TWO-WAY PATTERN ANALYSIS OF A LARGE DATA SET TO EVALUATE GENOTYPIC ADAPTATION

D. E. BYTH, R. L. EISEMANN and I. H. DE LACY

Department of Agriculture, University of Queensland, St Lucia, Queensland, Australia, 4067

Received 19.ii.76

Summary

A method for the analysis of genotype × environment interaction in large data sets is presented and applied to yield data for 49 wheat cultivars grown in each of 63 international environments. Pattern analysis using numerical classification defined separately groups of cultivars and groups of environments, based on similarities in yield performance. The group structure for cultivars was interpreted in terms of similarities and differences in cultivar mean yield and/or cultivar yield response patterns across environments. In addition, the cultivar groups reflected differences in genetical and selectional origin. Environment groups largely reflected differences in the average mean yield of the set of cultivars, but some groups showed differences in response patterns related to differential rust incidence.

The cultivar and environment groupings were superimposed on the original data matrix, reducing it to a 100 cell 10 × 10 matrix of group means. The efficiency of the reduction process was measured by comparing the variation retained in the reduced matrix with the total variation available in the original data matrix. Further study of the information retained by the 10 × 10 matrix was made by plotting cultivar group yields and cultivar group interaction effects against an environment group index. The reduction process achieved a size reduction of 97 per cent with the loss of 18 per cent of the total variation available in the original unreduced matrix. Partitioning was used to identify the nature of this loss. However, the information retained in the reduced matrix maintained the integrity of the cultivar group yield response patterns and allowed comparison of cultivars on a group basis across the environmental range. This reduced greatly the complexity of analysis of cultivar performance and interaction patterns, and simplified the identification and specification of differences in response among cultivars.

1. Introduction

MULTI-ENVIRONMENT testing of genotypes may involve a broad range of environments, large numbers of genotypes, and substantial genotype × environment (G × E) interactions. Efficient and objective comparison and interpretation of the resulting genotypic performance is a major problem in plant breeding. Various analytic approaches may be useful, and Mungomery, Shorter and Byth (1974) formalised some of these in relation to the analysis of genotypic performance across environments.

A linear regression analysis, first applied by Finlay and Wilkinson (1963), has subsequently been developed into a powerful and informative method of G × E analysis (Perkins and Jinks, 1968a, b; Freeman and Perkins, 1971). A basic objective of this technique is to identify systematic variation in performance, and it can be very informative where genotype × environment interactions have high linear association with the environmental index. However, where a low degree of linearity exists, the technique may be at best uninformative, and at worst misinformative regarding genotypic performance over environments.

Multivariate methods have been applied in the analysis of G × E interaction and genotypic adaptation. Principal component analysis was used by Suzuki (1968) in studying strain adaptability, and Perkins (1972) compared this technique with linear regression for *Nicotiana rustica* data. Freeman and Dowker (1973) analysed carrot variety performance with a two-way principal component analysis. Mungomery *et al.* (1974) showed that pattern analysis methods were a useful alternative means of studying the performance of large sets of cultivars over environments, using soybean data. Pattern analysis is a general term encompassing the use of both cluster analysis and ordination to examine data structure. Cluster analysis was used by Chuang-Sheng Lin and Thompson (1975) to delineate groups of genotypic performance types, based on differences among linear regression coefficients. In this context, Perkins and Jinks (1968b) used a simple binary grouping system to characterise the non-linear portion of G × E interaction, based on the correlation of the deviations from linear regression.

Environmental aspects of G × E interactions and adaptation were examined by Horner and Frey (1957), using a grouping system which restricted the range of environments, and by Abou-El-Fittouh, Rawlings and Miller (1969), who used cluster analysis to define homogeneous regions for cotton variety tests in the U.S.A. Goodchild and Boyd (1975) analysed wheat yields over a 40-year period in Western Australia by principal component analysis.

The difficulties of detailed analysis and interpretation of large data sets are particularly acute in international plant improvement programmes. These programmes involve annual evaluation of many cultivars of wheat, maize, rice and other crops in numerous international environments, and have a major influence on world agriculture. However, objective and rapid analysis of such data has been inhibited by the complexity of the task and the lack of suitable procedures.

This paper extends the pattern analysis approach of Mungomery *et al.* (1974) to provide an objective method of analysis of cultivar adaptation and environmental contribution in larger data sets.

2. METHODS

An analysis of data from the fourth international spring wheat yield trials * conducted in 1967-68 (MacKenzie, Mexas, Finlay and Borlaug, 1971) was attempted. These data are representative of the large sets of plant adaptation data collected annually by various international programmes. The experiment involved the evaluation of 49 cultivars of wheat in each of 63 international environments. These cultivars represented the principal types of spring wheat grown throughout the world. Their parentage, origin and aspects of their habit, disease resistance and quality have been described (MacKenzie *et al.*, 1971). The trial sites were located in wheat-growing areas in all continents, and varied widely in latitude, climate and altitude.

Six row plots 2·5 m long were grown with three replications in each trial. Various agronomic, phenological and disease response characters were recorded. The cultivar means for a number of characters along with their

* Fourth International Spring Wheat Yield Nursery (ISWYN 4) coordinated by CIMMYT (International Maize and Wheat Improvement Centre), Mexico.

standard errors, coefficients of variation and LSD values were presented for each trial (MacKenzie *et al.*, 1971). This study involved analysis of seed yield (t/ha) only, using the cultivar means for each trial.

Pattern analysis, specifically numerical classification, was applied in the present study in the manner suggested by Mungomery *et al.* (1974). The approach was extended to incorporate a two-way classificatory analysis of the ISWYN 4 data. Use was made of the CSIRO computer program MULTCLAS employing an agglomerative, hierarchical clustering technique (Williams, 1971) with an incremental sum-of-squares sorting strategy (Burr, 1970). Variance-standardised squared Euclidean distance was the dissimilarity measure between elements to be classified (Burr, 1968).

