Heredity **56** (1986) 243–253

Received 29 July 1985

Some methods of analysing genotype–environment interaction

Brian Westcott

Centro Internacional de Mejoramiento de Maiz y Trigo (CIMMYT), Mexico.*

Methods of analysing genotype–environment interaction were extensively reviewed by Freeman (1973) and Hill (1975). A large number of papers involving such analyses have been published since then, some of them providing new methods, particularly of the multivariate type.

While making no pretensions to be a comprehensive review, the present paper attempts an examination of some of these new techniques. It also offers a critique of some established methods, while pointing out others which have been neglected. The methods considered include the linear regression approach and related stability parameters, cluster analysis, principal components analysis, geometrical methods and stochastic dominance.

In all these methods, environments are measured by the mean value of the genotypes grown in them: a case is made for research into the use of environmental variables.

LINEAR REGRESSION AND RELATED STABILITY PARAMETERS

In earlier times, methods of analysing genotype-environment interaction were associated with the linear regression approach. This was first introduced by Mooers (1921) and was given prominence by Yates and Cochran (1938), who used the mean performance of all genotypes grown in an environment as a suitable index of its productivity. The performance of each genotype was plotted against this index for each environment, and simple linear regression fitted by least squares to summarise the genotype's response, the mean regression slope being 1·0.

An identical technique was used by Mandel (1959) and Mandel and Lashoff (1959) to compare the results of tests on a number of materials at several different laboratories. The use of this approach as a basis for an analysis of variance and associated tests of hypotheses was discussed by Mandel (1961), who showed it to be an extension of Tukey's one degree of freedom for non-additivity (Tukey, 1949; Scheffe, 1959, pp 129–134). Translating Mandel's results to genotype-environment interaction analysis, if t is the number of genotypes, he removed a component with $t-1$ degrees of freedom from the interaction sum of squares. This component represents the sum of squares for heterogeneity of t regression slopes and Mandel proved that if the slopes are identical, it is distributed as chi-squared and is independent of environmental effects. Thus an F-test can be made of the existence of different genotype slopes, the null hypothesis being that the slopes are all the same.

Finlay and Wilkinson (1963) used the regression technique in examining the yield stability of various barley genotypes, although they claimed that better fits were obtained with log-transformed yields. In assessing stability, they considered that simply comparing regression slopes was not enough: the overall yield level of a genotype also had to be taken into account. The slope of the regression line for each genotype was, accordingly, plotted against its mean yield over environments. Genotypes with a slope near 1·0 and a high mean yield were regarded as being well adapted to all environments. As mean yield decreased, genotypes with high or low slopes were regarded as being specifically adapted to favourable or unfavourable environments, respectively.

Eberhart and Russell (1966) also used a linear regression approach. They regarded deviations

* Present address: Plant Breeding Institute, Trumpington, Cambridge CB2 2LQ, UK.

from the regression line as another important component of varietal stability, a stable variety being one with a regression line of slope near to 1·0 with a small sum of squared deviations. They seemed to be unaware, however, that this sum of squares is not independent of the slope (see, for example, Hardwick and Wood, 1972).

Tai (1971) used an essentially similar technique to Eberhart and Russell (1966); however, he employed an alternative method of fitting, using maximum likelihood estimation of a structural relationship where an appropriate joint normal distribution had been assumed.

All the authors mentioned so far assumed an expected linear response to environment. The relative stability approach of Hanson (1970) also assumed a specific function of environment as expected response, although this did not need to be linear.

Plaisted and Peterson (1959) computed an analysis of variance for every pair of genotypes so as to estimate the interaction variance for every combination of two genotypes. The mean of the interaction variances obtained for each genotype was used as an indicator of the contribution of that genotype to the total genotype-environment interaction.

Wricke (1962, 1964) proposed, in a similar analysis, that the contribution of a genotype to the interaction sum of squares in a two-way analysis of variance be used as a measure of its instability. This was criticised by Freeman and Perkins (1971) on the grounds that, given t genotypes and s environments, consideration of the alternative form of the interaction sum of squares in terms of totals shows that there is no way of dividing it into t groups. Further, the degrees of freedom are $(t-1)(s-1)$, and this number is not in general divisible by t.

St-Pierre, Klinck and Gauthier (1967) tried a completely different approach: they defined the percentage adaptability of a genotype to be that percentage of the environments being tested in which its performance was better than the mean performance of all genotypes.

Easton and Clements (1973) carried out an experiment on wheat genotypes in which differences in environment were, they claimed, due entirely to measurable amounts of nitrogen fertiliser. Some of the genotypes gave atypical (non-linear) responses to fertiliser amount. The linear regression parameters of Finlay and Wilkinson (1963) and the parameter of St-Pierre *et al.* (1967) failed to identify the aberrant genotypes. Although the parameters of Plaisted and Peterson (1959), Wricke (1962, 1964) and Eberhart and Russell (1966) gave poor stability measures to these genotypes, given the conditional nature of these measures (Knight, 1970; Witcombe and Whittington, 1971), they concluded that "caution should be exercised in describing as unstable those genotypes with high values of these parameters." As an illustration, they examined a subset of six entries out of the 25 originally analysed; five of the six were aberrant in that the highest nitrogen levels were supra-optimal, the other entry showing a sustained yield response up to the highest nitrogen level. This latter entry, when evaluated by means of the Eberhart-Russell parameter in the original set of 25 genotypes, was regarded as a very stable line, a low value being obtained. However, when the subset of six entries was considered as a separate group, this same entry appeared to be unstable, giving a comparatively large value when Eberhart-Russell parameters were calculated. Thus, a variety could have marked deviations from linear regression, not because it was inherently irregular, but because it showed a different response pattern from the majority of the group with which it was being compared. It is clear, then, that the stability measures mentioned above are far from satisfactory and, consequently, are not recommended.

