ADVANCES IN AGRONOMY, VOL. 44

STATISTICAL ANALYSES OF MULTILOCATION TRIALS

Jose Crossa

Biometrics and Statistics Unit
International Maize and Wheat Improvement Center (CIMMYT)
06600 Mexico D. F., Mexico

I. INTRODUCTION

Multilocation trials play an important role in plant breeding and agronomic research. Data from such trials have three main agricultural objectives: (*a*) to accurately estimate and predict yield based on limited experimental data; (*b*) to determine yield stability and the pattern of response of genotypes or agronomic treatments across environments; and (*c*) to provide reliable guidance for selecting the best genotypes or agronomic treatments for planting in future years and at new sites.

Agronomists in particular use multilocation trials to compare combinations of agronomic factors, such as fertilizer levels and plant density, and

on this basis make recommendations for farmers. Breeders compare different improved genotypes to identify the superior ones.

Variation in yield responses among certain agricultural production alternatives (genotypes, agronomic treatments, and cropping systems), when evaluated in different environments, is known in the classical sense as interaction. This interaction is part of the behavior of the genotype or agronomic treatment and confounds its observed mean performance with its true value. Assessing any genotype or agronomic treatment without including its interaction with the environment is incomplete and thus limits the accuracy of yield estimates. Therefore, a significant portion of the resources of crop breeding and agronomy programs is devoted to determining this interaction through replicated multilocation trials.

Data collected in multilocation trials are intrinsically complex, having three fundamental aspects: (*a*) structural patterns; (*b*) nonstructural noise; and (*c*) relationships among genotypes, environments, and genotypes and environments considered jointly. Pattern implies that a number of genotypes respond to certain environments in a systematic, significant, and interpretable manner, whereas noise suggests that the responses are unpredictable and uninterpretable.

The function of the experimental design and statistical analyses of multilocation trials is to eliminate as much as possible of the unexplainable and extraneous variability (noise) contained in the data. Gauch (1988) mentions that statistical analysis of multilocation trials may have two different objectives. With the first, *postdictive criteria*, a statistical model is constructed for a data set, and success is measured in terms of the model's ability to fit this data set, with consideration of parsimony (reduced model with minimal degrees of freedom). With the second, *predictive criteria*, the data from a yield trial are partitioned into (*a*) data used to construct a model (modeling data) and (*b*) data used to validate the model (validation data). Success is measured in terms of the model's ability to fit the validation observations. Gauch points out that predictive assessment of a multilocation trial is expected to provide more accurate yield estimates than a model chosen by postdictive criteria.

Although many countries conduct extensive trials, little attention has been devoted to the most effective analyses of the data generated. This chapter reviews some of the conventional statistical analyses and stability methods for yield trials. Statistical and biological limitations are discussed. New methodologies for analyzing multilocation trials as well as multivariate analyses for assessing yield stability are presented.

II. CONVENTIONAL ANALYSIS OF VARIANCE

Consider a trial in which the yield of G genotypes is measured in E environments each with R replicates. The classic model for analyzing the total yield variation contained in GER observations is the analysis of variance (Fisher, 1918, 1925). The within-environment residual mean square measures the error in estimating the genotype means due to differences in soil fertility and other factors, such as shading and competition from one plot to another.

After removing the replicate effect when combining the data, the GE observations are partitioned into two sources: (*a*) additive main effects for genotypes and environments and (*b*) nonadditive effects due to genotype–environment interaction.

The analysis of variance of the combined data expresses the observed (Y_{ij}) mean yield of the i^{th} genotype at the j^{th} environment as

$$Y_{ij} = \mu + G_i + E_j + GE_{ij} + e_{ij} \tag{1}$$

where μ is the general mean; G_i, E_j, and GE_{ij} represent the effect of the genotype, environment, and genotype–environment interaction, respectively; and e_{ij} is the average of the random errors associated with the r^{th} plot that receives the i^{th} genotype in the j^{th} environment.

The nonadditivity interaction as defined in (1) implies that the expected value of the i^{th} genotype in the j^{th} environment (Y_{ij}) depends not only on the levels of G and E separately but also on the particular combination of levels of G and E.

The presence of a significant genotype–environment interaction complicates interpretation of the results. Freeman (1985) mentioned two intuitive ways of overcoming this problem. The first is to examine the data and see if the interaction is due to one observed outlier. A high residual might be due to a recording error. In that case the outlier may be replaced by its expected value, using a model such as (1) without considering the interaction term. One can then determine whether the interaction in model (1) disappears. The second approach is to find a different scale on which the results can be expressed. Frequently, the logarithm transformation of the data removes interaction without rank changes and achieves both equality of variance within treatments and normality of residuals. However, crossover or rank change interactions cannot be removed by log transformation.

When genotype–environment interaction exists and is not due to one outlier and cannot be removed by a suitable transformation, it should be assumed that an underlying model such as (1) is representing the data.

A. Limitations

Hill (1975) outlined the advantages of analysis of variance in obtaining unbiased estimates of genetic and genotype–environment interaction variance components. He failed, however, to recognize its limitations in describing further structures in the nonadditive component.

In practice, if there is little variation in residual mean squares from one environment to another and the experiments are of equal size, the pooled error variance is found by averaging the residual mean squares of all environments. This combined experimental error is used to test the null hypothesis that the genotype differences are the same in all environments.

This analysis is open to criticism, however, if error variances are heterogeneous. The F-test of the genotype–environment interaction mean squares against the pooled error variance is biased toward significant results. Cochran and Cox (1957 p. 554) point out that in agricultural experimentation, loss of sensitivity of the F-test is equivalent to discarding 10% to 20% of the data.

A correct test of significance, by weighting each genotype mean by the inverse of its estimated variance, has been used by Yates and Cochran (1938) and Cochran and Cox (1957). The weighted analysis gives less weight to environments that have a high residual mean square. The sum of squares for genotype–environment interaction is inflated by errors in the weights; however, it can be reduced to a quantity that is distributed approximately as chi-square.

The disadvantage of weighted analysis is that weights may be correlated to environment yield responses (with high-yielding environments showing higher error variance and low-yielding sites presenting lower error variances). This would mask the true performance of some genotypes in certain environments. It is recommended that less weight be assigned to agricultural environments of less importance (Patterson and Silvey, 1980).

The genotype mean square is influenced by the pooled error variance, the variance of genotype–environment interaction, and the variance among true genotype means. The ratio of the genotype mean square to the genotype–environment interaction provides a test for the null hypothesis that there are no differences among the true genotype means. A criticism of this F-test is that, if the interaction variance is not the same for all of its components (some components of the interaction are much higher than others), too many significant results are obtained. In a trial of genotypes, this may occur when some genotypes are relatively unresponsive to a change in environment whereas others have a marked response. It is

recommended that the genotypes be further partitioned into a set of orthogonal components and that all of these components be tested for their interaction with the environment (Cochran and Cox, 1957).

Often, the analysis of variance test of the significance of the genotype–environment interaction declares it not significant when in fact it is agronomically or genetically important and its sum of squares accounts for a large proportion of the total variation (Zobel *et al.*, 1988). This may occur because the interaction contains a large number of degrees of freedom.

One of the main deficiencies of the combined analysis of variance of multilocation yield trials is that it does not explore any underlying structure within the observed nonadditivity (genotype–environment interaction). Analysis of variance fails to determine the pattern of response of genotypes and environments. The valuable information contained in (*G*-1) (*E*-1) degrees of freedom is practically wasted if no further analysis is done.

Since the nonadditive structure of a data matrix has a nonrandom (pattern) and random (noise) component, the advantages of the additive model are lost if the pattern component of the nonadditive structure is not further partitioned into functions of one variable each. Williams (1952), Mandel (1961, 1969, 1971), and Gollob (1968) have delineated methods for analyzing and interpreting two-way tables with interaction. They show that the sum of squares for interaction can be further partitioned in multiplicative components related to eigenvalues. The interaction part of Eq (1) can be expressed in the form

$$GE_{ij} = k_1 v_{1i} s_{1j} + k_2 v_{2i} s_{2j} + k_3 v_{3i} s_{3j} + \ldots \tag{2}$$

then $GE_{ij} = \sum_{n-1}^{h} k_n v_{ni} s_{nj}$, and

$$Y_{ij} = \mu + G_i + E_j + (\sum_{n-1}^{h} k_n v_{ni} s_{nj}) + e_{ij} \tag{3}$$

where k_n is the singular value of the n^{th} axis (k_n^2 is the eigenvalue), v_{ni} is the eigenvector of the i^{th} genotype for the n^{th} axis, s_{nj} is the eigenvector of the j^{th} environment for the n^{th} axis, and $\sum_{n-1}^{h} v_{ni} = \sum_{n-1}^{h} s_{nj} = 1$. This result links the analysis of variance with the principal components analysis. This analysis is called Additive Main effect and Multiplicative Interaction (AMMI) and is considered in this chapter in the discussion of nonconventional analysis of variance.

