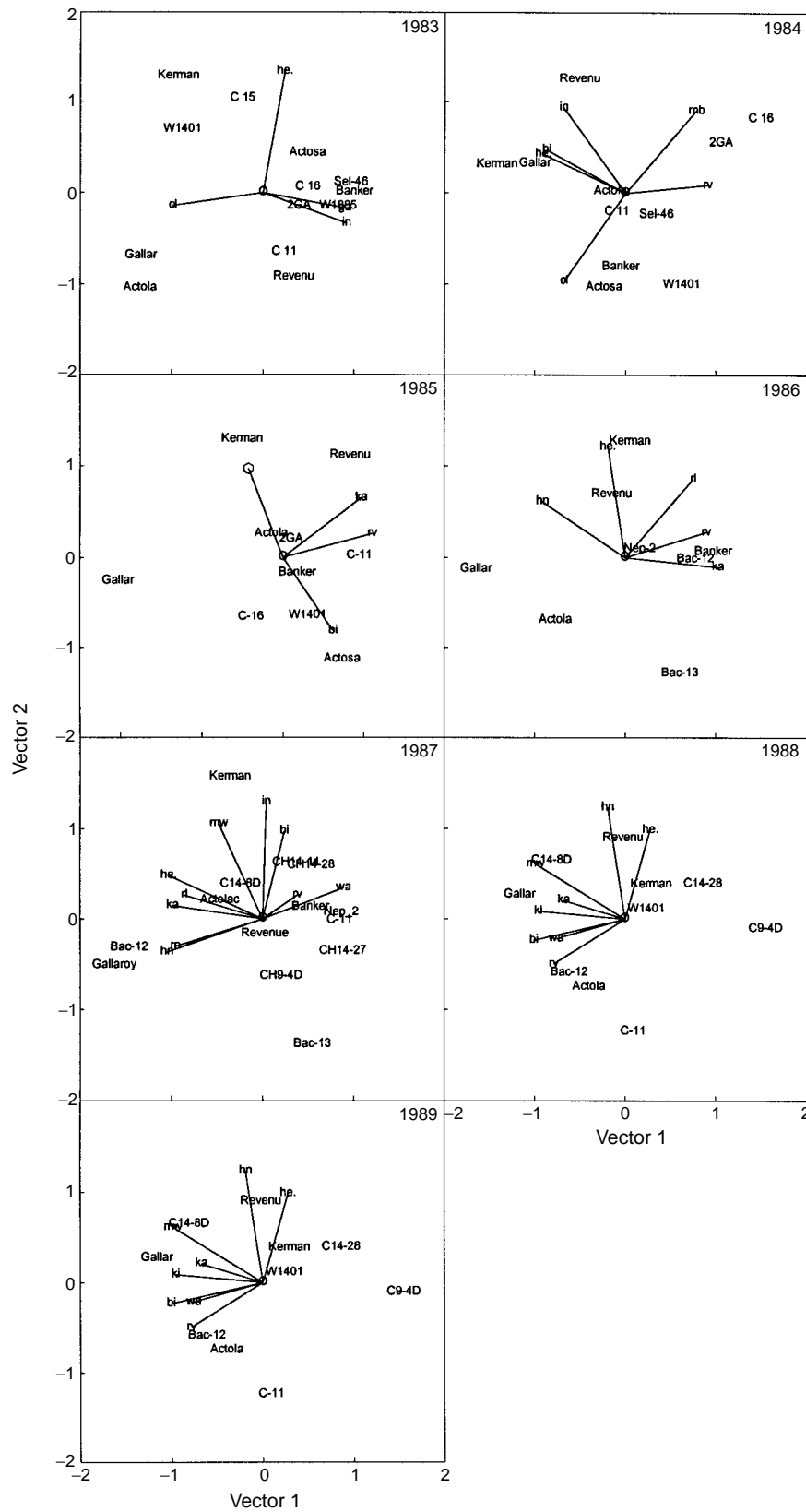


Fig. 3. Bi-plots using singular value decomposition (SVD) of environment standardised mean yield from the ordination of environment standardised mean yield data for navy bean lines for the 1983–1989 seasons.

highest yielding environment this reflected a much wider variance over environment means and greater heterogeneity of variance within environments. The analyses using location centred data were dominated by environment means yields and since the data were location-centred this must reflect differences in variances associated with mean yield. However, when standardised data were used, Kairi was seen to have a greater similarity to Redvale for $L \times E$ discrimination.