Some authors, such as Baker (1969) and Byth, Eisemann and Delacy (1976) have criticised the linear regression approach on the grounds that in an analysis of variance, the proportion of the genotype-environment interaction sum of squares attributable to linear regression on the environmental indices may be very small.

Mungomery, Shorter and Byth (1974), in advocating the application of cluster analysis, emphasised "that the definition of an 'expected', or a 'suitable' or 'ideal' response is unnecessary, and that actual responses are examined". It is clear that not using such definitions avoids arguments about the relevance of particular expected response functions.

Hill (1975) attempted to counter the criticisms of the linear regression approach made by Easton and Clements (1973) and Mungomery *et al.* (1974): he stated that "there is not, as these workers have supposed, any *a priori* reason for believing that the interaction of the genotype and environment will be a linear function of the environment". To support this statement, he put forward the argument that "indeed, the *null hypothesis* (his italics) being tested by the linear regression analysis is that no relationship exists between the interaction of genotype and environment and the additive

environmental component apart from that due to chance variation". Unfortunately, Hill fails to appreciate the provisional role of the null hypothesis in statistical tests. Furthermore, it is clear that he misunderstands, not only the nature of the null hypothesis, but also the role of the alternative hypothesis, which indeed, in this case, asserts the existence of a linear relationship.

In the case of the linear regression approach to genotype-environment interaction analysis, even if a good counter-argument could be produced along these lines, it would be nugatory. Whether regarded as revealing a true linear relationship, or whether linearity is regarded as merely empirical, an assumption of convenience or however regarded, in every known instance where this methodology has been used, a linear regression was fitted to the data, usually by least squares. It is also a fact that such fits can be largely determined by one or two data points (see, for example, Weisberg, 1980, Chapter 5, and Daniel and Wood, 1980, Chapter 7), and particularly by points which are some distance from the centroid. This means that stability statistics of a variety may be unduly influenced by its performance in only one or two environments, and, to that extent, may be seriously misleading.

Consider, for example, data used by Yates and Cochran (1938), reproduced in table 1.

In this example, worse fits were obtained from log-transformed data and so the original values were used. The fitted regressions for three of the varieties are displayed in fig. 1; the fitted lines for the other two varieties were effectively parallel to that of variety 3 but were lower down and, as they did not cross any of the other lines, they were left out of the figure to improve its clarity. The regression coefficients and the residual mean squares for each variety are shown in table 2, and the same statistics are shown for the regressions fitted after, firstly, excluding the highest yielding site and, secondly, excluding the lowest yielding site.

It can be seen from this table that the slopes of the lines fitted for varieties 1 and 3 were relatively insensitive to omission of either of these points. The highest yielding site was 69·9 units above the overall mean. Leaving out this site from varieties 2 and 5 caused their fitted lines to swing towards the horizontal, whereas the line fitted for variety 4 became appreciably steeper. Omission of the lowest yielding site, however, caused the line for variety 4 to flatten. Although the yield of variety 4 at the lowest yielding site was 151·2, its predicted

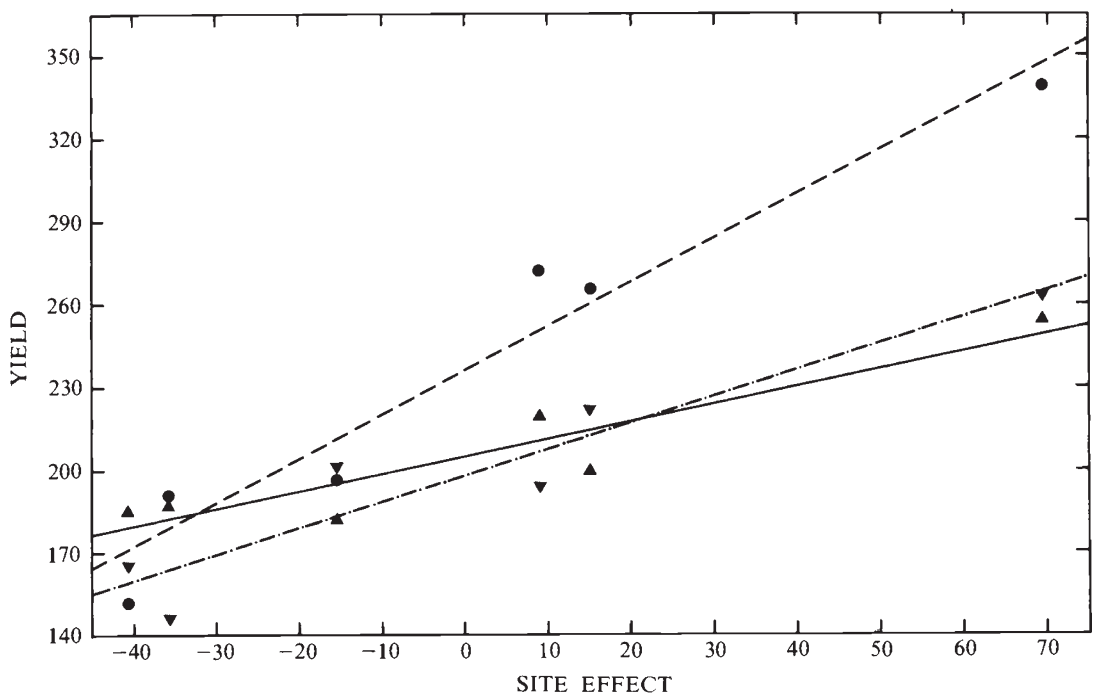

Figure 1 Regression lines for barley varieties 3, 4 and 5 with data points. Regressions are of yield against site effect (site mean − overall mean). The data values are taken from Yates and Cochran (1938) and are displayed in table 1. The regression coefficients are shown in table 2. Key for regression lines: Variety 3, · — · — ·; Variety 4, — — —; Variety 5, ——. Key for data points: Variety 3, ▼; Variety 4, ●; Variety 5, ▲.