B. Components of Variance

Analysis of variance of multilocation trials is useful for estimating variance components related to different sources of variation, including genotypes and genotype–environment interaction. Variance components have been widely used in genetics and plant breeding (Comstock and Moll, 1964; Cockerham, 1964; Gardner, 1964).

In general, variance component methodology is important in multilocation trials, since errors in measuring the yield performance of a genotype arise largely from genotype–environment interaction. Therefore, knowledge of the size of this interaction is required to (*a*) obtain efficient estimates of genotype effects and (*b*) determine optimum resource allocations, that is, the number of plots and locations to be included in future trials.

In a breeding program, variance component methodology is used to measure genetic variability and to estimate the heritability and predicted gain of a trait under selection.

For balanced multilocation trials, that is, those with the same number of experimental units (genotypes or agronomic treatments) observed per site, estimation of the variance component is accomplished using the analysis of variance method. Each of the mean squares is known to estimate a linear function of the variance components defined in the model. These linear functions are called expected mean squares. By solving simultaneous equations, linear functions of the mean squares can be obtained that estimate each variance component. This method is limited to balanced data, and its main advantage is that it produces the best unbiased point estimators of the variance components (Graybill and Hultiquist, 1961).

However, there is nothing intrinsic in the method to prevent negative estimates. The interpretation of a negative estimate of a nonnegative parameter creates controversy. In practice, the negative estimate can be accepted and used or a value of zero can be used instead. Thompson (1961, 1962) gives some rules for ignoring the negative component and reestimating the others.

Genetic and genetic–environment variance components can be estimated by the maximum likelihood method. The disadvantage of these estimators, in the case of balanced data, is that they are biased downward (Patterson and Thomson, 1975). This problem can be overcome by using the restricted maximum likelihood (REML) method (Robinson, 1987). This method is analogous to the analysis of variance, and both produce identical estimators for balanced data.

For unbalanced experiments, including incomplete block designs, estimating the expected mean squares can be difficult, and the analysis of variance method for variance component estimation is not necessarily a

desirable approach. Unbalancedness in multilocation trials can have many different causes, including shortage of seed, testing of some genotypes only at some locations (or in some years), and the addition of new genotypes to the trial system and discarding of others. General methods for calculating variance components in nonorthogonal data by means of REML analysis have been developed by Patterson and Thomson (1971, 1975).

III. JOINT LINEAR REGRESSION

Another important model for analyzing and interpreting the nonadditive structure (interaction) of two-way classification data is the joint linear regression method. This approach has been extensively used in genetics, plant breeding, and agronomy for determining yield stability of different genotypes or agronomic treatments.

The genotype–environment interaction is partitioned into a component due to linear regression (b_i) of the i^{th} genotype on environmental mean and a deviation (d_{ij}):

$$(GE)_{ij} = b_i E_j + d_{ij} \tag{4}$$

and

$$Y_{ij} = \mu + G_i + E_j + (b_i E_j + d_{ij}) + e_{ij} \tag{5}$$

This model uses the marginal means of the environments as independent variables in the regression analysis and restricts the interaction to a multiplicative form. It was first proposed by Yates and Cochran (1938) in their analysis of a barley yield trial. The method divides the (G-1) (E-1) df for interaction into G-1 df for heterogeneity among genotype regressions and the remainder (G-1) (E-2) for deviation. Further details about interaction are obtained by regressing the performance of each genotype on the environmental means. Eberhart and Russell (1966) proposed pooling the sum of squares for environments and genotype–environment interactions and subdividing it into a linear effect between environments (with 1 df), a linear effect for genotype–environment (with G-1 df), and a deviation from regression for each genotype (with E-2 df).

Thus, not until the 1960s was it possible to solve the intractable problem of genotype by environment interaction by means of a regression approach. Part of the genotype's performance across environments or *genotype stability* is expressed in terms of three empirical parameters: the mean performance, the slope of the regression line, and the sum of squares

deviation from regression. Although joint regression has been principally used for assessing the yield stability of genotypes in a plant breeding program, it may also be used for agronomic treatments. It has also been used to estimate biometrical genetical parameters (Bucio Alanis *et al.*, 1969).

When attention is focused on environments, the converse analysis may be performed by regressing each environment's yields on the genotype means (Fox and Rathjen, 1981).

Freeman (1973), Hill (1975), and Westcott (1986) have provided comprehensive reviews of regression methods for studying genotype–environment interactions. Several statistical and biological limitations of the regression method should be noted.

A. Statistical and Biological Limitations

The first statistical criticism of regression analysis is that the genotype mean (x variable) is not independent of the marginal means of the environments (y variable). Regressing one set of variables on another that is not independent violates one of the assumptions of regression analysis (Freeman and Perkins, 1971; Freeman, 1973). This interdependence may be a major problem for small numbers, but not when the number of genotypes is large (say 15 to 20). If the standard set for stable yield is based on very few genotypes (say 4), each estimated stability coefficient involves regressing one genotype on an average to which it contributes one-fourth.

Biological and algebraic interdependency also exists between slopes and sums of squares due to deviations from regression. Hardwick and Wood (1972) concluded that this is a necessary adjunct of the line-fitting procedure.

The second statistical limitation is that errors associated with the slopes of genotypes are not statistically independent, because the sum of squares for deviation, with $(G\text{-}1)$ $(E\text{-}2)$ df, cannot be subdivided orthogonally among the G genotypes.

The third statistical problem with regression analysis is that it assumes a linear relationship between interaction and environmental means. When this assumption is violated,the effectiveness of the analysis is reduced, and results may be misleading (Mungomery *et al.*, 1974). In fact, the analysis requires that a high proportion of the genotype by environment effects should be attributable to linear regression (Perkins, 1972; Freeman, 1973).

A nonlinear relationship between interaction and environmental effects has been proposed by Pooni and Jinks (1980), and Hill and Baylor (1983) have used an orthogonal contrast analysis of variance that subdivides the

variation over environments (years and sites) for each entry into sources due to environment linear and quadratic effects.

Freeman and Perkins (1971) have criticized Eberhart and Russell's partitioning of the pooled sum of squares for environments and genotype–environment interaction, noting that the 1 df sum of squares for the linear component between environments is the same as the total sum of squares for environments with E-1 df.

A major biological problem with regressing genotype means on environmental means arises when only a few very low or very high yielding sites are included in the analysis. The fit of a genotype may be largely determined by its performance in those few extreme environments, with possibly misleading results (Hill and Baylor, 1983; Westcott, 1986). An example is presented by Westcott (1986) from the barley yield trial data of Yates and Cochran (1938), in which regression coefficients for the yield of the genotypes were calculated for all of the trials and for all except the highest- and lowest-yielding site (Table I). The exclusion of one extreme point had a strong influence on the slope of genotypes 2, 4, and 5, even though the lowest-yielding site was only 41.1 units apart from the grand mean.

Crossa (1988) found that excluding 1 very low yielding site out of 20 or 1 high-yielding site out of 17 influenced the estimates of slopes and deviations from regression for some genotypes. The performance of some genotypes at only one site overshadowed their general response at most of the other sites. The author concluded that regression analysis should be used with caution when the data set includes results from a few extremely low or high yielding locations.

Another biological criticism of the regression method is that the relative stability of any two genotypes depends not only on the particular set of locations included in the analysis but also on the other genotypes that are included in the regression calculation. It has been shown that the stability of a genotype depends on the mean performance of the group with which

Table I

Regression Coefficients of Five Barley Genotypes[a]

Genotype	All sites included	Excluding highest-yielding site	Excluding lowest-yielding site
1	0.84	0.87	0.93
2	0.99	0.79	0.96
3	0.95	1.00	0.99
4	1.61	1.92	1.46
5	0.61	0.43	0.65

[a] From Wescott (1986).

that entry is being compared (Knight, 1970; Witcombe and Wittington, 1971; Mead *et al.*, 1986; Crossa, 1988). Furthermore, it is possible that the ranking of two genotypes' stability coefficients may be reversed when they are compared with two other sets of genotypes.

The stability of a particular genotype is unsatisfactory if its response is different from the mean response of the group with which it is being compared (Easton and Clements, 1973). This can be seen in Table II, which gives the deviations from regression for 6 entries, considered (*a*) as members of the original set of 25 entries and (*b*) as an isolated group. It can be seen that the entry Raven × 65 RN 85 was originally a stable line (604) but appears unstable when considered as a member of the subset of 6 entries (6078).