Initially, a classification of the 49 cultivars was constructed using the 63 environment scores for each cultivar as attributes. For the purposes of this study, the hierarchy was truncated arbitrarily at the 10 group level. A classification of the 63 environments, using the 49 cultivar scores in each environment as attributes, was also constructed. The hierarchy was again truncated at an arbitrary 10 groups. This two-way analysis paralleled the "normal" and "inverse" analyses developed for ecological classification (Williams and Lambert, 1961).

The groups defined by the two classificatory analyses were superimposed on the original (49 × 63) data matrix, so that at the group level it was reduced to a 100 cell (10 × 10) two-way table of group means. Comparison of the variation in the reduced matrix with that available in the original unreduced data matrix was used to interpret the value of the two-way classification procedure in retaining useful information. This comparison was achieved by (i) analyses of the variation existing among and within the defined groups, (ii) study of the performance of the 10 cultivar groups across the 10 environment groups.

3. RESULTS AND DISCUSSION

A pooled analysis of variance across cultivars and environments (table 1) indicated substantial differences in cultivar and environment mean yield, and in cultivar × environment (C × E) interaction. Further, linear regression accounted for only 8·9 per cent of the C × E interaction, implying that

TABLE 1

Pooled analysis of variance for seed yield (t/ha) across cultivars and environments, ISWYN 4

Source	d.f.	S.S.	M.S.
Environments (E)	62	4833·46	77·96
Cultivars (C)	48	564·01	11·75
Cultivar × Environment (C × E)	2976	1501·65	0·50
Heterogeneity among regressions	48	133·55	2·78
Remainder	2928	1368·10	0·47
Error*	—	—	0·09
Percentage total S.S. in C × E		21·8	
Percentage C × E S.S. accounted for by regression		8·9	
Grand mean		2·903 t/ha	
Range of environment means		0·745- 5·749 t/ha	
Range of cultivar means		1·785- 3·548 t/ha	
Range of coefficients of variation		8·0- 35·0%	

* For details of error estimate, see text.

systematic variation, if it existed, could not be characterised effectively by this technique.   MacKenzie *et al.* (1971) did not present error variances for each environment.   The coefficient of variation and environment mean were therefore used to obtain an estimate of error for each environment, and these values were pooled to obtain an overall estimate on a mean plot basis (table 1).

(i) *Cultivar classification*

The dendrogram corresponding to the hierarchy for the 10 cultivar groups is shown in fig. 1*a*.   In the major discontinuity of this classification (node A in fig. 1*a*), mean yield differences were clearly important, in that groups I-V had lower mean yields than groups VI-X.   However, it was apparent that some nodes within these two major sub-groups depended on factors other than mean yield.   For example, groups VIII, IX and X had similar mean yields but VIII and IX were fused relatively earlier, suggesting they were more similar in their yield response patterns over the 63 environments, while group X had the most dissimilar yield response pattern of these three groups.   In contrast, groups VI and VII exhibited yield superiority.

From a study of the individual cultivars within groups (data not presented here), it was apparent that substantial overlap of cultivar mean performance occurred among some groups.   These data *in toto* suggested that the classification identified groups of cultivars which differed in mean yield across the 63 environments or in the pattern of yield response over these environments, or both.
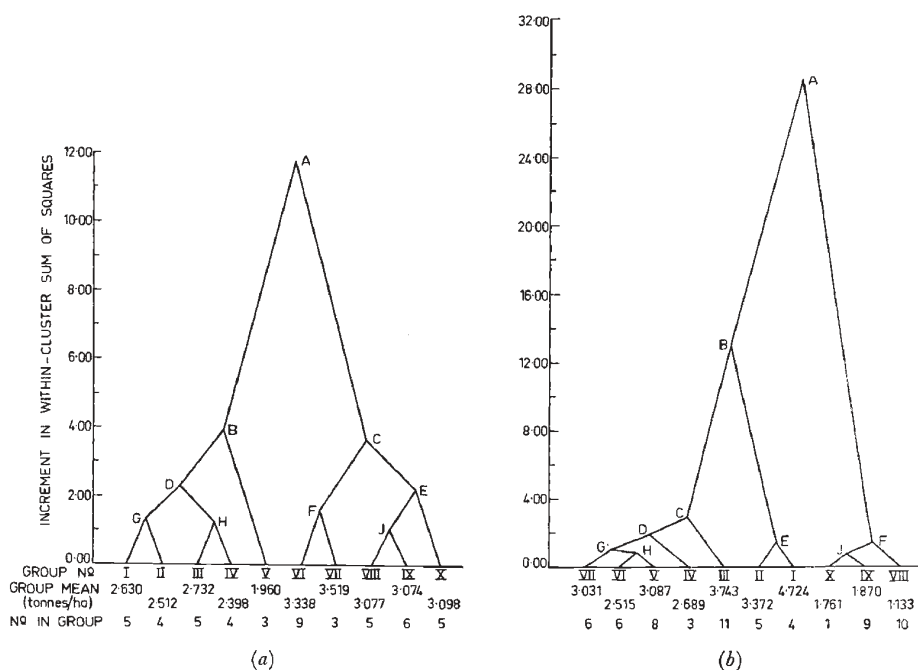

FIG. 1.—Dendrograms for the cultivar and environment classification of seed yield: (*a*) cultivar classification, (*b*) environment classification.