Table 1 Data on five barley varieties grown in six sites, used by Yates and Cochran (1938)

Site	Variety 1	2	3	4	5	Mean	Site Effect
1	161·7	187·7	200·1	196·9	182·5	185·8	−16·4
2	247·0	257·5	262·9	339·2	253·8	272·1	69·9
3	185·4	182·4	194·9	271·2	219·2	210·6	8·4
4	218·7	183·3	220·2	266·3	200·5	217·8	15·6
5	165·3	138·9	165·8	151·2	184·4	161·1	−41·1
6	154·6	143·8	146·3	193·6	190·1	165·7	−36·5
Mean	188·8	182·3	198·4	236·4	205·1	202·2	

Table 2 Regression coefficients for the yields of the barley varieties in table 1 against site effects (residual mean squares in parentheses)

		Variety 1	2	3	4	5
All sites	a	187·8 (175·6)	182·3 (198·6)	198·4 (207·0)	236·4 (339·6)	205·1 (136·7)
	b	0·84	0·99	0·95	1·61	0·61
Excluding highest yielding site	a	177·1 (233·4)	167·2 (213·6)	185·5 (272·6)	215·8 (332·5)	195·3 (139·6)
	b	0·87	0·79	1·00	1·92	0·43
Excluding lowest yielding site	a	193·5 (168·6)	190·9 (260·5)	204·9 (255·2)	253·4 (260·2)	209·2 (171·3)
	b	0·93	0·96	0·99	1·46	0·65

value from the regression omitting this site was 181·3. This was remarkable, since the lowest yielding site was not so far from the overall mean (41·1 units). To summarise, it is clear that the fitted lines for three of the varieties were strongly influenced by one or two data points.

Because of such possibilities, the linear regression aproach to analysing genotype–environment interaction can not be regarded as trustworthy.

CLUSTER ANALYSIS

Abou-El-Fittouh, Rawling and Miller (1969) applied cluster analysis to classify locations used in cotton variety trials in the U.S.A. They used a distance coefficient and a correlation coefficient as dissimilarity measures and a variable group clustering strategy. However, they did not apply cluster analysis to classifying varieties.

Mungomery *et al.* (1974) employed an unstandardised squared Euclidean distance as dissimilarity measure, using an unweighted group average link clustering strategy (Sokal and Michener, 1958). The calculations were done using the general agglomerative algorithm of Lance and Williams (1967).

Byth *et al.* (1976) attempted to use clustering methods to analyse CIMMYT data from the 4th International Spring Wheat Nursery. They used a variance-standardised squared Euclidean distance as dissimilarity measure and an incremental sum of squares clustering strategy. The calculations for this method can also be done using the general agglomerative algorithm of Lance and Williams (1967).

Although Mungomery *et al.* (1974) and Byth *et al.* (1976) used clustering methods as a result of rejecting the linear regression approach, Lin and Thompson (1975) attempted to make use of cluster analysis to extend this approach. They took as the dissimilarity measure for a subset of t genotypes the variance ratio for testing the null hypothesis of a common linear regression line against the alternative hypothesis of t independent regression (Mandel, 1961). They proved that this dissimilarity measure equalled the mean of the measures for all possible pairs of genotypes in the subset. Thus, the index conformed to the conditions set by Sokal and Michener (1958) for use of their unweighted group average link strategy for clustering.

Lin and Thompson (1975) tried their method on the data of Yates and Cochran (1938) which was used earlier. As can be seen from tables 1 and 2, there are errors in their calculations: the mean for variety 5 should be 205·08 (not 209·08) and the slope for variety 1 should be 0·8440 (not 0·8840). However, their calculation of the dissimilarity matrix was correct and their method gave two distinct groups: the first group containing variety 4 alone and the second group containing the other four varieties.

This result need not be the most useful from a practical point of view, particularly if stability in low-yielding sites is of interest. Although the fitted line for variety 4 had the highest slope and this variety also had the highest mean yield, variety 5 performed much better than variety 4 at the lowest yielding site (see fig. 1 and table 1), where it was clearly the highest yielding variety. At the lowest but one yielding site (site 6), the two varieties performed about the same. Although variety 4 outyielded variety 5 in the lowest but two yielding site (site 1), it was not by much and variety 5 had a higher mean yield than variety 4 over the three lowest yielding sites, where it also exhibited remarkably similar yields (182·5, 190·1 and 184·4).

Lin (1982), dropping the linear regression approach but retaining the clustering strategy, proposed a dissimilarity measure for a pair of genotypes to be the squared distance between them adjusted for the average effects of genotypes. Using the Yates and Cochran data again, his calculated dissimilarity matrix was not quite correct (the measure for varieties 2 and 3 should be 87·5, not 88·5); this, however, did not affect the conclusion, which was to find the same two groups: variety 4 and the four other varieties.

Ramey and Rosielle (1983) modified Lin's method by minimising the total sum of squares for genotype-environment interaction within clusters at each fusion cycle.

Fox and Rosielle (1982) observed that "the scale of observations may have marked influences on determination of distances". They considered a diagram similar to fig. 2 in which four environments are plotted on axes representing the performance of two genotypes.

Fox and Rosielle (1982, p. 647) remarked: "Environment E1 ranks genotypes 1 and 2 equally, E3 ranks genotype 1 above genotype 2 while E4 ranks genotype 2 above genotype 1. In terms of squared Euclidean distance these 3 environments are equally separated and closer to each other than to E2 which, like E1, ranks the genotypes together. It is important to emphasise that though E1 and E2 are ranking genotypes identically they are the furthest apart in terms of distance because of the large difference in their environmental main effects".