Crossa (1988) estimated Eberhart and Russell's stability parameters for genotypes considered, along with others, as a subset of the original group of 27 entries. When 7 genotypes were considered in isolation, deviation from regression of some genotypes changed drastically. This result confirmed that the yield stability of one entry, as determined by regression, varied according to the average response of the rest of the group. The author also pointed out that, in trying to determine which genotype is superior, plant breeders have difficulty reaching a compromise between the yield mean, slope, and deviation from regression, because the genotype's response to environments is intrinsically multivariate and regression tries to transform it into a univariate problem (Lin *et al.*, 1986).

An alternative approach to overcoming the dependency present in the regression analysis—one especially suitable for agronomic treatments—is to consider the joint distribution of a pair of treatments, say A and B, and to regress the yield differences (A-B) on the mean yield (A−B/2) (R. Mead,

Table II

Deviation from Regression of 6 Genotypes when Considered as Members of the Original Group of 25 Genotypes and as an Isolated Group[a]

Genotype	Member of 25 genotypes	Isolated group
Aotea	8,942	1,347
Arawa	5,267	4,839
1020, 01	9,645	3,486
Nadadores 63	8,235	6,425
Hi-61 × Aotea	12,402	9,543
Raven × 65 RN 85	604	6,078

[a] From Easton and Clements (1973).

personal communication). Assuming an approximately linear relationship between both treatments, a positive slope would indicate that B is more stable than A.

If a large percentage of the genotype–environment interaction sum of squares can be explained by the heterogeneity of regressions, then the joint regression method can efficiently describe the pattern of adaptation in the response of genotypes. However, Baker (1969), Byth *et al.* (1976), Eagles and Frey (1977), and Shorter (1981) reported that a very small portion (9–16%) of the genotype–environment sum of squares is attributable to linear regression in various situations. Shorter (1981) concluded that, if this is the most common situation in field crops, the joint regression method of analysis is of little value.

Moll *et al.* (1978) studied the interaction of several populations of maize with environments, using the Eberhart and Russell procedure with the modification of Mather and Caligari (1974). The interaction sum of squares was divided into two parts: differences among genotypes in their variability among environments and differences in correlations among pairs of entries. Moll *et al.* found that heterogeneity of regression coefficients among genotypes may be due to heterogeneity of variance.

Using results from Bruckner and Frohberg (1987) on kernel weight of 20 spring wheats tested in 15 environments, Baker (1988a) pointed out that the high correlation between regression coefficients and estimated variances over environments suggests that heterogeneity of slopes is explained by heterogeneity of variance.

B. Other Measurements of Yield Stability

Other methods of determining genotype stability are based on genotype–environment interaction effects and are briefly examined next.

Plaisted and Peterson (1959) computed combined analysis of variance for each pair of genotypes included in a trial. The variance component of the genotype–environment interaction is estimated for each pair and each genotype. The genotype with the smallest mean variance component contributes less to the total interaction and is considered the most stable.

Wricke (1962, 1964) defined the concept of ecovalence as the contribution of each genotype to the genotype–environment interaction sum of squares. The ecovalence (W_i) or stability of the i^{th} genotype is its interaction with environments, squared and summed across environments, and expressed as

$$W_i = [\bar{Y}_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} - \bar{Y}_{..}]^2 \tag{6}$$

where $\bar{Y}_{ij}$ is the mean performance of genotype i in the j^{th} environment and

$\bar{Y}_{i.}$ and $\bar{Y}_{.j}$ are the genotype and environment mean deviations, respectively, and $\bar{Y}_{..}$ is the overall mean. Accordingly, genotypes with low ecovalance have smaller fluctuations from the mean across different environments and are therefore more stable.

Shukla (1972) defined the stability variance of genotype i as its variance across environments after the main effects of environmental means have been removed. Since the genotype main effect is constant, the stability variance is based on the residual ($GE_{ij} + e_{ij}$) matrix.

Lin and Binns (1988) defined the superiority measure (P_i) of the i^{th} genotype as the mean square of distance between the i^{th} genotype and the genotype with maximum response as

$$P_i = [n(Y_{i.} - M_{.})^2 + (Y_{ij} - Y_{i.} + M_{j.} + M_{..})^2]/2n \quad (7)$$

where Y_{ij} is the average response of the i^{th} genotype in the j^{th} environment, $Y_{i.}$ is the mean deviation of genotype i, M_j is the genotype with maximum response among all genotypes in the j^{th} location, and n is the number of locations. The first term of the equation represents the genotype sum of squares, and the second term is the genotype–environment sum of squares. The smaller the value of P_i, the less its distance to the genotype with maximum yield and the better the genotype. A pairwise genotype–environment interaction mean square between the maximum and each genotype is also calculated. This method is similar to that of Plaisted and Peterson (1959), except that (*a*) the stability statistics are based on both the average genotypic effects and genotype–environment interaction effects and (*b*) each genotype is compared only with the one maximum response at each environment.

Lin *et al.* (1986) reviewed nine stability measurements frequently used in biological research and grouped them into four categories, depending on whether they are based on the deviations from the average genotype effect or on the genotype–environment interaction effects. The authors defined three different parametric concepts of stability statistics. A genotype is stable if (*a*) its among-environment variance is small; (*b*) its response to environment is parallel to the mean response of all genotypes included in the trial; and (*c*) the residual mean square from the regression model on the environmental index is small. Stability methods based on the genotype–environment interaction sum of squares correspond to type *b*, whereas the Eberhart and Russell method is type *c*. As the authors point out, these parametric concepts of stability are relatively simple and address only some aspects of stability without giving an overall picture of the genotype's response. A genotype may be considered to have type *b* stability and simultaneously type *c* instability. Since a genotype's response to environment is multivariate, Lin *et al.* (1986) proposed using cluster analysis to classify genotypes.

C. Stability, Risk, and Economic Analysis

One of the main aims of breeders and agronomists is to recommend to farmers new agriculture production alternatives (genotypes, agronomic treatments, and cropping systems) that are stable under different environmental conditions and minimize the risk of falling below a certain yield level.

Subsistence farmers using low levels of inputs in unfavorable environments tend to be reluctant to adopt new technology. Given the uncertainty of their circumstances, these farmers' main concern is not so much to increase production as to avert catastrophe.

Conventional regression analysis considers only three components of stability: (*a*) response to changing environment (regression coefficient); (*b*) yield variability; and (*c*) mean yield level. However, this assessment of stability is incomplete and inappropriate unless it is related to risk probability (Barah *et al.*, 1981; Mead *et al.*, 1986).

The concept of risk efficiency of a particular genotype involves a tradeoff between its average yield and variance. A genotype is risk efficient if no other genotype has the same yield with lower variance or the same variance with higher mean yield. The mean–standard deviation analysis provides a method in which the benefits of reduced yield variability are measured against loss in yield (Binswanger and Barah, 1980). This analysis requires that the breeder or farmer specify how mean and standard deviation are "trade off." Mean–standard deviation analysis translates the stability parameters of a genotype (slope and deviation from regression) into economic benefit for the farmer.

Mean–standard deviation analysis and regression analysis were compared on yield data of pearl millet genotypes tested for 5 years in India and Pakistan (Witcombe, 1988). The results of both analyses were similar in all the environments and the standard deviation predicted well the values of deviation from regression.

In comparing the risk stability of two cropping systems (two crops versus one crop), Mead *et al.* (1986) define risk as the probability of yield falling below certain prespecified levels. The authors describe a general method of expressing stability related to risk probability by adjusting a bivariate distribution to the data and then estimating a theoretical continuous risk curve. The method can be used for assessing the risk stability of any two genotypes or agronomic treatment.

The stochastic dominance procedure (Anderson, 1974; Menz, 1980) ranks different agricultural alternatives according to farmers' risk aversion and selects those with high risk efficiency. It is assumed that each alternative has a probability distribution of yield, $f(i)$, and therefore a cumulative distribution function, $F(i)$. Then, the $f(i)$ is said to dominate $f(j)$ by first-

degree stochastic dominance if all the values of the yield distribution of alternative *i* are greater than those of alternative *j*. Second- and third-degree stochastic dominance appear when the distributions of yields are not easily separated. The importance of stochastic dominance is, unlike mean–standard deviation analysis, that the breeder or farmer does not have to specify the trade-off between average yield and variance.