The efficiency of the classification in formulating groups differing in mean yield or patterns of response was evaluated by partitioning the variation for cultivars and C×E interaction into among- and within-group sources for the 10 cultivar groups (table 2a). Clearly, variation for cultivars represented differences in mean yield across the 63 environments, and C×E interaction variation represented differences in the patterns of performance over these environments. Variation among the 10 cultivar groups accounted for 91·3 per cent of the cultivar sum of squares and its mean square was approximately 45 times that of the within-group mean

TABLE 2

Partition of variation for seed yield (t/ha) into the among- and within-group components for (a) cultivar, (b) environment and (c) two-way cultivar-environment classification

	Source	d.f.	S.S.	M.S.	Partition of S.S. Among and Within groups (%)
(a)	Cultivars (C)	48	564·01	11·75	—
	Among C groups	9	514·79	57·20	91·3
	Within C groups	39	49·22	1·26	8·7
	Environments (E)	62	4833·46	77·96	—
	C×E	2976	1501·65	0·50	—
	Among C groups×E	558	801·16	1·44	53·4
	Remainder	2418	700·49	0·29	46·6
(b)	Cultivars (C)	48	564·01	11·75	—
	Environments (E)	62	4833·46	77·96	—
	Among E groups	9	4661·03	517·89	96·4
	Within E groups	53	172·43	3·25	3·6
	C×E	2976	1501·65	0·50	—
	Among E groups×C	432	710·82	1·65	47·3
	Remainder	2544	790·83	0·31	52·7
(c)	Cultivars (C)	48	564·01	11·75	—
	Among C groups	9	514·79	57·20	91·3
	Within C groups	39	49·22	1·26	8·7
	Environments (E)	62	4833·46	77·96	—
	Among E groups	9	4661·03	517·89	96·4
	Within E groups	53	172·43	3·25	3·6
	C×E	2976	1501·65	0·50	—
	Among C groups×Among E groups	81	478·21	5·90	31·8
	Remainder	2895	1023·44	0·35	68·2

square. For C×E interaction, the among-groups partition accounted for 53·4 per cent of the interaction sum of squares and its mean square was approximately five times that of the within-group mean square. It was apparent that classification of the cultivars effectively partitioned variation in both mean yield and response pattern over the 63 environments among the 10 cultivar groups, and that the cultivars within the groups were relatively homogeneous for these aspects.

The group composition reflected, in part, the origin and parentage of the cultivars. All 28 cultivars within node C (fig. 1a) were of Mexican or South American origin, and 25 of these were of Mexican origin or involved parentage derived from the CIMMYT programme in Mexico. All cultivars within the highest yielding group VII were derived in Mexico. In contrast, all 21 cultivars within node B were of non-Mexican origin, and none

involved parentage tracing to the CIMMYT programme or to Norin 10 ancestry. Clearly, a marked difference in performance over environments existed between the " Mexican " and " non-Mexican " cultivars, and this was identified objectively by pattern analysis.

Within the " non-Mexican " material (node B, fig. 1*a*), group V contained three spring wheats of U.S.A. or Canadian origin. Two other North American cultivars were separated from the group V cultivars in classification, and this may have been related to their parentage which involved winter wheat characteristics. Cultivars in groups I-IV were of more diverse origin, and will not be discussed here.

Relating associated information to the classification in this way indicated that parentage, selectional origin and even differences in physiological response had sensible meaning within the cultivar group structure defined by the numerical classification of seed yield.

(ii) *Environment classification*

It was apparent from the dendrogram for the classification of the 63 environments (fig. 1*b*) that mean yield differences were important in constructing the classification. The major discontinuity (node A, fig. 1*b*) showed a large mean difference between branches. Environment groups I-VII exhibited a distinct mean yield superiority over groups VIII-X. Further, node B reflected a separation of high-yielding groups I and II from the moderate-yielding groups III-VII. A majority of the remaining nodes also reflected varying differences in mean yield within each of the high-, moderate- and low-yielding nodes.

These large differences in mean yield among the environment groups overshadowed evidence of similarity of yield response patterns over the 49 cultivars. However, groups IV-VII within node D (fig. 1*b*) showed that similarity of mean yields was not the sole criterion for fusion. Groups V and VI were the most disparate in mean yield, and yet linked first at node *H*. This combination then joined group VII which was again marginally the more disparate in mean yield. Finally, group IV fused to give node D. This order was a reversal of the sequence expected on mean yield alone, and could only be interpteted in terms of the relative similarity of yield response patterns in these groups.

An evaluation of the efficiency of classification, similar to that followed in the cultivar classification, is shown in table 2*b*. The among-environment groups source of variation accounted for 96·4 per cent of the environment sum of squares, with a mean square 159 times that of the within-group mean square. For the C × E interaction, the corresponding among-groups partition accounted for 47·3 per cent of the interaction sum of squares, and its mean square was approximately five times that of the within-groups source. These data showed that the classification produced groups of environments which differed in their average level of performance or in the pattern of response they elicited from the 49 cultivars, or both. Within groups, environments were relatively consistent in average yield and cultivar response.

In general, the composition of the environment groups did not reflect the location of the trials. The over-riding effect of the differences in mean yield was apparent, with the 63 environments ranging from 745 to 5749 kg/ha

(table 1). On the basis of scant information on the trial sites (MacKenzie *et al.*, 1971), it appeared that the low-yielding environment groups VIII, IX and X were characterised by a wide range in length of growing season, latitude and altitude (in and out of the tropics), variable disease incidence and further abnormalities such as late or out-of-season sowing, drought, heavy rain, and heat. In contrast, the high-yielding groups I and II involved sites which were in a relatively narrow latitudinal range (27·5°-38·5°N), mostly irrigated, and had low to normal disease incidence. Group I included sites greater than 700 m above sea-level, and was lower in mean yield and higher in disease incidence than Group II.

The medium-yield groups III-VII were diverse in latitude. It was noted previously that the fusion of groups within node D (fig. 1*b*) was different from that expected on the basis of mean yield. It is possible that the classificatory structure was influenced by differential cultivar performance in these environments, induced by variation in disease incidence. For example, Group VII environments experienced good heavy rust development, while group V environments were relatively free of rust. Further, Group VI environments varied in rust incidence, and the high altitude (>2600 m) group IV environments had good stem and stripe rust development.