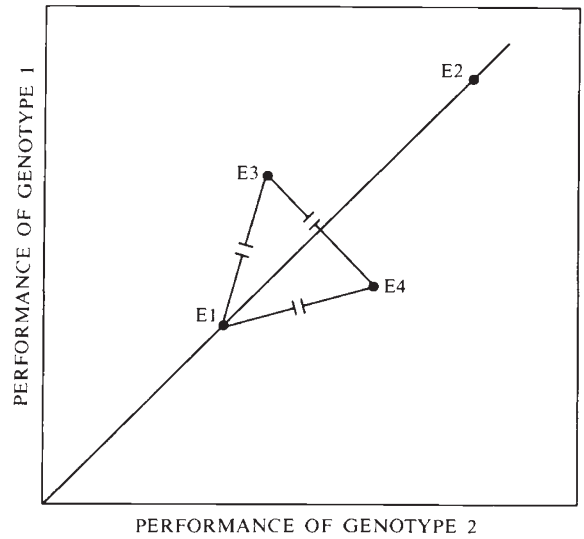

Figure 2 Four environments E1, E2, E3 and E4 plotted in two dimensions representing the performance of two genotypes.

After examining coded, ratio and standardised data, they concluded that standardisation was the most effective procedure for removing potentially dominant effects of environmental means on squared Euclidean distance measures of dissimilarity.

Johnson (1977) also employed cluster analysis as part of his examination of the yield and stability of a set of maize hybrids. He used what he described as weighted Euclidean distance as measure of similarity and maximum distance between clusters as clustering metric.

Ghaderi, Everson and Cress (1980) used cluster analysis to classify environments and genotypes in wheat, with a distance coefficient as dissimilarity measure and a complete link clustering strategy.

Ghaderi, Adams and Saettler (1982) also rejected the regression approach and used clustering methods to analyse data on beans. The correlation coefficient of genotypes over environments was taken as similarity measure and a complete link clustering strategy employed.

As we have seen, when using cluster analysis to elucidate genotype-environment interaction, many different dissimilarity measures and clustering strategies have been used. More are possible: in fact, Cormack (1971), in a review of

classification methods, listed 10 different dissimilarity measures and eight different clustering strategies. Particular choices of these can result in different cluster groups and the acceptance or rejection of any particular choice may be difficult to justify.

Another damning criticism of clustering methods made by Gordon (1981, Chapter 5) is that they "can force unwarranted structure on a data set, suggesting misleading results. It is a useful precaution also to employ methods of analysis which are not obliged to present their results in the form of a certain number of groups. As Cormack (1971, p. 340) remarks: 'When the data have not been forced into clusters, the observer can assess better whether clusters exist' ".

PRINCIPAL COMPONENTS ANALYSIS

Williams (1952) considered an equivalent model to the linear regression approach. He showed, in effect, that least squares estimation of the regression coefficients was equivalent to extracting the first principal component of the genotypic performances. The validity of the model can then be assessed by extracting further principal components or, more informally, by inspection of the residual correlation matrix after extracting the first component. Perkins and Jinks (1968), in examining plant height of some inbred lines of *Nicotiana rustica*, found that the genotypes could be divided into two groups on this basis, those genotypes showing positive residual correlations with each other forming one group, those showing negative correlations forming the other. A new regression analysis within each group showed a better fit to the data, although significant deviations still existed. The groups coincided with a well-established major gene difference, mop-head (m)/non-mop-head (M); however, in other cases, such a clear-cut biological explanation is seldom apparent.

Mandel (1971) considered the principal components approach further, using a multiplicative model. This method indicated the number of dimensions necessary to contain the genotypic variation and gave estimates of the corresponding coefficients, without, however, any prior knowledge of which factors these dimensions represented. When the deviations from regression on the environmental mean are substantial but no environmental variables have been measured, Hardwick and Wood (1972) suggested that Mandel's method may prove particularly valuable. Principal components analysis was used, among others, by Freeman and Dowker (1973), Goodchild and Boyd (1975) and Hill and Goodchild (1981). However, the chief difficulty with this approach is in the interpretation of the resulting principal components, which may not bear any obvious relationship to environmental conditions (see, for example, Silvey, 1982).

ENVIRONMENTAL VARIABLES

In the case when deviations are large and measurements of environmental variables are available, Williams' (1952) approach was extended by Hardwick and Wood (1972) and further elaborated by Wood (1976). However, environmental measurements are very seldom available. This is remarkable, considering that in many countries or states it is common experience to find large variation in crop yields: the maximum yield obtained for a crop usually exceeds greatly the mean yield over the whole area of the country or state. In view of such large variation, studies are needed to understand the relative importance of its causes, whether it be due to unchangeable differences in soils or climate, or to changeable differences in the management methods which are used or in the varieties which are grown: such studies are of even greater importance at the international level. Much effort has been devoted by research workers in all crops to gain this kind of understanding, although comparatively little has been done by international agricultural research centres. It is possible that such research could be aided by the measurement and appropriate analysis of environmental variables.

Many workers have emphasised the importance of genotype-environment interaction in both breeding and variety testing (see, for example, Arnold, 1972, Campbell and Lafever, 1977, Hamblin, Fisher and Ridings, 1980, Patterson and Silvey, 1980, Pederson and Rathjen, 1981, and Wright, 1976) and some have attempted to quantify those environmental factors responsible for observed interactions (see, for example, Beckett, 1982). However, even when environments for breeders' trials or variety recommendation trials are deliberately chosen to represent appropriate ecological zones and management systems, reports of trial results almost never include an account of the agency factors (those under the direct control of the farmer), in the form of the managements adopted and the treatments applied, the edaphic factors, in the form of comprehensive soil analyses

and the climatic factors, in the form of full meteorological data. Even when some of this information is available, it may not be utilised: further work may therefore be necessary to discover how this can best be done.