Under yield uncertainty a major problem is how to make trade-offs among conventional stability statistics, for example, mean yield, slope, and deviation from regression. The central concept in safety-first decision strategies is the assumption that breeders and farmers prefer genotypes with a small chance of producing small yields. Eskridge (1990) addressed this issue by developing safety-first selection indices based on four different stability approaches: (*a*) the variance of a genotype across environments (EV); (*b*) the regression coefficient used by Finlay and Wilkinson (1963) (FW); (*c*) the stability variance of Shukla (1972) (SH); and (*d*) the regression coefficient and deviation from regression defined by Eberhart and Russell (1966) (ER). The rank correlations between the mean genotype rankings and the four selection indices show that FW, SH, and ER produce similar rankings (>0.65). The mean ranking, on the other hand, is poorly correlated with EV (0.152) and only moderately correlated with FW, SH, and ER (0.45 < rank correlations < 0.7). Only one genotype was ranked near the bottom for all indices. A safety-first index is useful for selecting genotypes in the presence of genotype–environment interaction (Eskridge, 1990) because: (*a*) it weights the importance of stability relative to yield; (*b*) it can be used with different types of univariate stability statistics for any trait; and (*c*) it is more likely to identify superior varieties when high costs are associated with low yields.

IV. CROSSOVER INTERACTIONS

Interaction in the classic sense exists because the responses of genotypes are not parallel over all environments. In agricultural production, changes in a genotype's rank from one environment to another are important. These are called crossovers or qualitative interactions, in contrast to noncrossovers or quantitative interactions (Baker, 1988b,c; Gail and Simon, 1985). With a qualitative interaction, genotype differences vary in direction among environments, whereas with quantitative interactions, genotypic differences change in magnitude but not in direction among environments. If significant qualitative interactions occur, subsets of genotypes are to be recommended only for certain environments, whereas

with quantitative interactions the genotypes with superior means can be used in all environments. Therefore, it is important to test for crossover interactions (Baker, 1988b).

Consider, for example, genotypes A and B tested in environments 1 and 2 Let $\overline{Y}_{A1}$ and $\overline{Y}_{A2}$ be the means of genotype A in environments 1 and 2, respectively, and $\overline{Y}_{B1}$ and $\overline{Y}_{B2}$ be the means of genotype B in environments 1 and 2, respectively. The genotype effects in each environment are defined as

$$d_1 = \overline{Y}_{A1} - \overline{Y}_{B1}$$
$$d_2 = \overline{Y}_{A2} - \overline{Y}_{B2}$$

No interaction and all types of interactions can be illustrated in the space of genotype effects by plotting d_2 and d_1 (Fig. 1).

Then, the line where $d_1 = d_2$ represents the situation in which there are differences in genotype means but not in genotype–environment interac-

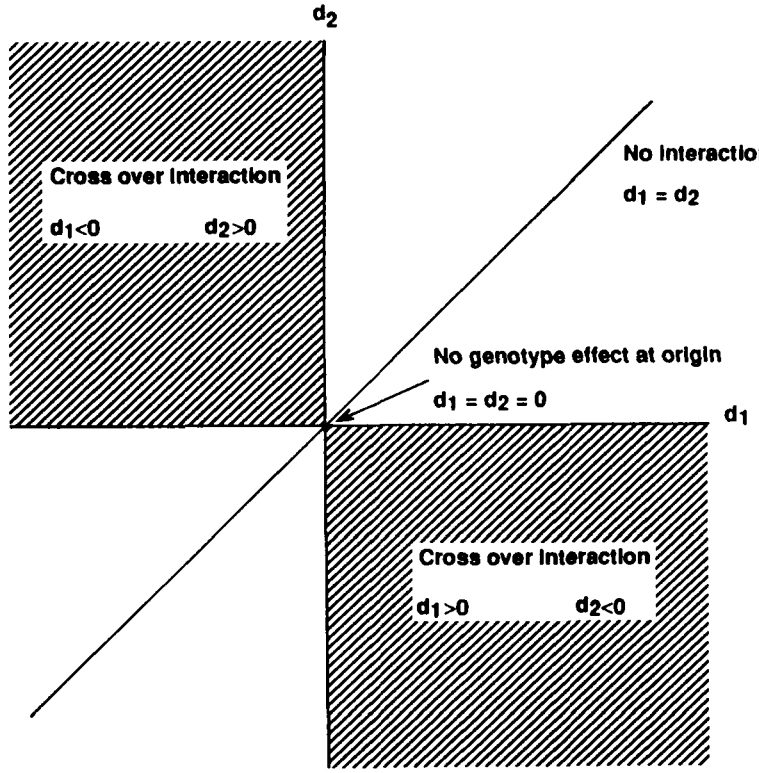

FIG. 1. The space of genotype effects d_1 and d_2. From Gail and Simon (1985).

tion; when $d_1 = d_2 = 0$ (at the origin), there are no differences in genotype means and no interaction. Qualitative or crossover interaction will occur in the second and fourth quadrants when $d_1<0$ and $d_2>0$ or $d_1>0$ and $d_2<0$. The rest of the cases are noncrossover interactions and consist of all the points in the first and third quadrants, excluding those lying on the line where $d_1 = d_2$. For the case of n environments, there will be two quadrants without a crossover and 2^{n-1} with a crossover interaction (Gail and Simon, 1985).

The hypothesis that there are no crossover interactions is equivalent to the hypothesis that the vector $d_1, d_2, \ldots, d_n$ lies in the quadrant in which all components are nonnegative or in the quadrant in which all components are nonpositive (see Fig. 1). For the case of two treatments, the likelihood ratio test of this hypothesis was developed by Gail and Simon (1985).

For the case of various genotypes evaluated in different environments, it is recommended that this test be restricted to *a priori* comparisons among genotypes (i.e., two genotypes across different environments); comparison among all possible genotypes will increase the experimentwise error rate (Baker, 1988b).

Azzalini and Cox (1984) described a test for the absence of crossover interactions that requires calculating the difference between all pairs of genotypes in all possible pairs of environments. Since this test is conservative, it is recommended that a significance level of .1 be used (Baker, 1988b).

V. MULTIVARIATE ANALYSES OF MULTILOCATION TRIALS

Multivariate analysis has three main purposes: (*a*) to eliminate noise from the data pattern (i.e., to distinguish systematic from nonsystematic variation); (*b*) to summarize the data; and (*c*) to reveal a structure in the data. In contrast with classic statistical methods, the function of multivariate analysis is to elucidate the internal structure of the data from which hypotheses can be generated and later tested by statistical methods (Williams and Gillard, 1971, cited by Gauch, 1982b).

Multivariate analyses are appropriate for analyzing two-way matrices of G genotypes and E environments. The response of any genotype in E environments may be conceived as a pattern in E-dimensional space, with the coordinate of an individual axis being the yield or other metric of the genotype in one environment.

Two groups of multivariate techniques have been used to elucidate the internal structure of genotype–environment interaction:

1. Ordination techniques, such as principal components analysis, principal coordinates analysis, and factor analysis, assume that data is continuous. These techniques attempt to represent genotype and environment relationships as faithfully as possible in a low-dimensional space. A graphical output displays similar genotypes or environments near each other and dissimilar items are farther apart. Ordination is effective for showing relationships and reducing noise (Gauch, 1982a, 1982b).

2. Classification techniques, such as cluster analysis and discriminant analysis, seek discontinuities in the data. These methods involve grouping similar entities in clusters and are effective for summarizing redundancy in the data.

A. Principal Components Analysis

Principal components analysis is one of the most frequently used multivariate methods (Pearson, 1901; Hotelling, 1933; Gower, 1966). Its aim is to transform the data from one set of coordinate axes to another, which preserves, as much as possible, the original configuration of the set of points and concentrates most of the data structure in the first principal components axes. In this process of data reduction, some original information is inevitably lost.

Principal components analysis assumes that the original variables define a Euclidean space in which similarity between items is measured as Euclidean distance. This analysis can effectively reduce the structure of a two-way genotype–environment data matrix of G (genotypes) points in E (environments) dimensions in a subspace of fewer dimensions. The matrix can also be conceptualized as E points in G dimensions.

The model is written as

$$Y_{ij} = \mu + \sum_{n=1}^{h} k_n v_{ni} s_{nj} \tag{8}$$

where the terms are defined as in (2). Under certain conditions, principal components analysis is a generalization of the linear regression analysis (Williams, 1952; Mandel, 1969; Johnson, 1977; Digby, 1979).

Mandel (1971) analyzed a two-way data matrix by applying the AMMI analysis. The first step in his solution was to conduct an analysis of variance with terms for two main effects and the interaction between rows and columns. The residual table (i.e., the row–column interaction) was partitioned into multiplicative terms where eigenvalues and eigenvectors are obtained. Finally, the relationships between the first two eigenvectors, which accounted for most of the variation, were examined.

Freeman and Dowker (1973) used principal components analysis to interpret the causes of genotype–environment interaction in carrot trials. Hirosaki *et al.* (1975) found that principal components analysis was more efficient than the linear regression method in describing genotypic performance. On the other hand, Perkins (1972) reported that principal components analysis was not useful for studying the adaptation of a group of inbred lines of *Nicotiana rustica*.

