Theor Appl Genet (1992) 84:193–200

Alternative partitioning of the genotype-by-environment interaction *

W. Muir [1,], W. E. Nyquist [2], and S. Xu [1]**

[1] Deaprtment of Animal Sciences and [2] Department of Agronomy, Purdue University, West Lafayette, IN 47907, USA

Received June 25, 1991; Accepted October 17, 1991
Communicated by L. D. Van Vleck

Summary. Alternative methods for partitioning the genotype-by-environment interaction, for an arbitrary number of genotypes or environments, were examined. Partitioning of the interaction is important in order to determine the nature of the interaction. Two methods of partitioning were examined; both separated the interaction into two types: (1) due to heterogeneous variances or (2) due to imperfect correlations. Method 1 was based on heterogeneity among environments in the scaling of differences among genotypes. Method 2 was based on heterogeneity among genotypes in the scaling of differences among environments. Any remaining interaction arises from deviations from the perfect positive correlation of genotypic rankings among environments (Method 1) or of environmental rankings among genotypes (Method 2).

Method 1 is more appropriate for random genotypes that are to be tested in either fixed or random environments. With Method 1, the interactions that arise mainly from heterogeneity of genotypic scaling among environments are generally unimportant. However, if environments are fixed and interactions are mainly due to imperfect correlations of rankings, specialized lines may be indicated for each environment. Method 2 is more useful in evaluating fixed genotypes for sensitivity to random environments. A partitioning of the interaction into that due to the type of interaction within each genotype was shown to be useful in that situation.

Key words: Interactions – Stability – Sensitivity – Clustering – Similarity coefficient

Introduction

The importance of genotype-by-environment interactions in the context of plant and animal breeding is well recognized (see Freeman 1973 for extensive review; Kang 1990). Statistical methods for detecting these interactions are abundant (Freeman 1973; Lin et al. 1986). However, once an interaction has been detected, it is equally important to determine its nature.

The interaction sum of squares can be separated into two parts, that associated with heterogeneous genetic variances measured in each environment and that due to differences in genetic correlations of the same trait measured in different environments. This partitioning was first shown by Robertson (1959) and extended by Dickerson (1962), Yamada (1962), Cockerham (1963), Eisen and Saxton (1983), and Yamada et al. (1988). In contrast, Moll et al. (1978) partitioned the interaction into that due to heterogeneous environmental variances measured for each genotype and that due to differences in environmental correlations between genotypes. In either case, the interaction can be separated into that due to differences in scale and that due to imperfect correlations or changes in rank. In the first case, the scales and correlations are genetic, whereas in the second case, the scales and correlations are environmental.

The purpose of this paper is to examine alternative methods for partitioning of the interaction to determine if one is more appropriate or informative than the other for a given situation and to extend the method of Moll et al. (1978) to allow separation of those effects for each genotype.

Statistical methods

Partitioning the interaction into two types

In the extreme case, an interaction can be exclusively due to either a change in scale or rank. For a change in scale only, the genotypic values change proportionally from environment to environment, whereas for a change in rank only, the scale remains the same. However, most interactions exhibit characteristics of both types of interaction. This is especially true with multiple genotypes or environments where a partitioning is useful in quantifying the proportion of the interaction sum of squares due to each type.

* Journal Paper No. 123737 of the Purdue University Agricultural Experiment Station
** Correspondence to: W. Muir

Assume an experiment in which either genotypes, environments, or both were random, and that each of g genotypes was reared in e environments. Performance on n individuals within each combination was measured. The model for analysis is:

$$Y_{ijk} = \mu + G_i + E_j + GE_{ij} + \varepsilon_{(ij)k}$$
$$i = 1, \ldots, g;\ j = 1, \ldots, e;\ k = 1, \ldots, n$$

where Y_{ijk} is the performance of the kth individual in the jth environment of the ith genotype. Expected mean squares for interaction and within sources of variation are given in Table 1. Compositions of the expected mean squares for environment and genotypes are not given since each depends on whether the other factor is fixed or random. Furthermore, their compositions are irrelevant in the present context.