Several authors, in attempting to relate performance to particular environmental variables, have computed simple linear regressions of plant characters, such as yield, on to each environmental variable in turn (see, for example, Jones, 1979, and Beckett, 1982). When one independent variable dominates the situation, then, of course, the corresponding simple linear regression will show this correctly. However, when several independent variables are influential (even when acting additively), or when several of them are intercorrelated, such an analysis is no longer appropriate (see, for example, Daniel and Wood, 1980, section 4.1, and Weisberg, 1980, section 2.1).

Where many factors are involved, analysis using multiple regression has been suggested (see, for example, Knight, 1970, p. 230). In using this technique, more research is needed on which variables to include in the full regression equation. For example, in addition to agency, edaphic and climatic variables, the times at which the various physiological growth stages are reached may also need to be present. Another difficulty is that if many environmental variables are considered together with interaction and polynomial terms (in an attempt to cope with non-additivity), many environments will be needed in order to fit the regressions satisfactorily. However, as already mentioned, large yield variation is common and so, despite the cost, more environments are often needed in any case for the results to be applicable over an adequate range and so be unbiassed.

Climatic variables have been included in a multiple regression analysis by various authors, although they were used to predict overall yields of a crop rather than the yields of particular genotypes. Recent papers include those by Feyerherm and Paulsen (1981) on wheat and by Haun (1982) on maize. These authors used the stepwise regression technique for selection of a subset of regression variables, presumably because this is the only subset selection computer program available in some widely-used statistical packages. However, the stepwise regression procedure is known to have serious pitfalls (see, for example, Berk, 1978). None of the authors on this subject seem to be aware of more modern methods for selection of subsets in multiple regression (see, for example, Miller, 1984).

Clearly, much more needs to be done in this important area of research. However, since environmental data are not often available at present, the rest of this paper will be devoted to analyses in which an environment is measured by the mean yield of the genotypes grown in it.

GEOMETRICAL METHODS

In so-called geometrical methods, the basic aim is to represent each object (genotype or environment, in this case) by a point in some Euclidean space so that objects which are similar to one another are represented by points which are close together. The configuration of points is then investigated in an attempt to detect any underlying structure. Thus, unlike cluster analysis, no structure is forced on the data. Geometrical methods include principal coordinates analysis, non-metric multidimensional scaling, the biplot method and correspondence analysis.

Principal coordinates analysis was developed by Schoenberg (1935), Young and Householder (1938) and Torgerson (1952) and is sometimes referred to as "classical scaling". It was popularised by Gower (1966), who noted some of its links with other statistical procedures. Given a matrix whose elements are known (or assumed) to represent the squared distances between points in some Euclidean space, the procedure determines the coordinates of the points. If a set of coordinates is found, another set with the same interpoint distances can be found by translating the axes; this indeterminacy is removed by requiring the centroid of the set of points to lie at the origin of coordinates. The matrix of squared distances is first transformed into a symmetric inner product matrix. The requirement that the centroid of the final configuration of points should be at the origin is equivalent to centring this inner product matrix by subtracting from each element the means of the row and column in which it lies and adding the grand mean. The required coordinates are then found from the eigenvalues and eigenvectors of the centred matrix: the axes turn out to be just the principal axes.

If a similarity matrix, S, is used as starting point, the resulting principal coordinates reproduce a matrix of squared "distances" equal to $2(J-S)$, where J is a matrix of 1's of the same size as S (see, for example, Gordon, 1981, section 5.2). However, if S, when centred, is not non-negative definite, negative eigenvalues will be found and the resulting "distances" cannot be represented accurately in Euclidean space. In any

event, a representation can be found in a space of smaller dimension, r, by ignoring small eigenvalues. By analogy with principal components analysis, the adequacy of such a representation could be measured by the proportion of the sum of the eigenvalues due to the first r. However, this can be misleading if there are any large negative eigenvalues (indicating that the interpoint distances are not Euclidean).

The dimension of the Euclidean space used in an adequate representation is less than the number of genotypes, but may be quite large. As the human brain finds it difficult to imagine high-dimensional data sets, the first two coordinates can be represented in a two-dimensional scatter diagram. In order to discard as little as possible of the information contained in the remaining coordinates, metroglyphs (Anderson, 1960) can be used, where the information in extra dimensions is represented by rays radiating from the scatter point, the length of a ray depicting the absolute value of the coordinate which it represents. It needs some practice to assimilate the information in metroglyphs quickly: Anderson (1960) thought that the eye worked most efficiently with between three and seven rays.

Another insurance against discarding information needlessly is to superimpose the minimum spanning tree (Gower and Ross, 1969) of the whole configuration on the two-dimensional scatter diagram. Given n genotypes in a graph, a tree spanning these genotypes is a set of edges such that the genotypes form a connected graph which contains no loops. When each of the $n(n-1)/2$ possible edges is assigned a length equal to the dissimilarity between the corresponding pair of genotypes, a minimum spanning tree is a spanning tree for which the sum of the edge lengths is a minimum. Even if most of the variability is present in the two-dimensional representation, the positions of a few genotypes could still be inaccurate: genotypes which are nearer to each other could be less similar than some which are further apart. Superimposing the minimum spanning tree can identify genotypes whose relationships are distorted in the geometrical representation.

Principal coordinates analysis was used in the study of genotype–environment interaction by Mungomery *et al.* (1974), Shorter, Byth and Mungomery (1977) and Fox and Rosielle (1982). A squared Euclidean distance was used in these cases, but this was arbitrary, in the sense that it was not obviously appropriate and yet no justification of it was attempted. Configurations resulting from equally plausible metrics could perhaps be compared by Procrustes analysis (Gower, 1975; ten Berge, 1977) to examine the robustness of any observed structure.