Principal components analysis combined with cluster analysis was effective in forming subgroups among 29 populations of faba bean (*Vicia faba* L.), which differed in mean performance and response across environments (Polignano *et al.*, 1989). Principal components have also been used by Suzuki (1968), Goodchild and Boyd (1975), and Hill and Goodchild (1981).

Zobel *et al.* (1988) presented analysis of variance and principal components analysis for seven soybean genotypes yield-tested in 35 environments (Table III). The genotype by environment interaction sum of squares of the analysis of variance was large but not significant. The principal components analysis with the first three principal axes accounting for 76% of the variation is found to be statistically efficient but undesirable for describing the additive main effects.

Kempton (1984) used AMMI analysis for summarizing the pattern of genotype responses across environments with different levels of nitrogen. The first principal component is the axis that maximizes the variation among genotypes. The second principal component is perpendicular to the first and maximizes the remaining variation. The display of the genotypes and environments along the first two principal component axes for the interaction table of residuals is called the biplot (Gabriel, 1971, 1981).

Table III

Analysis of Variance and Principal Components Analysis of a Soybean Trial[a]

Analysis of variance			Principal components analysis		
Source	df	Mean square	Source	df	Mean square
Treat	244	574***	Treat	244	574***
Geno	6	1499***	PCA 1	41	2599***
Env	34	3105***	PCA 2	39	471***
GE	204	125	PCA 3	37	264***
Error	667	111	Res	127	44
			Error	667	111

[a] From Zobel *et al.* (1988).

*** Significant at 0.001 probability level.

Figure 2 represents the biplot of 12 genotypes and 14 environments (7 sites each with low and high nitrogen levels). The component of the total interaction due to nitrogen level is small, since the biplot shows that high- and low-nitrogen trials are closely associated. Environments represented by vectors of similar orientation but different length usually give similar genotype rankings.

Zobel *et al.* (1988) and Crossa *et al.* (1990) used the same model to analyze a series of soybean and maize trials, respectively. Additive main effects for genotype and environments are first fitted by the analysis of variance. Then, multiplicative effects for genotype by environment interaction are calculated by principal components analysis. The biplot of the model helps to visualize the overall pattern of response as well as specific interactions between genotypes and environments.

Ordination techniques such as principal components analysis may have limitations. First, in reducing dimensionality of multivariate data, distortions may sometimes occur. If the percentage of variance accounted for by the first principal components axes is small, individuals that are really far apart may be represented by points that are close together (Gower, 1967).

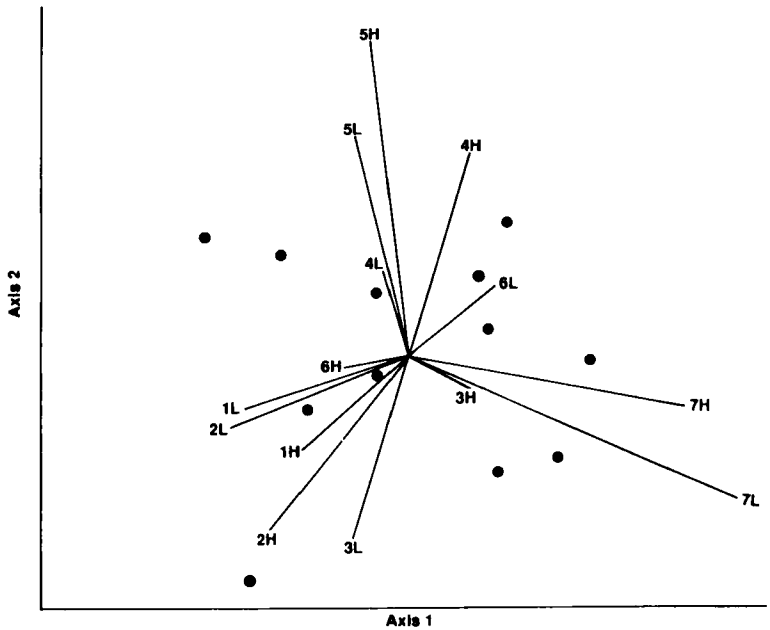

FIG. 2. First two principal component axes for genotypes (•) and environments based on residual yields. (Sites are coded from 1 to 5; L and H are trials with low and high nitrogen levels, respectively.) From Kempton (1984).

In this case, higher axes may be inspected to identify points with large displacement not revealed in lower dimensions. Second, a lack of correlation between variables prevents few dimensions from accounting for most of the variation (Williams, 1976). Third, sometimes the components do not have any obvious relationship to environmental factors. Fourth, contrary to analysis of variance, which assumes a complete additive model and treats the interaction as a residual, principal components analysis assumes a complete multiplicative model without any description of the main effects of genotypes and environments (Zobel *et al.*, 1988). This is important in the context of multilocation trials, in which genotype means are of primary interest. Principal components analysis confounds the additive (main effects of genotypes and environments) structure of the data with the nonadditivity (genotype–environment interaction). The fifth limitation is that nonlinear association in the data prevents principal components analysis from efficiently describing the real relationships between entities (Williams, 1976).

The linear regression method uses only one statistic, the regression coefficient, to describe the pattern of response of a genotype across environments, and, as mentioned previously, most of the information is wasted in accounting for deviation. Principal components analysis, on the other hand, is a generalization of linear regression that overcomes this difficulty by giving more than one statistic, the scores on the principal component axes, to describe the response pattern of a genotype (Eisemann, 1981).

B. Principal Coordinates Analysis

Principal coordinates analysis (Gower, 1966) is a generalization of principal components analysis in which any measure of similarity between individuals can be used. Its objective and limitations are similar to those of principal components analysis.

Principal coordinates analysis was used in combination with cluster analysis ("pattern" analysis) to study the adaptation of soybean lines evaluated across environments in Australia (Mungomery *et al.*, 1974; Shorter *et al.*, 1977). The authors found these analyses to be useful for helping breeders choose among test sites for early screening of breeding lines.

Principal coordinates analysis was employed to examine the use of a reference set of genotypes to monitor genotype–environment interaction (Fox and Rosielle, 1982a) and also to assess methods for removing environmental main effects to provide a description of environments (Fox and Rosielle, 1982b).

A spatial method for assessing yield stability, in which principal coordinates analysis is based on a suitable measure of similarity between genotypes, has been proposed by Westcott (1987). As pointed out by Crossa (1988), the method has several advantages: (*a*) it is trustworthy when used for data that include extremely low or high yielding sites; (*b*) it does not depend on the set of genotypes included in the analysis; and (*c*) it is simple to identify stable varieties from the sequence of graphic displays. The spatial method has been extensively used by Crossa *et al.* (1988a,b, 1989) to assess the yield stability of CIMMYT's maize genotypes evaluated across international environments.

C. Factor Analysis

Factor analysis is an ordination procedure related to principal components analysis, the "factors" of the former being similar to the principal components of the later. A large number of correlated variables is reduced to a small number of main factors (Cattell, 1965), and variation is explained in terms of general factors common to all variables and in terms of factors unique to each variable. The axes of the general factors may be rotated to oblique positions to conform to hypothetical ideas.

Factor analysis has been used to understand relationships among yield components and morphological characteristics of crops (Walton, 1972; Seiler and Stafford, 1985). Jardine *et al.* (1963) used an oblique rotation to indicate four relatively independent factors related to bread wheat baking quality.

Peterson and Pfeiffer (1989) applied principal factor analysis to study the underlying structures and relationships of test sites, based on winter wheat performance. The authors grouped the original 56 locations into seven regions, which can be considered megaenvironments for winter wheat adaptation. The association between secondary factors was used to identify transitional environments between the seven major regions.

D. Cluster Analysis

Cluster analysis is a numerical classification technique that defines groups or clusters of individuals. Two types of classification can be distinguished. The first is nonhierarchical classification, which assigns each item to a class. Relationships among classes are not characterized, so this type is useful in the early stages of data analysis. The second type is hierarchical

classification, which groups individuals into clusters and arranges these into a hierarchy for the purpose of studying relationships in the data.

Cluster analysis requires a measure of similarity between the individuals to be classified, and it imposes a discontinuity in the data. The method has been used to study genotype adaptation by simplifying the pattern of responses and to subdivide genotypes and environments into more homogeneous groups. Comprehensive reviews of the application of cluster analysis to the study of genotype–environment interactions can be found in Lin *et al.* (1986) and Westcott (1987).