The total sum of squares for the G × E source of variation is computed as

$$SS(G \times E) = n\sum_{i}^{g}\sum_{j}^{e}(\bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...})^2 \quad (1)$$

and by principle of symmetric differences (Casella and Berger 1990, p 237–8)

$$= n\sum_{j}^{e}\sum_{i}^{g}\sum_{i' \neq i}^{g}(R_{ij} - R_{i'j})^2/2g$$

where $R_{ij} = \bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...}$, and after symplifying

$$SS(G \times E) = n\sum_{j}^{e}\sum_{i}^{g}\sum_{i' \neq i}^{g}[(\bar{Y}_{ij.} - \bar{Y}_{i..}) - (\bar{Y}_{i'j.} - \bar{Y}_{i'..})]^2/2g$$
$$= n\sum_{i}^{g}\sum_{i' \neq i}^{g}[S_i^2 + S_{i'}^2 - 2S_{ii'}]/2g$$

Table 1. Analysis of variance and expected mean squares

Source	*df*	Sum of squares	Mean square	Expected mean square
Environments (E)	e-1			
Genotypes (G)	g-1			
G × E	(g-1) (e-1)	SS (G × E)	MS (GE)	$\sigma^2 + n\sigma^2_{GE}$
Individuals within	ge (n-1)	SS (within)	MS (E)	σ^2

where $S_i^2 = \sum_{j}^{e}(\bar{Y}_{ij.} - \bar{Y}_{i..})^2$, and $S_{ii'} = \sum_{j}^{e}(\bar{Y}_{ij.} - \bar{Y}_{i..})(\bar{Y}_{i'j.} - \bar{Y}_{i'..})$ or re-expressed in terms of correlations

$$SS(G \times E) = n\sum_{i}^{g}\sum_{i<i'}^{g}[(S_i - S_{i'})^2 + 2(1 - r_{ii'})S_iS_{i'}]/g \quad (2)$$

$$= n\left[\sum_{i}^{g}(S_i - \bar{S}_.)^2 + 2\sum_{i<i'}^{g}\sum^{g}(1 - r_{ii'})S_iS_{i'}/g\right] \quad (3)$$

where

$$S_i = \sqrt{S_i^2},\ \bar{S}_. = \textstyle\sum_i S_i/g, \quad \text{and} \quad r_{ii'} = S_{ii'}/S_iS_{i'}.$$

Note that $\sigma_i^2 = S_i^2/(e-1)$ is the variance among environments within the ith genotype. For the special case of $n=1$ and $g=2$, these results are as given by Moll et al. (1978). This method of computing the sum of squares will hereafter be referred to as Method 2.

By similar argument and expansion on subscript j, rather than i, Eqs. 2 and 3 could have been formulated as:

$$SS(G \times E) = n\sum_{j}^{e}\sum_{j<j'}^{e}[(Z_j - Z_{j'})^2 + 2(1 - r_{jj'})Z_jZ_{j'}]/e \quad (2')$$

$$= n\left[\sum_{j}^{e}(Z_j - \bar{Z}_.)^2 + 2\sum_{j<j'}^{e}\sum^{e}(1 - r_{jj'})Z_jZ_{j'}/e\right] \quad (3')$$

where

$$Z_j^2 = \sum_{i}^{g}(\bar{Y}_{ij.} - \bar{Y}_{.j.})^2,\ Z_{jj'} = \sum_{i}^{g}(\bar{Y}_{ij.} - \bar{Y}_{.j.})(\bar{Y}_{ij'.} - \bar{Y}_{.j'.}),$$
$$Z_j = \sqrt{Z_j^2},\ \bar{Z}_. = \textstyle\sum_j Z_j/e \quad \text{and} \quad r_{jj'} = Z_{jj'}/Z_iZ_{i'}.$$

Note that $\sigma_j^2 = Z_j^2/(g-1)$ is the variance among genotypes within the jth environment. This method of computing the sum of squares will hereafter be referred to as Method 1. This formulation for Method 1, although in terms of sums of squares, is analogous to the formulations of Cockerham (1963, p 88), Robertson (1959), Dickerson (1962), and Yamada (1962) in terms of variances and covariances.

With either method of analysis, the sum of squares for the genotype-by-environment interaction can be partitioned into two types of interaction: (1) due to heterogeneity in scaling of either genetic or environmental effects (HV), and (2) due to deviations from a perfect positive correlation (IC) of genotypic rankings among environments (Method 1) or of environmental ranking among genotypes (Method 2), as shown in Table 2.

Partitioning interaction into two types within genotypes

With fixed genotypes or lines and random environmental effects, further partitioning of the interaction into that due to each

Table 2. Sum of squares of alternative partitioning of the genotype-by-environment interaction

Type of interaction	Method 1[a]	Method 2[b]
Heterogeneous variances: SS (HV)	$n\sum_{j}^{e}(Z_j - \bar{Z}_.)^2$	$n\sum_{i}^{g}(S_i - \bar{S}_.)^2$
Imperfect correlation: SS (IC)	$2n\sum_{j<j'}^{e}(1 - r_{jj'})[Z_jZ_{j'}]/e$	$2n\sum_{i<i'}^{g}(1 - r_{ii'})S_iS_{i'}/g$
Total: SS (G × E)	$n\sum_{i}^{g}\sum_{j}^{e}(\bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...})^2$	$n\sum_{i}^{g}\sum_{j}^{e}(\bar{Y}_{ij.} - \bar{Y}_{i..} - \bar{Y}_{.j.} + \bar{Y}_{...})^2$