Both Hermitage locations (Hermitage standard and Hermitage narrow) clustered together except when there was another narrow row trial at Redvale in 1987 (Fig. 1). In this instance, row spacing had a large effect on environment grouping. In the absence of Redvale, Hermitage standard and Hermitage narrow clustered together and often separated from other locations at the first clustering split (1986, 1988, and 1989).

The 90° relation between the Clifton and Hermitage standard location vectors in 1983 and 1984 (Fig. 3) indicated that genotypic interactions were unrelated between these 2 locations despite their being geographically proximate. Both locations were rainfed and similarly late-sown, both largely escaped frost before harvest in June, and rainfall distribution was broadly similar. This $L \times E$ difference was best demonstrated in the bi-plot, with confirmation in the performance plot (Fig. 2).

The irrigated locations (Inglewood and Biloela) were not closely clustered; however, Biloela and Wallaville were closely grouped for $L \times E$ interaction in 1988. Inglewood and Biloela differ in soil type (sand *v.* clay) and in climate (Butler *et al.* 2000); Biloela and Wallaville are similar in climate but also differ in soil type (clay *v.* sand). The proximate Biloela and Teakle environments were similar both edaphically and in management, and were broadly similar in discrimination amongst lines, differing mainly for effects upon the line Revenue. The only association of Kingaroy locations occurred in 1988 with Kingaroy and Redvale, otherwise there was no geographic affinity for $L \times E$ expression. The results indicate that previous environment history and management, e.g. mixed farming (Mt Wooroolin, Clifton, Rocklea) *v.* pure crop enterprises (Teakle, Kumbia, and other research stations), affects discrimination amongst lines. They also indicate that trial management contrasts of irrigated *v.* rainfed or wide *v.* narrow rows fluctuate seasonally for effects upon line discrimination. Research station environments (especially Hermitage) tend to differ from local farm environments for line discrimination especially under rainfed conditions.

This analysis of environments does not indicate major redundancies which would allow reduction of test environments whilst retaining the major component of $L \times E$ interaction. The effect of narrow rows under rainfed conditions (Hermitage narrow and Redvale narrow row) suggests that 1 narrow row environment per season would be adequate, although only 1 season's data for comparison of narrow row

environments was available. A query remains over the appropriateness of Hermitage as an environment. It is away from bean-growing areas and differed for $L \times E$ interaction from Clifton, which is close geographically and in soil type. The Rocklea location in Brisbane is unrelated to any nearby production zone. The locations from Kingaroy (Redvale, Kingaroy, Kumbia, Mt Wooroolin) northward (Biloela, Wallaville, Kairi, Mareeba) are all within production areas. This study suggests that the number of test environments for bean trials is below the critical level for identifying redundancies, and that provided environments are relevant to bean production, the number of test environments could be increased to optimise $L \times E$ information.

Clustering of genotypes over years was inconsistent (Figs 1, 3). Gallaroy and Actolac, both early old Queensland varieties, were similar in $L \times E$ discrimination in 1983, 1986, and 1989; however, no other close line clusterings were associated with these years. Both 2GA and C16 clustered very similarly in 1983 and 1985, while Actolac and Banker were associated in 1984 and 1985. BAC134 and Kerman separated widely on Vector 2 in both 1986 and 1987, thus differed markedly in $L \times E$ interaction, but at different locations each year. Newly bred selections clustered variously with old Queensland and introduced lines in both 1987 and 1988; erect lines (C11, C16 Banker, plus CH14 selections) closely clustered with non erect lines in all years, and maturity only partly separated groups in 1986 and 1989 with the early BAC125 having a different $L \times E$ pattern from Actolac/Gallaroy (all 3 being early).