A detailed analysis of the characteristics of the environments within and among groups is not attempted here. However, the relative homogeneity of environments within groups for yield (table 2*b*) indicated that they were similar in the pattern of response of all 49 cultivars. The cause of such similarity of response is not known, but it is possible that it may reflect basic characteristics common to these environments.

(iii) *Two-way cultivar-environment classification*

Superimposition of the group structure on the original data matrix allows variation in cultivar yields to be represented by a linear model of the following form:

$$y_{ijkl} = m + C_i + c_{ij} + E_k + e_{kl} + (CE)_{ik} + (Ce)_{ikl} + (cE)_{ijk} + (ce)_{ijkl} \tag{1}$$

Let n_i = number of cultivars in cultivar group i, $\sum_i n_i = 49$; and m_k = number of environments in environment group k, $\sum_k m_k = 63$. We may then define the terms in model (1) as follows:

y_{ijkl} = mean yield of a cultivar in an environment

m = general mean

C_i = additive effect of the ith cultivar group; $i = 1, ..., 10$; $\sum_i n_i C_i = 0$.

c_{ij} = additive effect of the jth cultivar within the ith cultivar group;

$j = 1, ..., n_i$; $\sum_j c_{ij} = 0$ for all i

E_k = additive effect of the kth environment group; $k = 1, ..., 10$;

$\sum_k m_k E_k = 0$

e_{kl} = additive effect of the lth environment within the kth environment

group; $l = 1, ..., m_k; \sum_l e_{kl} = 0$ for all k.

$(CE)_{ik}, (Ce)_{ikl}, (cE)_{ijk}, (ce)_{ijkl}$ are the effects arising from the interaction of corresponding additive effects above.

$$\sum_{ik} n_i m_k (CE)_{ik} = 0$$

$$\sum_{il} n_i (Ce)_{ikl} \quad = 0 \text{ for all } k$$

$$\sum_{jk} m_k (cE)_{ijk} \quad = 0 \text{ for all } i$$

$$\sum_{jl} (ce)_{ijkl} \quad = 0 \text{ for all } i \text{ and } k.$$

All effects are uncorrelated. This model allows the sum of squares of the original data to be partitioned into a set of orthogonal components. No error term is associated with the values in the original 49×63 data matrix.

The schematic presentation of the classificatory procedures (fig. 2) illustrates the reduction in size achieved in the data matrix. Values in the reduced 10×10 table of group means obtained from the original data can be represented by the relevant parts of the model (1) as follows:

$$\bar{y}_{i.k.} = m + C_i + E_k + (CE)_{ik}, \tag{2}$$

where $\bar{y}_{i.k.}$ = mean yield of the ith cultivar group in the kth environment group.

The remaining effects are defined as for model (1).

Partition of the variation associated with model (1) is shown in table 2c. From models (1) and (2), it follows that three sources of variation in this table correspond to the variation in the reduced 10×10 matrix; viz. (i) among cultivar groups, (ii) among environment groups, (iii) among cultivar groups × among environment groups. These three sources accounted for 91·3, 96·4 and 31·8 per cent of the sums of squares for cultivars, environments and C × E interaction, respectively (table 2c). Overall, 82·0 per cent of the original total sums of squares was retained by the 10×10 matrix.

Two aspects are involved here—matrix reduction for ease of interpretation, and information retention. By necessity, these aspects conflict to some degree. Cultivar, environment and two-way classification reduced the effective size of the matrix by approximately 80 per cent, 84 per cent and 97 per cent respectively (fig. 2). Despite this reduction, it was apparent that classification into 10 groups resulted in only marginal loss of information on variation among the cultivar and environment means (table 2a and b). The loss was relatively lower for the environment classification than for the cultivar classification. This was expected on the basis of the earlier contention that the environment classification was based primarily on mean yield. This difference between the classifications was also reflected in the relative loss of informaion on C × E interaction, in that relatively more information from this source was lost by the environment classification.
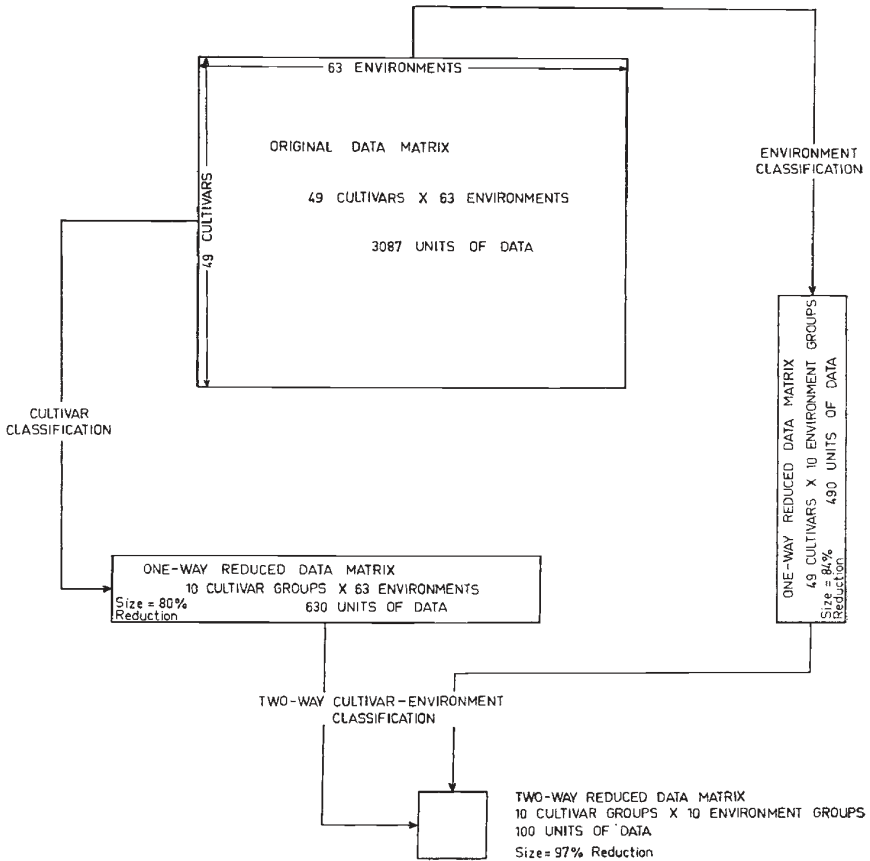
FIG. 2.—Matrix reduction (to scale) resulting from cultivar, environment, and two-way classification of the original data matrix.