[a] $Z_j^2 = \sum_{i}^{g}(\bar{Y}_{ij.} - \bar{Y}_{.j.})^2$, $Z_{jj'} = \sum_{i}^{g}(\bar{Y}_{ij.} - \bar{Y}_{.j.})(\bar{Y}_{ij'.} - \bar{Y}_{.j'.})$, $r_{jj'} = Z_{jj'}/[Z_jZ_{j'}]$

[b] $S_i^2 = \sum_{j}^{e}(\bar{Y}_{ij.} - \bar{Y}_{i..})^2$, $S_{ii'} = \sum_{j}^{e}(\bar{Y}_{ij.} - \bar{Y}_{i..})(\bar{Y}_{i'j.} - \bar{Y}_{i'..})$, $r_{ii'} = S_{ii'}/[S_iS_{i'}]$

Table 3. Sum of squares for partitioning of the genotype-by-environment interaction sum of squares into types of interaction within fixed genotypes based on Method 2

Genotype i	Type of interaction[a]		Total
	Heterogeneous variances $SS(HV)_i$	Imperfect correlations $SS(IC)_i$	
1	$n\sum_{i'\neq 1}(S_1-S_{i'})^2/2g$	$n\sum_{i'\neq 1}(1-r_{1i'})S_1S_{i'}/g$	$SS(G\times E)_1$
2	$n\sum_{i'\neq 2}(S_2-S_{i'})^2/2g$	$n\sum_{i'\neq 2}(1-r_{2i'})S_2S_{i'}/g$	$SS(G\times E)_2$
⋮			
i	$n\sum_{i'\neq i}(S_i-S_{i'})^2/2g$	$n\sum_{i'\neq i}(1-r_{ii'})S_iS_{i'}/g$	$SS(G\times E)_i$
⋮			
g	$n\sum_{i'\neq g}(S_g-S_{i'})^2/2g$	$n\sum_{i'\neq g}(1-r_{gi'})S_gS_{i'}/g$	$SS(G\times E)_g$
Total	$SS(HV)$	$SS(IC)$	$SS(G\times E)$

[a] $S_i^2=\sum_j^e(\bar{Y}_{ij.}-\bar{Y}_{i..})^2$ and $r_{ii'}=\sum_j^e(\bar{Y}_{ij.}-\bar{Y}_{i..})(\bar{Y}_{i'j.}-\bar{Y}_{i'..})/S_iS_{i'}$

genotype and type of interaction within genotypes is desirable to compare their relative stabilities.

For Method 2, the interaction sum of squares associated with any given pair of genotypes is:

$$SS(G\times E)_{ii'}=n[(S_i-S_{i'})^2+2(1-r_{ii'})S_iS_{i'}]/g. \quad (4)$$

This sum of squares can be further partitioned into parts associated with heterogeneous variance, $SS(HV)_{ii'}$, and that due to imperfect positive correlation of the pair, $SS(IC)_{ii'}$, i.e.,

$$SS(HV)_{ii'}=n(S_i-S_{i'})^2/g \quad (5)$$

$$SS(IC)_{ii'}=2n(1-r_{ii'})S_iS_{i'}/g \quad (6)$$

The sum of squares for the ineraction of the *i*th genotype with all other genotypes is:

$$SS(G\times E)_i=\sum_{i'\neq i}^{g}SS(G\times E)_{ii'}=\sum_{i'\neq i}^{g}SS(HV)_{ii'}+\sum_{i'\neq i}^{g}SS(IC)_{ii'} \quad (7)$$
$$=SS(HV)_i+SS(IC)_i$$

Thus, each genotype interaction sum of squares can be partitioned into that due to heterogeneous variances, $SS(HV)_i$, and that due to imperfect correlations, $SS(IC)_i$, as shown in Table 3.

The equivalent partitioning for Method 1 would separate the interaction into that due to each environment, but is generally not useful.

Discussion

Because the partitions of the sum of squares were not constructed as the sum of squared deviations, the partitions are not distributed as chi-squares, and test of significance have not been developed. Nevertheless, partitioning of the interaction is important to conceptualize the interaction and to provide breeders with information necessary for deciding on alternative breeding strategies. For example, if the nature of the interaction is due to differences in scaling only, the interaction is not important to the breeder.

Partitioning interaction into two types

With Method 2, Moll et al. (1978) refers to the first partition as that due to differential responsiveness or differential environmental sensitivity among the entires, while the second partition is due to differences in correlation among pairs of entries. In contrast, with Method 1 the first partition is associated with the degree of interaction due to scale, while the second partition is associated with the genetic correlation of the same trait measured in alternative environments. The problem of which method to use depends on the question being asked. The types of questions being addressed by breeders can generally be grouped by which factors are fixed or random.