The results do suggest that erect non lodging genotypes perform relatively well in narrow rows. Clustering patterns were, therefore, complex and somewhat unique for each year, within which certain environment performances for clusters were readily interpretable and others unexplainable. The newly bred lines tended to be low yielding (standardised data) and often were positioned on the opposite side of the general mean away from the environment mean in the vector analyses. This was not apparent from an examination of the raw data and seasonal mean yields. In general, neither pedigree (old Queensland, introduced, or newly bred) nor growth habit (erect *v.* spreading, determinate *v.* indeterminate, early *v.* late, or disease resistances) consistently affected the pattern of clustering for $L \times E$ effects within years, but where such groupings occurred they could be identified using pattern analyses.

Discussion

This is the first report for pattern and ordination analyses of navy bean trials in Australia. These analyses indicate that relatively small breeding programs which target a wide geographic area can be assisted with data interpretation through pattern analyses to reveal underlying patterns of adaptation.

It is preferable to examine the importance of sources by generalised mixed models rather than coefficients of deter-

mination. However, as original data were not available and the effectiveness of pattern analysis models is commonly evaluated by proportion of total SS explained by the model (these are R^2 values or coefficients of determination), these statistics were given for comparison purposes. The advantage of coefficients of determination is that they are easily understood. A major disadvantage is that the statistics are firmly rooted in fixed model concepts and hence do not have the flexibility and robustness of the general mixed model. Currently the pattern analyses methods are being generalised to mixed models and these are likely to replace the present fixed model approaches.

Line \times environment interaction dominated genotypic expression in all years and, given the inadequacy of the joint linear regression approach, is effectively described using a combination of dendrograms, performance plots, and bi-plots. Dendrograms display hierarchical relationships among line or environment groups. The group performances are displayed by aggregating lines with similar performances into groups, which are then plotted against an environmental index (often environment mean yields). Bi-plots enable both the relation of each genotype to each environment to be examined from perpendicular plots of line values onto respective vectors for environments and the relationships of the discrimination provided by the environments to be compared by examining the length and angles of the environment vectors. These diagrams provide insights into differential modes of genotypic interactions with environments that are not available or difficult to visualise from the normal tables of means or ranks of genotypic mean values.

The breeding lines were only from the first round of selection and further analyses of 4th–5th recurrent selection cycle entries (Redden and Usher 1993) may provide a better analysis of the contribution of plant breeding.

Alternatively, analysis of genotypes using untransformed data can be considered. Unlike environment analyses for which there is theoretical justification for the use of standardised data to investigate L \times E interaction (Cooper and DeLacy 1994), there is no compelling justification for using standardised data rather than some other transformation for analyses of lines. For example, with pattern analysis with location centred data, observations were that Gallaroy was consistently lower yielding especially at high-yielding environments, erect genotypes had an advantage under narrow row spacing, late maturity was an advantage at irrigated environments, and early maturity an advantage in dry seasons at rainfed environments. The interpretation of genotypic yield performance with location centred data was somewhat simpler in relation to phenology, growth habit, and diseases than the analyses using standardised data. Which interpretation is more meaningful may depend on the purpose of the analysis. The analyses of line performance using location centred data emphasised attributes of the line performance important for commercial bean production, noting irrigation,

row spacing, and overall farm management strategies. Analyses using standardised data emphasised attributes of the line performance important for selection. Hence, the purpose of the analyses of line and of L \times E interactions influences the appropriateness of an analytical method in interpreting trial data.