Loss of variation due to C × E interaction was more substantial following the combination of the two classifications (table 2c). If this loss of information regarding cultivar group yield response over environments were critical, then it would follow that the differences among cultivar groups would also be reduced or lost. Consequently, the meaningfulness of the variation retained in the reduced matrix was examined further by plotting the yield of each cultivar group against an environment group index (fig. 3). This index involved 10 elements, each defined as the mean yield of all cultivars in an environment group.

The original classification of cultivars was based on performance in all 63 environments. Despite the reduction to a 10-element environment group index, it was apparent that, in general, the response patterns of the cultivar groups over the index (fig. 3) confirmed the observations on group similarity or dissimilarity made as a result of the original cultivar classification (fig. 1a); that is, superimposition of the environment classification caused little disruption of the cultivar group performances.

Within the "Mexican" cultivars, groups VIII and IX were more similar in response pattern over the 10 environment groups than either
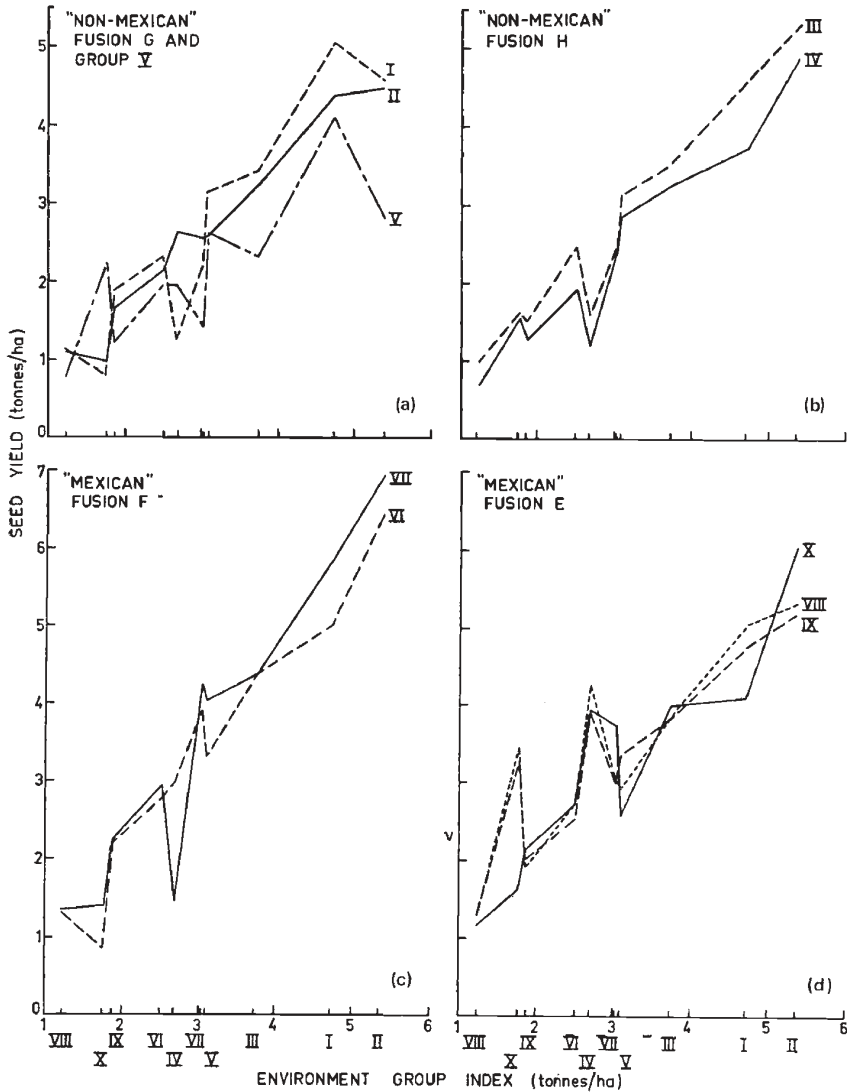
FIG. 3.—Yield performance of cultivar groups across the environment group index for
(a) groups I, II and V, (b) III and IV, (c) VI and VII, (d) VIII, IX and X.

was to group X (fig. 3d). Clearly, the classificatory structure (fig. 1a)
reflected response differences in this case. Similarly, groups VI and VII
were higher yielding and classified separately from VIII, IX and X (fig. 1a),
and revealed a greater response to more favourable environmental conditions
(fig. 3c and d). Groups VI and VII were markedly different in their response
to environment group IV (fig. 3c). This may be related to differential
rust reaction. The high incidence of stripe and stem rust in these environ-
ments (MacKenzie et al., 1971) apparently depressed the yield of otherwise
high-yielding and responsive cultivars of group VII, whereas the group VI
cultivars revealed lower susceptibility to rust attack.

It is interesting to speculate on the origin of differences in response among the groups within the "Mexican" material. Eight of the 12 cultivars within node F (groups VI and VII; fig. 1a), and all cultivars within the high-yielding and responsive group VII, were derived in Mexico. The occurrence of other Mexico-derived cultivars within other lower yielding and less responsive groups (VIII, IX, X) suggested that different selection objectives existed within the CIMMYT programme, or that the selection procedures used were sufficiently broad to encompass a range of objectives. Regardless of the explanation, it is apparent that markedly different response types were produced, deliberately or inadvertently. Cultivar classification allowed identification of the differences, and the two-way classification allowed some generalisations regarding the nature and possible causes of these differences.