The singular value decomposition of Eckart and Young (1936) is a well-known and useful procedure for the canonical decomposition of two-way tables and leads to an expression for the table matrix involving, in a product, although not in this order, a diagonal matrix of singular values and two other matrices which are orthogonal. The orthogonality means that, for all r, the first r columns of the matrices involved in the product give rise to the least squares fit to the table in r dimensions. The extension of this procedure to a three-way table leads to a decomposition involving, in a product, a diagonal matrix and three other matrices. Difficulties arise because these three other matrices are not necessarily orthogonal and the first r columns of the matrices involved will not, generally, give the least squares fit to the table in r dimensions. Carroll and Chang (1970), however, developed an individual scaling algorithm based on this procedure. Basford (1982) used a non-metric version of their algorithm (Takane, Young and de Leeuw, 1977) to match the weighted Euclidean distance between genotypes (referred to assumed underlying axes) to the observed distance between genotypes within an environment, where each genotype was represented by several attributes and the estimated weights reflected the relative importance of the underlying axes for the various environments. He used the data previously analysed by Mungomery *et al.* (1974) and Shorter *et al.* (1977) and claimed that, by considering six attributes simultaneously in a single analysis, information was revealed which was not forthcoming from separate analyses of two of these attributes.

Other non-metric multidimensional scaling methods (see, for example, Gordon, 1981, section 5.3) do not seem to have been used for analysing genotype–environment interaction and may repay further study.

The biplot method was proposed by Gabriel (1971). A two-dimensional approximation to an $r \times c$ table can be obtained from the least squares fit to the table in two dimensions. As mentioned earlier, this can be obtained by using the first two columns of the matrices involved in the singular value decomposition of the table matrix. The adequacy of this rank 2 approximation may be measured by that proportion of the sum of squared singular values which is due to the first two. Assuming that the approximation is good, the original table can thus be approximated by the product of

an $r \times 2$ matrix G and the transpose of a $c \times 2$ matrix H (see, for example, Gordon, 1981, section 5.6.1). For a table consisting of genotype-environment residuals, the rows of G and H are vectors with two coordinates which can be plotted in the same graph (the biplot) to represent separate components due to genotypes and environments. The distances between genotype points in the biplot correspond to actual Euclidean distances between genotypes if the table columns (corresponding to genotypes) have been relocated to zero mean. Kempton (1984) used this approach to form principal components biplots and also extended the previously mentioned plots of Finlay and Wilkinson (1963). By providing a simple way of using the biplot to give the expected response of a genotype in an environment, he made the information available from conventional regression and principal components analysis more accessible. However, the criticisms which were made earlier of these latter techniques also apply, by extension, to this use of biplots.

Correspondence analysis (reviewed by Hill, 1974), which would also involve the simultaneous representation of genotypes and environments, seems to have been relatively neglected in the context of genotype-environment interaction studies.

Further geometrical methods were described by Gnanadesikan (1977), but their use also seems to have been unexplored in this context.

STOCHASTIC DOMINANCE

Another technique which has been little used in analysing genotype-environment interaction and genotypic stability is the stochastic dominance procedure. New technologies in general and new crop varieties in particular may often be regarded by farmers as more risky than more traditional ones. Risk may then tend to act as an impediment to their adoption. Improved varieties that would be preferred by "risk-averse" farmers can be identified by the stochastic dominance procedure under certain assumptions. Anderson (1974, section 2) performed such an exercise on the sixth International Spring Wheat Yield Nursery administered by CIMMYT. He made three assumptions: firstly, that it makes sense to talk about a world (or large regional) probability distribution of wheat yields; secondly, that the selection of sites, cooperators, fields and growing and disease conditions is representative of the relevant world (or regional) domain of production and thirdly, that yield *per se* provides a reasonable surrogate for the argument of the average farmer's utility function. This assumption involves ignoring variable production costs. Anderson claimed that such an assumption "is unavoidable in processing international nursery data since each trial is in general grown under differing regimes of irrigation (where practised), tillage, fertilizers and weed and pest control that are most difficult to cost".

Menz (1980) used the cluster analysis of Byth *et al.* (1976) to analyse CIMMYT International Spring Wheat Yield Nurseries over five years and also, for comparison purposes, used stochastic dominance. He found that variety groups based on each method showed a considerable degree of coincidence and concluded that, while the stochastic dominance procedure is no panacea, it "appears to be useful as an adjunct to other analytical procedures for the interpretation of plant breeding or other trial data". It may prove instructive to perform comparative analyses of stochastic dominance with other techniques: indeed, the whole area needs further research. A FORTRAN program for the stochastic dominance method was provided by Anderson, Dillon and Hardaker (1977, Chapter 9).

DISCUSSION AND CONCLUSIONS

Methods involving the linear regression approach and related stability parameters cannot be recommended, nor can the defects of these methods be overcome by the use of either cluster analysis or principal components analysis.

As prophesied by Freeman (1973), new multivariate techniques have appeared in view of the available range of modern computers. However, continued research is needed on the use of environmental variables. In assessing series of trials, at present most or all of the effort is concentrated on measuring the genotypes, while little or none is devoted to measuring the environments. It is little wonder, then, that detailed, and hence more useful, knowledge of genotype-environment interaction is difficult to obtain. However, the question of how best to analyse environmental data has still not been properly addressed and more research is needed in an attempt to find a solution to this statistical problem.

In the cases where environmental variables are unlikely to be measured, further research is needed on relatively unexplored techniques like stochastic dominance, multidimensional scaling and correspondence analysis. Other geometrical methods

may also repay study in the context of genotype-environment interaction and stability studies.

Acknowledgement This work was carried out at CIMMYT while on sabbatical leave from the Plant Breeding Institute, Cambridge, England.