Some of the disadvantages of cluster analysis are: (*a*) numerous hierarchical grouping algorithms exist, and each of them may produce different cluster groups; (*b*) the truncation level of the classificatory hierarchies may be decided arbitrarily; (*c*) many different similarity measures can be used (Lin *et al.*, 1986, listed nine), yielding different results; and (*d*) cluster analysis may produce misleading results by showing structures and patterns in the data when they do not exist (Gordon, 1981, cited by Westcott, 1987).

VI. AMMI ANALYSIS

The additive main effect and multiplicative interaction (AMMI) method integrates analysis of variance and principal components analysis into a unified approach (Bradu and Gabriel, 1978; Gauch, 1988). It can be used to analyze multilocation trials (Gauch and Zobel, 1988; Zobel *et al.*, 1988; Crossa *et al.*, 1990).

AMMI analysis first fits the additive main effects of genotypes and environments by the usual analysis of variance and then describes the nonadditive part, genotype–environment interaction, by principal components analysis. The AMMI model is given by Eq. (3).

The AMMI method is used for three main purposes. The first is model diagnosis. AMMI is more appropriate in the initial statistical analysis of yield trials, because it provides an analytical tool for diagnosing other models as subcases when these are better for a particular data set (Bradu and Gabriel, 1978; Gauch, 1985). The second use of AMMI is to clarify genotype–environment interactions. AMMI summarizes patterns and relationships of genotypes and environments (Kempton, 1984; Zobel *et al.*, 1988; Crossa *et al.*, 1990). The third use is to improve the accuracy of yield estimates. Gains have been obtained in the accuracy of yield estimates that are equivalent to increasing the number of replicates by a factor of two to five (Zobel *et al.*, 1988; Crossa *et al.*, 1990). Such gains may be used to

reduce costs by reducing the number of replications, to include more treatments in the experiment, or to improve efficiency in selecting the best genotypes. This last benefit has obvious implications for breeding programs and particularly for maize hybrid testing systems, in which designs with fewer replicates per location are used (Bradley *et al.*, 1988).

A. AMMI Analysis with Predictive Success

Traditional analysis of variance of multilocation trials is intended to forecast agricultural performance, but it focuses only on postdictive assessment of genotype yield responses without evaluating the model's predictive accuracy with validation data not used in constructing the model.

Gauch (1985, 1988) emphasized the model's success in predicting validation data (prediction criteria), in contrast to its success in fitting its own data (postdiction criteria). Because multilocation trials are used for selecting genotype or agronomic treatments for farmers' fields in new environments, model evaluation should measure predictive success. Gauch proposed that AMMI analysis be used with prediction criteria.

Prediction assessment consists of splitting data into two subgroups, *modeling data* and *validation data*, and comparing the success of several models by computing their sum of squared difference (SSD) between model predictions and validation data. A small value of SSD indicates good predictive accuracy. Several models are then constructed and compared empirically in terms of their ability to predict the validation data: AMMI0, which estimates only the additive main effects of genotypes and environments and retains none of the principal components axes (PCA); AMMI1, which combines the additive main effects from AMMI0 with the genotype–environment interaction effects estimated from the first principal component axis (PCA 1); AMMI2 and so on, up to the full model with all PCA axes. The predictive values of the full model are equal to the average of the replicates selected at random for modeling.

Results of postdictive AMMI analysis of a trial consisting of 15 soybean genotypes evaluated in 15 environments are given in Table IV (Gauch, 1988). The postdictive evaluation using F-test at 5% showed that three PCAs of the interaction are significant; therefore, the model, including the two main effects, has 103 df. However, this information includes pattern and noise (systematic and nonsystematic variation). Prediction assessment, on the other hand, does discriminate between pattern and noise and indicates AMMI with one interaction PCA as the best predictive model (Fig. 3). This model has 55 df—14 for genotypes, 14 for environ-

Table IV

AMMI Analysis for a Soybean Trial[a]

Source	df	SS	MS
Environment	14	38,798	2,771***
Genotype	14	2,552	182***
G × E	196	6,880	35***
PCA 1	27	2,348	87***
PCA 2	25	1,250	50***
PCA 3	23	1,010	44***
PCA 4	21	736	35
Residual	100	1,536	15
Error	210	4,649	22

[a] From Gauch (1988).
*** Significant at 0.001 probability level.

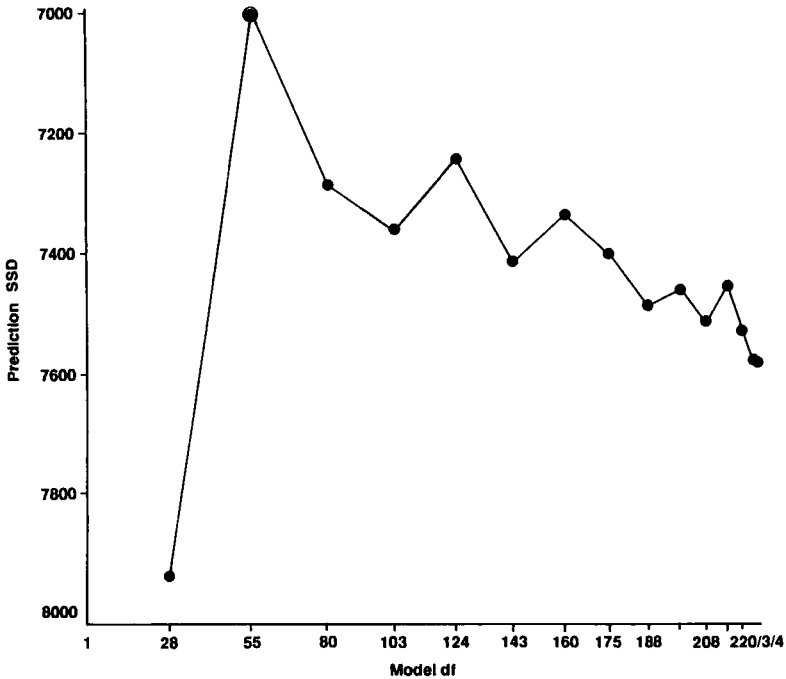

FIG. 3. Sum of squared difference (SSD) between model prediction and validation data for 15 models (AMMI0 with 28 df to the full model with 224 df). From Gauch (1988).

ments, and 27 for the interaction PCA 1. Further interaction PCAs will capture mostly noise and therefore do not help to predict validation observations. The interaction of 15 soybean genotypes with 15 environments is best predicted by the first principal component of genotypes and environments. Thus, the model is

$$Y_{ij} = \mu + G_i + E_j + k_1 v_{1i} s_{1j} + e_{ij} \quad (9)$$

From (9) it can be seen that, when a genotype or an environment has an interaction PCA score of nearly 0, it has a small interaction. When both have PCA scores of the same sign, their interaction is positive; if different, their interaction is negative.

For data in which AMMI1 is found to be the best predicted model, a graphical display of the genotype and environment interaction PCA 1 and their mean effects should be useful for revealing favorable patterns in genotype response across environments.

Figure 4 gives the mean on the x axis and the AMMI interaction PCA 1 scores on the y axis of 17 maize genotypes tested in 36 environments (Crossa *et al.*, 1990). Three groups of genotypes with different genetic composition can be seen: (*a*) group 1 includes genotypes 13, 14, 15, and 17, which contain temperate germplasm from the U.S. Corn Belt and southern Europe; (*b*) group 2 comprises genotypes 1, 2, 3, 4, and 5, which are from subtropical regions and have intermediate maturity; and (*c*) group 3 contains genotypes 6 to 12 and 16, which are derived from lowland tropical maize types from Mexico and the Caribbean islands. Interaction PCA 1 scores arrange the environments in a sequence from tropical environments

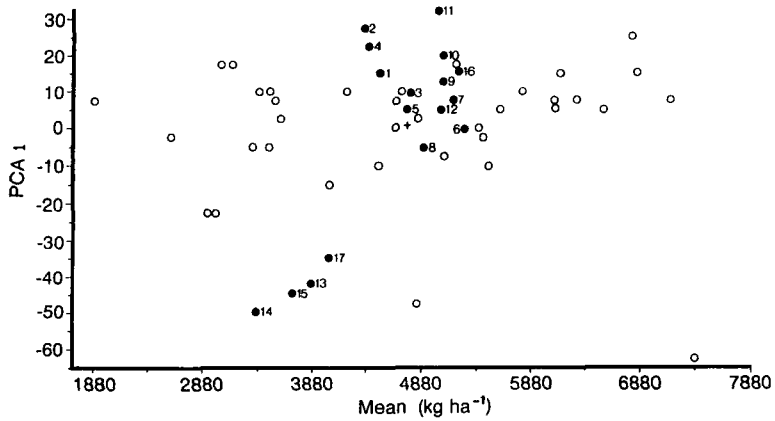

FIG. 4. Plot of the means (kg ha^{-1}) and PCA 1 scores of 17 maize genotypes (●) and 36 environments (○). From Crossa *et al.* (1990).