Differences in response across the environment group index occurred among the "non-Mexican" groups (fig. 3a and b). These are not discussed in detail here. However, groups III and IV revealed similar response patterns at different mean yield levels (fig. 3b), and both differed in response from that of groups I and II (fig. 3a). Group V differed from groups I-IV in both mean yield and response pattern. In general, these differences coincided with the patterns of group similarity implicit in the original classification of cultivars using the full data matrix (fig. 1a).

From the study of these response patterns, it was apparent that the cultivar classification identified groups of cultivars which differed meaningfully in environmental response, and that the hierarchy reflected the degree of similarity of cultivar group response. Furthermore, the groups maintained their general relationship in the reduced 10×10 matrix obtained by two-way classification.

(iv) Cultivar × environment interaction

Estimates of the $(CE)_{ik}$ effects obtained from model (1) and (2) are plotted against the environment group index in fig. 4, in sets of groups most closely related by classification (fig. 1a). These data represented the among-cultivar groups × among-environment groups interaction source of variation, which accounted for 31·8 per cent of the total interaction sum of squares in the original data matrix. Obviously, these effects contributed to the group response patterns $(\bar{y}_{i.k.})$ plotted in fig. 3, but were confounded with cultivar group mean effect (C_i), and environment group mean effect (E_k).

It was apparent that the patterns in group interaction effects across the environment group index (fig. 4) were similar for cultivar groups most closely related by classification (fig. 1a); for example, groups VI and VII (fig. 4c) were more similar to each other than they were to any other group of the other three sets. A similar situation existed for all other sets of groups (fig. 4). This indicated that the cultivar groups differed substantially in their environmental interaction, that the relationships among groups reflected the degree of difference, and that the group differences were retained despite the reduction from a 63-element environmental index to a 10-element environment group index. These results suggested that, despite the loss of 68·2 per cent of the interaction variation following two-way classification, much of the meaningful information on environmental interaction was retained among the groups.
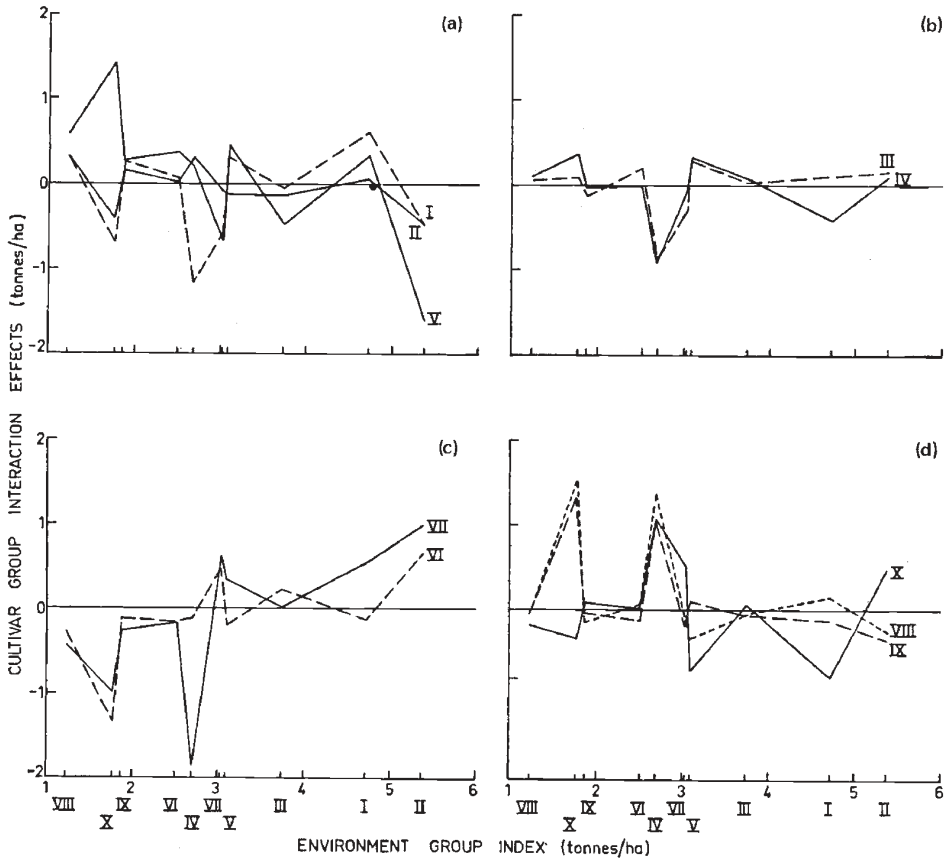
FIG. 4.—Changes in cultivar group interaction effects across the environment group index for (a) groups I, II and V, (b) III and IV, (c) VI and VII, (d) VIII, IX and X.