REFERENCES

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

ANDERSON, E. 1960. A semigraphical method for the analysis of complex problems. *Technometrics, 2*, 87-391.

ANDERSON, J. R. 1974. Risk-efficiency in the interpretation of agricultural production research. *Review of Marketing and Agricultural Economics, 42*, 131-184.

ANDERSON, J. R., DILLON, J. L. AND HARDAKER, J. B. 1977. *Agricultural Decision Analysis.* Iowa State University Press, Ames.

ARNOLD, M. H. 1972. Modal selection in BP52. *Cotton Growing Review, 49*, 107-125.

BAKER, R. J. 1969. Genotype-environment interactions in yields of wheat. *Canadian Journal of Plant Science, 49*, 743-751.

BASFORD, K. E. 1982. The use of multidimensional scaling in analysing multi-attribute genotype response across environments. *Australian Journal of Agricultural Research, 33*, 473-480.

BECKETT, J. L. 1982. Variety × environment interactions in sugar beet variety trials. *Journal of Agricultural Science, 98*, 425-435.

BERK, K. N. 1978. Comparing subset regression procedures. *Technometrics, 20*, 1-6.

BYTH, D. E., EISEMANN, R. L. AND DE LACY, I. H. 1976. Two-way pattern analysis of a large data set to evaluate genotype adaptation. *Heredity, 37*, 189-201.

CAMPBELL, L. G. AND LAFEVER, H. N. 1977. Cultivar × environment interactions in soft red winter wheat yield tests. *Crop Science, 17*, 604-608.

CARROLL, J. D. AND CHANG, J.-J. 1970. Analysis of individual differences in multidimensional scaling via an *n*-way generalisation of Eckart-Young decomposition. *Psychometrika, 35*, 283-319.

CORMACK, R. M. 1971. A review of classification (with discussion). *Journal of the Royal Statistical Society, A134*, 321-367.

DANIEL, C. AND WOOD, F. S. 1980. *Fitting Equations to Data.* Second Edition. Wiley, New York.

EASTON, H. S. AND CLEMENTS, R. J. 1973. The interaction of wheat genotypes with a specific factor of the environment. *Journal of Agricultural Science, 80*, 43-52.

EBERHART, S. A. AND RUSSELL, W. A. 1966. Stability parameters for comparing varieties. *Crop Science, 6*, 36-40.

ECKART, C. AND YOUNG, G. 1936. The approximation of one matrix by another of lower rank. *Psychometrika, 1*, 211-218.

FEYERHERM, A. M. AND PAULSEN, G. M. 1981. Development of a wheat yield prediction model. *Agronomy Journal, 73*, 277-282.

FINLAY, K. W. AND WILKINSON, G. N. 1963. The analysis of adaptation in a plant breeding programme. *Australian Journal of Agricultural Research, 14*, 742-754.

FOX, P. N. AND ROSIELLE, A. A. 1982. Reducing the influence of environmental main effects on pattern analysis of plant breeding environments. *Euphytica, 31*, 645-656.

FREEMAN, G. H. 1973. Statistical methods for the analysis of genotype-environment interactions. *Heredity, 31*, 339-354.

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.

GABRIEL, K. R. 1971. Biplot display of multivariate matrices with application to principal component analysis. *Biometrika, 58*, 453-467.

GHADERI, A., ADAMS, M. W. AND SAETTLER, A. W. 1982. Environmental response patterns in commercial classes of common bean (*Phaseolus vulgaris* L.). *Theoretical and Applied Genetics, 63*, 17-22.

GHADERI, A., EVERSON, E. H. AND CRESS, C. E. 1980. Classification of environments and genotypes in wheat. *Crop Science, 20*, 707-710.

GNANADESIKAN, R. 1977. *Methods for Statistical Data Analysis of Multivariate Observations.* Wiley, New York.

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. *Australian Journal of Agricultural Research, 26*, 209-217.

GORDON, A. D. 1981. *Classification.* Chapman and Hall, London.

GOWER, J. C. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika, 53*, 325-338.

GOWER, J. C. 1975. Generalized procrustes analysis. *Psychometrika, 40*, 33-51.

GOWER, J. C. AND ROSS, G. J. S. 1969. Minimum spanning trees and single linkage cluster analysis. *Applied Statistics, 18*, 54-64.

HAMBLIN, J., FISHER, H. M. AND RIDINGS, H. I. 1980. The choice of locality for plant breeding when selecting for high yield and general adaptation. *Euphytica, 29*, 161-168.

HANSON, W. D. 1970. Genotypic stability. *Theoretical and Applied Genetics, 40*, 226-231.

HARDWICK, R. C. AND WOOD, J. T. 1972. Regression methods for studying genotype-environment interactions. *Heredity, 28*, 290-222.

HAUN, J. R. 1982. Early prediction of corn yields from daily weather data and single predetermined seasonal constants. *Agricultural Meteorology, 27*, 191-207.

HILL, J. 1975. Genotype-environment interactions—a challenge for plant breeding. *Journal of Agricultural Science, 85*, 477-493.

HILL, J. AND GOODCHILD, N. A. 1981. Analysing environments for plant breeding purposes as exemplified by multivariate analysis of long term wheat yields. *Theoretical and Applied Genetics, 59*, 317-325.

HILL, M. O. 1974. Correspondence analysis: a neglected multivariate method. *Applied Statistics, 23*, 340-354.

JOHNSON, G. R. 1977. Analysis of genotypic similarity in terms of mean yield and stability of environmental response in a set of maize hybrids. *Crop Science, 17*, 837-842.

JONES, H. G. 1979. Effects of weather on spring barley yields in Britain. *Journal of the National Institute of Agricultural Botany, 15*, 24-33.

KEMPTON, R. A. 1984. The use of biplots in interpreting variety by environment interactions. *Journal of Agricultural Science, 103*, 123-135.