The analyses demonstrate the importance of taking season \times line interaction into account due to changes in correlation amongst environments when applying selection each year for lines for further testing. Modification of selection can be assisted by noting whether seasons are typical or not for pattern of environment relationships for discrimination of lines. The approach has implications for allocation of resources. Very few farms cultivate navy beans near Hermitage, and since Hermitage has distinct line discrimination, trials here may retard selection of widely adapted varieties.

Maturity level was significant for discrimination in some years and was not accounted for in standard analyses of variance (Butler *et al.* 2000); however, the initial selection of progeny was biased towards late maturity as this was generally positively associated with grain yield. Thus to breed for earliness a separate program would be desirable.

For high input agriculture erect habit with resistance to lodging is important both for the direct harvest option and to reduce risk of the disease sclerotinia.

Conclusions

The clustering and ordination analyses (pattern analysis) provide a good description of performance of lines and of environments in any one year. These analyses demonstrate applicability of these methodologies to a navy bean breeding program, consistent with their reported applicability to cereal, sugar cane, and cotton breeding programs.

This approach aids investigation of data for any one year, i.e. for selection and scanning of data. In 1987 the 2 narrow row environments were found to be more similar in discriminating among lines than corresponding trials with wide rows at the same environments, to group lines into introduced and new lines with erect habit *v.* traditional local varieties with low spreading canopies. Data from small-sized breeding programs with relatively small data sets and large imbalances of environments over years can be usefully interpreted with pattern analysis.

These analyses, through data reduction with minimal loss of information, revealed aspects of the data not readily apparent. In 1987, the lines with erect habit were disadvantaged at a very weedy environment (Inglewood) in comparison with traditional lines, and in 1986 drought conditions favoured early-maturing lines at the low-yielding Hermitage environments only. While in many cases the grouping of lines and of environments was biologically interpretable, not all patterns and bi-plots were explainable on the basis of available supplementary information.

The use of separate pattern analyses as a means of interpreting historical data over many environments and a series of years can, however, appear tedious and repetitious. To overcome the problem of an apparently excessive number of descriptive graphs and tables, methods of sequential analysis over environments and years have been applied and are described in paper 3 of this series (DeLacy *et al.* 2000). In particular this approach enables the evaluation and choice of environments for selection purposes in plant breeding programs.

Acknowledgments

The notable contributions of the following to the successful conduct of trials is especially appreciated: J. Hardy, D. Lack, B. Hall, A. Fernandes, G. Gibson, D. Younger.