Clearly, the differences among groups for group interaction effects ($(CE)_{ik}$, fig. 4) were of two forms; viz. (1) differences in magnitude and direction of the specific effects for particular environment groups, and (2) differences in any systematic variation in the effects across the environment group index. Certain examples of the former were discussed previously, e.g. the differential responses of the groups to environment group IV which probably reflected the influence of rust incidence and resistance (figs. 3c, 4c). It is emphasised that most of the group interaction effects (fig. 4) represent a number of cultivars and environments so that the differences among effects provide a firm basis for quantitative physiological analysis of environmental reaction. Further, it was apparent that systematic trends in the magnitude and direction of effects existed, and differed among the groups. For example, groups VI and VII (fig. 4c) revealed apparently increasing, and groups V, VIII and IX (fig. 4a and d) apparently decreasing, interaction effects with increase in the environment group index. Such systematic differences represented basic differences in adaptation, knowledge of which may have considerable physiological, breeding and genetic import- ance. Cultivar groups within sets related by classification were generally
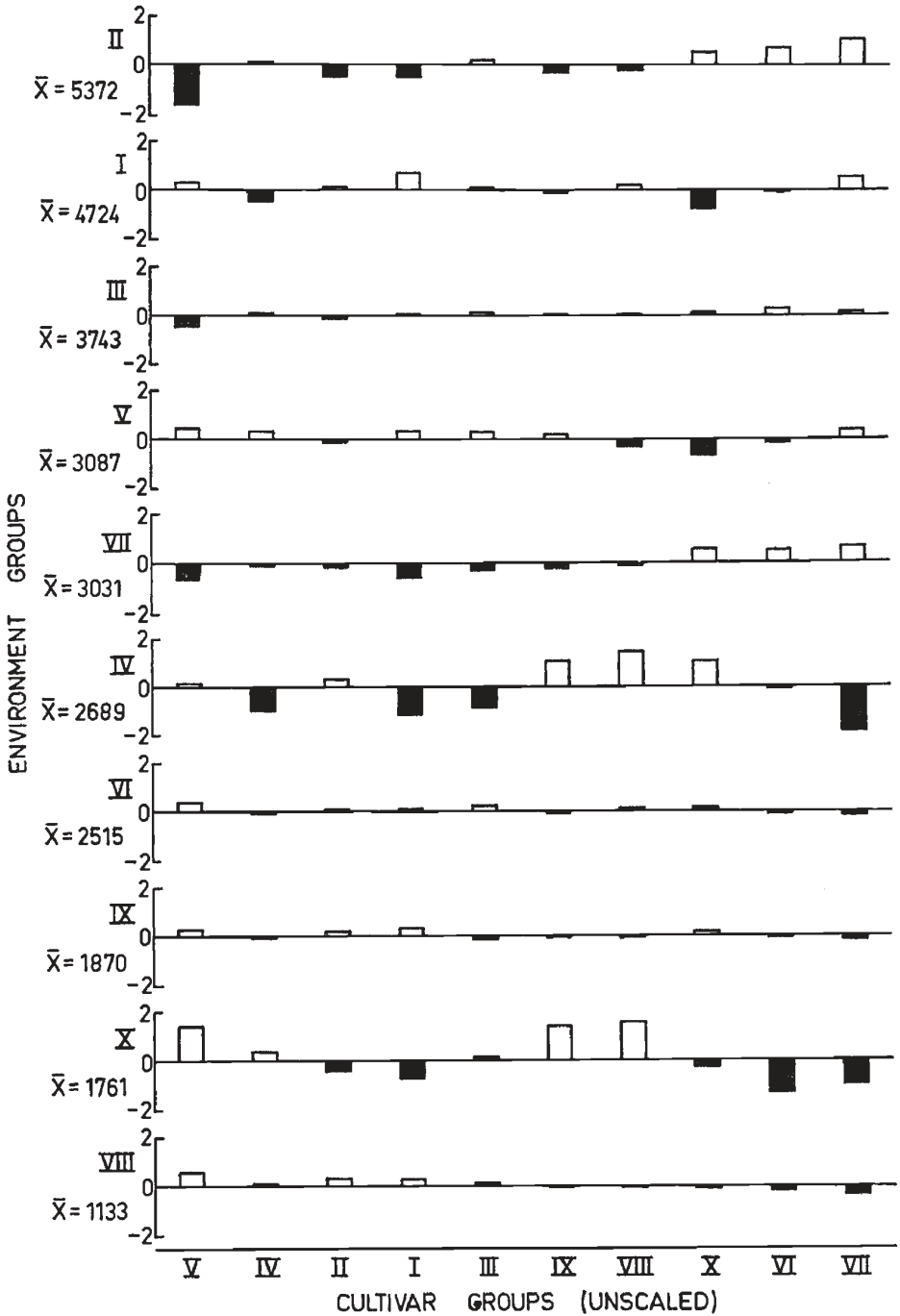
Fig. 5.—Interaction effects (t/ha) for all 100 combinations of the cultivar and environment groups. (Cultivar and environment groups ranked in increasing order of mean yield.)

similar in any systematic response pattern (fig. 4), and only rarely differed substantially from each other. The similar genetic or selectional background of these groups, and the number of cultivars and environments involved in each interaction effect, suggested that these differences and similarities in environment interaction may have practical implications in plant improvement.

An alternative diagrammatic presentation of the cultivar group × environment group interaction effects in histogram form (fig. 5) has been found to be convenient and informative, with respect to both the cultivar and environment groups. Some of the implications are discussed briefly here.

The magnitude and pattern of interaction effects of the various cultivar groups may be examined in relation to the performance level of each environment group (fig. 5). For these data, it was apparent that substantial interaction occurred in a number of the environment groups, regardless of mean performance of the group or of relationship in classification, and that the patterns of interaction effects differed among environment groups. As discussed previously, classification of the environments was influenced strongly by differences in mean yield (fig. 1b, table 2b). However, the marked differences in interaction effects between groups with similar mean yields (groups IX and X, I and II, and V and VII) emphasised the influence of response pattern over environments, rather than mean performance, at the lower levels of classification. Environment groups III, VI and IX were characterised by relatively small interaction effects for all cultivar groups, indicating that their use as test environments provided little useful information contrasting cultivar performance. It was also apparent that certain environment groups (VII and VIII, fig. 5) contrasted markedly in the magnitude of interaction effects relative to that of the cultivar group mean.

The relative similarity of interaction effects for cultivar groups related by classification (fig. 1a) was apparent (fig. 5), and is not discussed further here. However, such presentation allows easy identification of those environment groups providing the greatest contrast in performance; for example, environment group IV and X for cultivar groups VIII, IX and X versus VI and VII, and environment group IV for cultivar group VI versus VII.