(positive PCA 1) to temperate environments (negative PCA 1). The two temperate environments with the greatest negative PCA 1 scores favor temperate germplasm (group 1). At the other extreme of the diagram, tropical environments tend to favor genotypes from group 2 and 3.

VII. OTHER METHODS OF ANALYSIS

Many other approaches might be employed for studying genotype–environment interactions. Several of them have not been examined systematically or extensively used for different crops.

In most yield trials, environments are measured by the average yield of the genotypes or agronomic treatments. However, it is important to collect, analyze, and interpret physiological and environmental variables for (*a*) studying their relationships with genotype performance and (*b*) understanding the causes of the observed genotype–environment interaction (Westcott, 1986; Eisemann and Mungomery, 1981). The differential physiological responses of genotypes to edaphic and climatic factors, especially those related to nutrient efficiency and stress tolerance, are relevant to genotype–environment interaction (Baker, 1988a,c).

The multilinear regression method, in which environmental data are used as independent variables, can be employed for predictive purposes (Knight, 1970; Feyerherm and Paulsen, 1981; Haun, 1982). Hardwick (1972) and Hardwick and Wood (1972) used physiological and environmental variables to develop a predictive multiple linear regression model. Principal components analysis, combined with multiple regression, may be useful for reducing the number of environmental variables to be included in the final analysis (Perkins, 1972).

Principal components analysis was used by Holland (1969) to summarize and interpret environmental data. However, it is of limited use, because the importance of a certain variable in the analysis may not be related to the extent of genotype response (Eisemann and Mungomery, 1981).

Most of the exploratory or geometrical methods can be applied to the analysis of multilocation trials, although their use for this purpose has not been investigated. Ordination techniques, such as weighted average (Rowe 1956), polar ordination (Bray and Curtis, 1957), reciprocal average (Fisher, 1940), and detrended correspondence analysis (Hill and Gauch, 1980) have been used in community ecology to discover structures in data matrices (Gauch, 1982b). Their use in examining the pattern of genotype (or environment) responses needs investigation.

Canonical discriminant analysis has been used to allocate environments according to their interaction with genotypes (Seif *et al.*, 1979).

The stratified ranking method was used by Fox *et al.* (1990) for analyzing general adaptation of a large international triticale data base. The technique scores the number of locations for which each line occurred in the top, second, and bottom one-third of the entries in each trial. A line that occurred in the top one-third of the entries across locations was considered well adapted.

Unbalanced data often occur in multilocation trials as a result of (*a*) missing plots or (*b*) combining results of different experiments that do not have the same set of treatments. For incomplete data, missing plot values can be fitted, and the genotype–environment interaction sum of squares can be further partitioned into principal components (Freeman, 1975).

An algorithm for inputting missing values and then fitting the additive main effect and multiplicative interaction (AMMI) model has recently been developed (Gauch and Zobel, 1990).

VIII. GENERAL CONSIDERATIONS AND CONCLUSIONS

Data from multilocation trials help researchers estimate yields more accurately, select better production alternatives, and understand the interaction of these technologies with environments.

Several methodologies have been presented for efficient statistical analysis of such data. For geneticists, plant breeders, and agronomists, parametric stability statistics, obtained by linear regression analysis, are mathematically simple and biologically interpretable. However, this method has major disadvantages: (*a*) it is uninformative when linearity fails; (*b*) it is highly dependent on the set of genotypes and environments included in the analysis; and (*c*) it tends to oversimplify the different response patterns by explaining the interaction variation in one dimension (regression coefficient), when in reality it may be highly complex. There is a danger in sacrificing relevant information for easy biological and statistical interpretation.

A broad range of multivariate methods can be used to analyze multilocation yield trial data and assess yield stability. Although some of them overcome the limitations of linear regression, the results are often difficult to interpret in relation to genotype–environment interaction (as is the case with principal components analysis and cluster analysis). Certain multi-

variate techniques or a combination of them offer very relevant biological information and are statistically simple.

The integration of certain ordination and classification multivariate methods into so-called "pattern" analysis and factor analysis and the biplot method are valuable tools for grouping environments or genotypes showing similar response patterns.

The combination of analysis of variance and principal components analysis in the AMMI model, along with prediction assessment, is a valuable approach for understanding genotype–environment interaction and obtaining better yield estimates. Agronomic predictive assessment with AMMI can be used to analyze the results of on-farm trials. More research is needed to quantify the probability of successful selection of a genotype or agronomic treatment when using AMMI predictive value, compared with the probability of its selection based on the predictive value of other models.

Only qualitative or crossover interactions are relevant in agriculture. Therefore, appropriate statistical analysis for quantifying and testing changes in rank from one environment to another is required.

More attention has to be devoted to the collection, analysis, and interpretation of environmental and physiological variables. This will help to characterize particular genotypes and geographical regions and therefore better explain certain aspects of the interaction.

ACKNOWLEDGMENTS

The author thanks Drs. Kent Eskridge, Paul Fox, and Hugh Gauch for their helpful comments on the manuscript.

REFERENCES

Anderson, J. R. 1974. *Rev. Marketing Agric. Econ.* **42,** 131–184.
Azzalini, A., and Cox, D. R. 1984. *J. R. Stat. Soc.* **46,** 335–343.
Baker, R. J. 1969. *Can. J. Plant Sci.* **49,** 743–751.
Baker, R. J. 1988a. *In* "Proceedings of the Second International Conference on Quantitative Genetics." Sinauer, Sunderland, Massachusetts.
Baker, R. J. 1988b. *Can. J. Plant Sci.* **68,** 405–410.
Baker, R. J. 1988c. "ISI Atlas of Science: Animal and Plant Sciences."
Barah, B. C., Binswanger, H. P., Rana, B. S., Rao, N. P. 1981. *Euphytica* **30,** 451–458.
Binswanger, H. P., and Barah, B. C. 1980. Research Bulletin No. 3. International Crops Research Institute for the Semi-Arid Tropics (ICRISAT). Palanchero, India.
Bradley, J. P., Knittle, K. N., and Troyer, A. F. 1988. *J. Prod. Agric.* **1,** 34–38.
Bradu, D., and Gabriel, K. R. 1978. *Technometrics* **20,** 47–68.
Bray, J. R., and Curtis, J. T. 1957. *Ecol. Monogr.* **27,** 325–349.