KNIGHT, R. 1970. The measurement and interpretation of genotype-environment interactions. *Euphytica, 19*, 225-235.

LANCE, G. N. AND WILLIAMS, W. T. 1967. A general theory of classificatory sorting strategies. I. Hierarchical systems. *Computer Journal, 9*, 373-380.

LIN, C. S. 1982. Grouping genotypes by a cluster method directly related to genotype-environment interaction mean square. *Theoretical and Applied Genetics, 62*, 277-280.

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

MANDEL, J. 1959. The measuring process. *Technometrics, 1*, 251-267.

MANDEL, J. 1961. Non-additivity in two-way analysis of variance. *Journal of the American Statistical Association, 56*, 878-888.

MANDEL, J. 1971. A new analysis of variance model for non-additive data. *Technometrics, 13*, 1-18.

MANDEL, J. AND LASHOF, T. W. 1959. The interlaboratory evaluation of testing methods. *Bulletin of the American Society for Testing Materials, 239*, 53-61.

MENZ K. M. 1980. A comparative analysis of wheat adaptation across international environments using stochastic dominance and pattern analysis. *Field Crops Research, 3*, 33-41.

MILLER, A. J. 1984. Selection of subsets of regression variables (with discussion). *Journal of the Royal Statistical Society, A147*, 389-425.

MOOERS, C. A. 1921. The agronomic placement of varieties. *Journal of the American Society of Agronomy, 13*, 337-352.

MUNGOMERY, V. E., SHORTER, R. AND BYTH, D. E. 1974. Genotype × environment interactions and environmental adaptation. I. Pattern analysis—application to soya bean populations. *Australian Journal of Agricultural Research, 25*, 59-72.

PATTERSON, H. D. AND SILVEY, V. 1980. Statutory and recommended list trials of crop varieties in the United Kingdon (with discussion). *Journal of the Royal Statistical Society, A143*, 219-252.

PEDERSON, D. G. AND RATHJEN, A. J. 1981. Choosing trial sites to maximize selection response for grain yield in spring wheat. *Australian Journal of Agricultural Research, 32*, 411-424.

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

PLAISTED, R. L. AND PETERSON, L. C. 1959. A technique for evaluating the ability of selections to yield consistently in different locations or seasons. *American Potato Journal, 36*, 381-385.

RAMEY, T. B. AND ROSIELLE, A. A. 1983. HASS cluster analysis: a new method of grouping genotypes or environments in plant breeding. *Theoretical and Applied Genetics, 66*, 131-133.

SCHEFFE, H. 1959. *The analysis of variance*. Wiley, New York.

SCHOENBERG, I. J. 1935. Remarks to Maurice Frechet's article 'Sur la définition axiomatique d'une classe d'espace distancies vectoriellement applicable sur l'espace d'Hilbert'. *Annals of Mathematics (2nd Series), 36*, 724-732.

SHORTER, R., BYTH, D. E. AND MUNGOMERY, V. E. 1977. Genotype × environment interactions and environmental adaptation. II. Assessment of environment contributions. *Australian Journal of Agricultural Research, 28*, 223-235.

SILVEY, V. 1982. Analysis of crop variety adaptation from performance trials in England and Wales. In *Proceedings of the XIth International Biometric Conference*, pp. 157-163, Toulouse.

SOKAL, R. R. AND MICHENER, C. D. 1958. A statistical model for evaluating systematic relationships. *University of Kansas Science Bulletin, 38*, 1409-1438.

ST-PIERRE, C. A., KLINCK, H. R. AND GAUTHIER, F. M. 1967. Early generation selection under different environments as it influences adaptation of barley. *Canadian Journal of Plant Science, 47*, 507-517.

TAI, G. C. C. 1971. Genotypic stability analysis and its application to potato regional trials. *Crop Science, 11*, 184-190.

TAKANE, Y., YOUNG, F. W. AND DE LEEUW, J. 1977. Nonmetric individual differences multidimensional scaling: an alternating least squares method with optimal scaling features. *Psychometrika, 42*, 7-67.

TEN BERGE, J. M. F. 1977. Orthogonal procrustes rotation for two or more matrices. *Psychometrika, 42*, 267-276.

TORGERSON, W. S. 1952. Multidimensional scaling. I. Theory and method. *Psychometrika, 17*, 401-419.

TUKEY, J. W. 1949. One degree of freedom for non-additivity. *Biometrics, 5*, 232-242.

WEISBERG, S. 1980. *Applied Linear Regression*. Wiley, New York.

WILLIAMS, E. J. 1952. The interpretation of interactions in factorial experiments. *Biometrika, 39*, 65-81.

WITCOMBE, J. R. AND WHITTINGTON, W. J. 1971. A study of the genotype by environmental interaction shown by germinating seeds of *Brassica napus*. *Heredity, 26*, 397-411.

WOOD, J. T. 1976. The use of environmental variables in the interpretation of genotype-environment interaction. *Heredity, 37*, 1-7.

WRICKE, G. 1962. Über eine methode zur erfassung der ökologischen streubreite in feldversuchen. *Zeitschrift für Pflanzenzüchtung, 47*, 92-96.

WRICKE, G. 1964. Zur berechnung der ökovalenz bei sommerweizen und hafer. *Zeitschrift für Pflanzenzüchtung, 52*, 127-138.

WRIGHT, A. J. 1976. The significance for breeding of linear regression analysis of genotype-environment interactions. *Heredity, 37*, 83-93.

YATES, F. AND COCHRAN, W. G. 1938. The analysis of groups of experiments. *Journal of Agricultural Science, 28*, 556-580.

YOUNG, G. AND HOUSEHOLDER, A. S. 1938. Discussion of a set of points in terms of mutual distances. *Psychometrika, 3*, 19-22.