References

- Basford KE, Kroonenberg PM, DeLacy IH (1991) Three-way methods for multiattribute genotype × environment data: an illustrated partial survey. *Field Crops Research* **27**, 131–157.
- Bradu D, Gabriel KR (1978) The bi-plot as a diagnostic tool for models of two-way tables. *Technometrics* **20**, 47–68.
- Burr EJ (1968) Cluster sorting with mixed character types. I. Standardization of character values. *Australian Computer Journal* **1**, 97–99.
- Burr EJ (1970) Cluster sorting with mixed character types. II. Fusion strategies. *Australian Computer Journal* **2**, 98–103.
- Butler DG, Redden RJ, DeLacy IH, Usher T (2000). Analysis of line × environment interactions for yield in navy beans. 1. Components of variation. *Australian Journal of Agricultural Research* **51**, 597–605.
- Byth DE, Eisemann RL, DeLacy IH (1976) Two-way pattern analysis of a large data set to evaluate genotypic adaptation. *Heredity* **37**, 215–230.
- Cooper M, DeLacy IH (1994) Relationships among analytical methods used to study genotypic variation and genotype-by-environment interaction in plant breeding multi-environment experiments. *Theoretical and Applied Genetics* **88**, 561–572.
- Cooper M, DeLacy IH, Basford KE (1996) Relationships among analytical methods used to analyse genotypic adaptation in multi-environment trials. In 'Plant adaptation and crop improvement'. (Eds M Cooper, GL Hammer) pp. 193–224. (CAB International/IRRI/ICRISAT: Wallingford)
- DeLacy IH, Basford KE, Cooper M, Bull JK, McLaren CG (1996a) Analysis of multi-environment trials—an historical perspective. In 'Plant adaptation and crop improvement'. (Eds M Cooper, GL Hammer) pp. 39–124. (CAB International/IRRI/ICRISAT: Wallingford)
- DeLacy IH, M Cooper, KE Basford (1996b) Relationships among analytical methods used to study genotype-by-environment interactions and evaluation of their impact on performance to selection. In 'Genotype-by-environment interaction'. (Eds MS Kang, HG Gauch) pp. 51–84. (CRC Press: Boca Raton, FL)
- DeLacy IH, Eisemann RL, Cooper M (1990) The importance of genotype by environment interaction in regional variety trials. In 'Genotype-by-environment interaction and plant breeding'. (Ed MS Kang) pp. 287–300. (Louisiana State University: Baton Rouge, LA)
- DeLacy IH, Redden RJ, Butler DG, Usher T (2000) Analysis of line×environment interactions for yield in navy beans. 3. Pattern analysis of environments over years. *Australian Journal of Agricultural Research* **51**, 619–628.
- Eckart C, Young G (1936) The approximation of one matrix by another of lower rank. *Psychometrika* **1**, 211–218.
- Finlay KW, Wilkinson GN (1963) The analysis of adaptation in a plant breeding programme. *Australian Journal of Agricultural Research* **14**, 742–754.
- Fox PN, AA Rosielle (1982) Reducing the influence of environmental main-effects on pattern analysis of plant breeding environments. *Euphytica* **31**, 645–656.
- Gabriel KR (1971) The bi-plot-graphical display of matrices with application to principal component analysis. *Biometrika* **58**, 453–467.
- Gabriel KR (1981) Bi-plot display of multivariate matrices for inspection of data and diagnosis. In 'Interpreting multivariate data'. (Ed V Barnett) pp. 147–173. (John Wiley: Chichester, UK)
- Gauch HG (1988) Model selection and validation for yield trials with interaction. *Biometrics* **44**, 705–715.
- Gauch HG, Zobel RW (1996) AMMI analysis of yield trials. In 'Genotype-by-environment interaction'. (Eds MS Kang, HG Gauch) pp. 85–122. (CRC Press: Boca Raton, FL)
- Kempton RA (1984) The use of bi-plots in interpreting variety by environment interactions. *Journal of Agricultural Science* **103**, 123–135.
- Mirzawan PDN, Cooper M, DeLacy IH, Hogarth DM (1994) Retrospective analysis of the relationships among the test environments of the Southern Queensland sugarcane breeding program. *Theoretical and Applied Genetics* **88**, 707–716.
- Mungomery VE, Shorter R, Byth DE (1974) Genotype × environment interactions and environmental adaptation. I Pattern analysis—application to soya bean populations. *Australian Journal of Agricultural Research* **25**, 59–72.
- Redden RJ, Rose JL, Gallagher EC (1985) The breeding of navy and culinary beans in Queensland. *Australian Journal of Agricultural Research* **25**, 470–479.
- Redden RJ, Usher T (1993) Genetic diversity and multi-trait selection in a navy bean breeding program. In 'Focussed plant improvement towards responsible and sustainable agriculture. Proceedings: Tenth Australian Plant Breeding Conference'. Vol. 2. pp. 49–50. (Eds BC Inrie, JB Hacker) (University of Queensland Printers.)
- Ward JH (1963) Hierarchical grouping to optimise an objective function. *Journal of the American Statistical Association* **58**, 236–244.
- Williams WT (1976) 'Pattern analysis in agricultural science.' (Elsevier Scientific Publishing Company: Amsterdam)
- Wishart D (1969) An algorithm for hierarchical classification. *Biometrics* **25**, 165–170.

Manuscript received 15 September 1997, accepted 1 March 2000