Environmental analysis may provide a rationale for such differential response patterns, and suggest testing strategies or breeding objectives; for example, the cultivar groups VI and VII contrasted markedly in environmental group IV, in part because of differential rust reaction. More subtle differentiation and environmental analysis is possible, but is not examined here.

4. Conclusions

Pattern analysis of the ISWYN 4 data matrix provided a useful means of studying the yield performance of cultivars over the large set of environments, summarising the data effectively in terms of similarities of mean yield or patterns of response. The presence of these similarities in the group structure paralleled the results of Mungomery et al. (1974) for seed yield and protein percentage in soybeans. Further, classification based on

seed yield generally reflected the origin, parentage and agronomic perform-
ance of the cultivars in a biologically sensible way. The extension of the
analysis to incorporate a classification of environments and the subsequent
two-way reduction of the data matrix, further diminished the complexity
and size of the problem under study, but still permitted meaningful inter-
pretation of the dynamic yield response patterns of cultivars across environ-
ments on a group basis.

Resolution of the sources of variation in the 10×10 table showed that
despite reduction in the size of the matrix by approximately 97 per cent,
a substantial proportion of the variation in yield response was retained.
Detailed analysis of the reduced matrix revealed real and useful information
reflecting the variation in yield response in the original data.

It must be conceded that information was lost in the process. The
classificatory method is based on a " minimum-variance " grouping model
(Wishart, 1969). It is elementary that variation will be partitioned
preferentially into the among-groups source, since that is how the method
works. However, it then becomes a question of what sort of variation,
and how much, is partitioned in this way. There is clear indication herein
that the two-way analysis adopted here has been effective in formulating
cultivar groups and environment groups useful from a plant breeding view-
point and potentially useful for physiological analysis. This warrants
further investigation, including more optimum means of truncation and
combination of the separate classifications to maintain maximum information.

Finally, the usefulness of the method in reducing the scale and com-
plexity of the analysis cannot be overstated. No claim can be made of
the uniqueness of the groups derived since a change in the numerical model
may produce a different classification (Williams, 1967). However, the
result was an objective organisation of the data upon which to begin study
of dynamic adaptation responses. In general, the group structure reflected
sensibly existing biological and plant breeding knowledge of this population,
and this suggested that the techniques may have application in study of
breeding populations.

5. References

ABOU-EL-FITTOUH, H. A., RAWLINGS, J. O., AND MILLER, P. A. 1969. Classification of environ-
ments to control genotype by environment interactions with an application to cotton.
Crop Sci., *9*, 135-140.

BURR, E. J. 1968. Cluster sorting with mixed character types. I. Standardization of
character values. *Aust. Comput. J.*, *1*, 97-99.

BURR, E. J. 1970. Cluster sorting with mixed character types. II. Fusion strategies. *Aust.
Comput. J.*, *2*, 98-103.

CHUANG-SHENG LIN, AND THOMPSON, B. 1975. An empirical method of grouping genotypes
based on a linear function of the genotype-environment interaction. *Heredity*, *34*, 255-
263.

FINLAY, K. W., AND WILKINSON, G. N. 1963. The analysis of adaptation in a plant breeding
programme. *Aust. J. agric. Res.*, *14*, 742-754.

FREEMAN, G. H., AND DOWKER, B. D. 1973. The analysis of variation between and within
genotypes and environments. *Heredity*, *30*, 97-109.

FREEMAN, G. H., AND PERKINS, J. M. 1971. Environmental and genotype-environmental components of variability. VIII. Relations between genotypes grown in different environments and measures of these environments. *Heredity, 27*, 15-23.

GOODCHILD, N. A., AND BOYD, W. J. R. 1975. Regional and temporal variations in wheat yield in Western Australia and their implications in plant breeding. *Aust. J. agric. Res., 26*, 209-217.

HORNER, T. W., AND FREY, K. J. 1957. Methods for determining natural areas for oat varietal recommendations. *Agron. J., 49*, 313-315.

MACKENZIE, D. R., MEXAS, A. G., FINLAY, K. W., AND BORLAUG, N. E. 1971. Results of the Fourth International Spring Wheat Yield Nursery, 1967-1968. *International Maize and Wheat Improvement Centre Research Bull. No. 18.*

MUNGOMERY, V. E., SHORTER, R., AND BYTH, D. E. 1974. Genotype × environment interactions and environmental adaptation. I. Pattern analysis—application to soya bean populations. *Aust. J. agric. Res., 25*, 59-72.

PERKINS, J. M., AND JINKS, J. L. 1968a. Environmental and genotype-environmental components of variability. III. Multiple lines and crosses. *Heredity, 23*, 339-356.

PERKINS, J. M., AND JINKS, J. L. 1968b. Environmental and genotype-environmental components of variability. IV. Non-linear interaction for multiple inbred lines. *Heredity, 23*, 525-535.

PERKINS, J. M. 1972. The principal component analysis of genotype-environmental interactions and physical measures of the environment. *Heredity, 29*, 51-70.

SUZUKI, S. 1968. Studies on the evaluation of strain adaptability. *Proc. 12th int. Congr. Genet.,* 270.

WILLIAMS, W. T., AND LAMBERT, J. M. 1961. Multivariate methods in plant ecology. III. Inverse association analysis. *J. Ecol., 49*, 717-729.

WILLIAMS, W. T. 1967. The computer botanist. *Aust. J. Sci., 29*, 266-271.

WILLIAMS, W. T. 1971. Principles of clustering. *A. Rev. Ecol. System. 2*, 303-326.

WISHART, D. 1969. Mode analysis: a generalization of nearest neighbour which reduces chaining effects. In *Numerical Taxonomy* (ed. A. J. Cole). Academic Press, London.