Bruckner, P. L., and Frohberg, R. C. 1987. *Crop Sci.* **27,** 31–36.
Bucio Alanis, L., Perkins, J. M., and Jinks, J. L. 1969. *Heredity* **24,** 115–127.
Byth, D. E., Eisemann, R. L., and De Lacy, I. H. 1976. *Heredity* **37,** 215–230.
Cattell, 1965. *Biometrics* **21,** 190–215.
Cochran, W. G., and Cox, G. M. 1957. "Experimental Designs," 2nd Ed. Wiley, New York.
Cockerham, C. C. 1964. "Statistical Genetics and Plant Breeding," Vol. 982, pp. 53–93. National Academy of Science-National Research Council, Washington, D.C.
Comstock, R. E., and Moll, R. H. 1964. "Statistical Genetics and Plant Breeding," Vol. 982, pp. 164–196. National Academy of Science-National Research Council, Washington, D.C.
Crossa, J. 1988. *Theor. Appl. Genet.* **75,** 460–467.
Crossa, J., Westcott, B., and Gonzalez, C. 1988a. *Theor. Appl. Genet.* **75,** 863–868.
Crossa, J., Westcoot, B., and Gonzalez, C. 1988b. *Exp. Agric.* **24,** 253–263.
Crossa, J., Westcott, B., and Gonzalez, C. 1989. *Euphytica* **40,** 245–251.
Crossa, J., Gauch, H. G., and Zobel, R. W. (1990). *Crop Sci.* **30,** 493–500.
Digby, P. G. N. 1979. *J. Agric. Sci.* **93,** 81–86.
Eagles, H. A., and Frey, K. J. 1977. *Crop Sci.* **17,** 253–256.
Easton, H. S., and Clements, R. J. 1973. *J. Agric. Sci.* **80,** 43–52.
Eberhart, S. A., and Russell, W. A. 1966. *Crop Sci.* **6,** 36–40.
Eisemann, R. L. 1981. *In* "Interpretation of Plant Respnse and Adaptation to Agricultural Environments." Univ. of Queensland, St. Lucia, Brisbane.
Eisemann, R. L., and Mungomery, V. E. 1981. *In* "Interpretation of Plant Response and Adaptation to Agricultural Environments." Univ. of Queensland, St. Lucia, Brisbane.
Eskridge, K. M. 1990. *Crop Sci.* **30,** 369–374.
Feyerherm, A. M., and Paulsen, G. M. 1981. *Agron. J.* **73,** 277–282.
Finlay, K. W., and Wilkinson, G. N. 1963. *Aust. J. Agric. Res.* **14,** 742–754.
Fisher, R. A. 1918. *Trans. R. Soc. Edinburgh* **52,** 399–433.
Fisher, R. A. 1925. "Statistical Methods for Research Workers." Oliver & Boyd, London.
Fisher, R. A. 1940. *Ann. Eugenics* **10,** 422–429.
Fox, P. N., and Rathjen, A. J. 1981. *Aust. J. Agric. Res.* **32,** 691–702.
Fox, P. N., and Rosielle, A. A. 1982a. *Crop Sci.* **22,** 1171–1175.
Fox, P. N., and Rosielle, A. A. 1982b. *Euphytica* **31,** 645–656.
Fox, P. N., Skovmand, B., Thompson, B. K., Braun, H.-J., and Cormier, R. 1990. *Euphytica* **47,** 57–64.
Freeman, G. H. 1973. *Heredity* **31,** 339–354.
Freeman, G. H. 1975. *Appl. Stat.* **24,** 1–46.
Freeman, G. H., 1985. *J. Appl. Stat.* **12,** 3–10.
Freeman, G. H., and Dowker, B. D. 1973. *Heredity* **30,** 97–109.
Freeman, G. H., and Perkins, J. M. 1971. *Heredity* **27,** 15–23.
Gabriel, K. R. 1971. *Biometrika* **58,** 453–467.
Gabriel, K. R. 1981. "Interpreting Multivariate Data." Wiley, Chichester.
Gail, M., and Simon, R. 1985. *Biometrics* **41,** 361–372.
Gardner, C. O. 1964. "Statistical Genetics and Plant Breeding," Vol. 982. pp. 225–252. National Academy of Science-National Research Council, Washington, D.C.
Gauch, H. G. 1982a. *Ecology* **63,** 1643–1649.
Gauch, H. G. 1982b. "Multivariate Analysis in Community Ecology," 1st Ed. Cambridge Univ. Press, London and New York.
Gauch, H. G. 1985. Mimeo 85–7. Cornell Univ. Ithaca, New York.
Gauch, H. G. 1988. *Biometrics* **44,** 705–715.
Gauch, H. G., and Zobel, R. W. 1990. *Theor. Appl. Genet.* (in press).

Gauch, H. G., and Zobel, R. W. 1988. *Theor. Appl. Genet.* **76,** 1–10.
Gollob, H. F. 1968. *Psychometrika* **33,** 73–116.
Goodchild, N. A., and Boyd, W. J. R. 1975. *Aust. J. Agric. Res.* **26,** 209–217.
Gordon, A. D. 1981. "Classification." Chapman & Hall, London.
Gower, J. C. 1966. *Biometrika* **53,** 325–338.
Gower, J. C. 1967. *Statistician* **17,** 13–28.
Graybill, F. A., and Hultquist, R. A. 1961. *Ann. Math. Stat.* **32,** 261–269.
Hardwick, R. C. 1972. *Nature (London) New Biol.* **236,** 191–192.
Hardwick, R. C., and Wood, J. T. 1972. *Heredity* **28,** 209–222.
Haun, J. R. 1982. *Agric. Meteor.* **27,** 191–207.
Hill, J. 1975. *J. Agric. Sci.* **85,** 477–493.
Hill, J., and Goodchild, N. A. 1981. *Theor. Appl. Genet.* **59,** 317–325.
Hill, M. O., and Gauch, H. G. 1980. *Vegetatio* **42,** 47–58.
Hill, R. R., and Baylor, J. E. 1983. *Crop Sci.* **23,** 811–815.
Hirosaki, S., Okuno, T., and Shiyomi, M. 1975. "JIBP Synthesis," Vol. 6. Univ. of Tokyo Press, Tokyo.
Holland, D. A. 1969. *Exp. Agric.* **5,** 151–164.
Hotelling, H. 1933. *J. Educ. Psychol.* **24,** 417–441.
Jardine, R., Moss, H. J., and Mullaly, J. V. 1963. *Aust. J. Agric. Res.* **14,** 603–621.
Johnson, G. R. 1977. *Crop Sci.* **17,** 837–842.
Kempton, R. A. 1984. *J. Agric. Sci.* **103,** 123–135.
Knight, R. 1970. *Euphytica* **19,** 225–235.
Lin, C. S., and Binns, M. R. 1988. *Can. J. Plant Sci.* **68,** 193–198.
Lin, C. S., Binns, M. R., and Lefkovitch, L. P. 1986. *Crop Sci.* **26,** 894–900.
Mandel, J. 1961. *J. Am. Stat. Assoc.* **56,** 878–888.
Mandel, J. 1969. *J. Res. Natl. Bur. Stand. (U.S.)* **11,** 411–429.
Mandel, J. 1971. *Technometrics* **13,** 1–18.
Mather, K., and Caligari, P. D. S. 1974. *Heredity* **33,** 43–59.
Moll, R. H., Cockerham, C. C., Stuber, C. W., and Williams, W. P. 1978. *Crop Sci.* **18,** 641–645.
Mead, R., Riley, J., Dear, K., and Singh, S. P. 1986. *Biometrics* **42,** 253–266.
Menz, K. M. 1980. *Field Crop Res.* **3,** 33–41.
Mungomery, V. E., Shorter, R., and Byth, D. E. 1974. *Aust. J. Agric. Res.* **25,** 59–72.
Patterson, H. D., and Silvey, V. 1980. *J. R. Stat. Soc.* **143,** 219–252.
Patterson, H. D., and Thompson, R. 1971. *Biometrika* **58,** 545–554.
Patterson, H. D., and Thompson, R. 1975. *Proc. 8th Int. Biometric Conf.* pp. 197–207.
Pearson, K. 1901. *Philos. Mag.* **2,** 559–572.
Perkins, J. M. 1972. *Heredity* **29,** 51–70.
Peterson, C. J., and Pfeiffer, W. H. 1989. *Crop Sci.* **29,** 276–282.
Plaisted, R. L., and Peterson, L. C. 1959. *Am. Potato J.* **36,** 381–385.
Polignano, G. B., Ugenti, P., and Perrino, P. 1989. *Euphytica* **40,** 31–41.
Pooni, H. S., and Jinks, J. L. 1980. *Heredity.* **45,** 389–400.
Robinson, D. 1987. *Statistician* **36,** 3–14.
Rowe, J. S. 1956. *Ecology* **37,** 461–473.
Seif, E., Evans, J. C., and Balaam, L. N. 1979. *Aust. J. Agric. Res.* **30,** 1021–1026.
Seiler, G. J., and Stafford, R. E. 1985. *Crop Sci.* **25,** 905–908.
Shorter, R. 1981. *In* "Interpretation of Plant Response and Adaptation to Agricultural Environments." Univ. of Queensland, St. Lucia, Brisbane.
Shorter, R., Byth, D. E., and Mungomery, V. E. 1977. *Aust. J. Agric. Res.* **28,** 223–235.
Shukla, G. K. 1972. *Heredity* **29,** 237–245.

Suzuki, S. 1968. *Proc. Int. Congr. Genet. 12th, 1968* **1.**
Thompson, W. A. 1961. *Bull. Int. Inst. Stat.* **34,** 1–4.
Thompson, W. A. 1962. *Ann. Math. Stat.* **33,** 273–289.
Walton, P. D. 1972. *Crop. Sci.* **12,** 731–733.
Westcott, B. 1986. *Heredity* **56,** 243–253.
Westcott, B. 1987. *J. Agric. Sci.* **108,** 267–274.
Williams, E. J. 1952. *Biometrika* **39,** 65–81.
Williams, W. T. 1976. "Pattern Analysis in Agriculture Science." CSIRO, Melbourne and Elsevier, Amsterdam.
Williams, W. T., and Gillard, P. 1971. *Aust. J. Agric. Res.* **22,** 245–260.
Witcombe, J. R. 1988. *Euphytica* **39,** 11–18.
Witcombe, J. R., and Wittington, W. J. 1971. *Heredity* **26,** 397–411.
Wricke, G. 1962. *Z. Pflanzenzuecht.* **47,** 92–96.
Wricke, G. 1964. *Z. Pflanzenzuecht.* **52,** 127–138.
Yates, F., and Cochran, W. G. 1938. *J. Agric. Sci.* **28,** 556–580.
Zobel, R. W., Wright, M. J., and Gauch, H. G. 1988. *Agron. J.* **80,** 388–393.