Genotypes fixed, environments random. If genotypes are fixed and environments are random, the problem is either (1) to identify the genotype that is most 'stable' or (2) to choose the genotype that gives optimal performance averaged across the sampling space of environments, or both. If the objective is only the latter, the solution is simple, i.e., choose the genotype that gives optimal performance averaged across environments. However, for the following discussion, it will be assumed that the breeder is also interested in the more complex problem of stability. The merit of stability is particularly important in lesser developed countries where highest average yield may not be as important as having a 'guaranteed' yield. In that situation, one attempts to draw conclusions regarding stability of production of fixed genotypes in all possible environments in the sampling space from which the sample environments was derived. For example, genotypes may be exposed to a sample of environmental effects, such as locations or years, but the inference space is to all locations or years in the region of interest.

If the interaction is significant, the scaling partition of Method 2 indicates the magnitude of the interaction due to the differential environmental sensitivity of the genotypes. Neither partition by Method 1 relates to this question. Thus, partitioning by Method 2 is more appropriate for this situation. However, with more than two genotypes, neither method of partitioning is particularly useful since recommendations among the fixed genotypes cannot be made. Further partitioning of each type into that due to each genotype is required, as discussed later.

Genotypes random, environments fixed. If genotypes are random and environments are fixed, the problem is to determine if there is genetic variability for adaptation to specific environments. In this case, the first partition of Method 2 simply indicates the presence of differential genetic variability for environmental sensitivity, while the

second partition gives little useful information. In contrast, with Method 1, if the majority of the interaction is due to the first partition, i.e., variation among environments in the scaling of genetic effects, the interaction is generally considered to be unimportant since the ranks of the genotypes remain constant across environments. But, if the interaction is mainly due to the second partition, i.e., deviations from a perfect positive correlation of genotypic ranking among environments, a specific breeding recommendation can be made. Specialized genotypes or lines should be developed for each environment since re-ranking of genotypes in alternative environments is occurring. Thus, partitioning by Method 1 is more appropriate for this situation.

Genotypes random, environments random. With both factors random, only a general breed or single cultivar can be developed since a fixed environment cannot be provided. A general breed could be developed by measuring individuals in a random sample of environments from the inference space of environments. Since such a program is costly, the problem is to determine if selection based on performance in one environment is adequate to improve average performance across environments or, alternatively, can a general breed or cultivar be developed that will perform well over the inference space of environments by selecting in just one environment? If an interaction is due only to heterogeneous scaling of genetic effects, the best genotype will remain the best in all environments since re-ranking of genotypes generally did not occur in alternative environments, and thus performance needs to be measured only in one environment. However, if the correlation of genotypes among environments is seriously imperfect, the evaluation of genotypes averaged over a random sample of environments is essential, as stressed by Dickerson (1962) and Yamada (1962). Partitioning by Method 1 provides a breakdown based on these criteria, whereas Method 2 does not.

Genotypes fixed, environments fixed. If both factors are fixed, the solution is simple: choose the genotype that gives the optimal performance in each specific environment.

Thus, different questions are addressed with the alternative methods, differential environmental sensitivity of genotypes as opposed to differential scale of the genotypes in different environments. Partitioning by Method 2 is more appropriate for addressing questions posed with fixed genotypes, while Method 1 is more appropriate for random genotypes.

Numerical examples. Comparison of analysis and interpretation based on each method will be based on two hypothetical sets of data. Computations are simplified by calculating the total sum of squares for interaction and the sum of squares for the first partition. The sum of squares for the second partition is then found by subtraction.

Hypothetical data, for which $n=1$, are presented in Table 4 for two cases: (1) ranks of genotypes remain the same, but the magnitude of differences between genotypes differs among environments; (2) ranks and magnitude of differences between genotypes change among environments. Analysis for each case and method is presented in Table 5. Clearly, each method gives greatly different results. Note particularly the reversal in sum of squares between methods for Case 1. Interpretation is also entirely different for each method. For Case 1, both genotypes show response to the environment, but in opposite directions. Thus, partitions based on Method 2 indicate equal responsiveness, with all of the interaction

Table 4. Hypothetical data for two types of genotype-by-environment interactions

Case	Genotype	Environment				
		1	2	3	4	5
1[a]	A	10	11	12	13	14
	B	10	9	8	7	6
2[b]	A	8	9	10	11	12
	B	12	11	10	9	8

[a] The ranks of the genotypes remain the same, but the magnitude of the differences between the genotypes differs among environments

[b] The ranks and magnitude of the genotypes change among environments

Table 5. Sums of squares from alternative partitioning of the genotype-by-environment interaction for example data[a]

Type of interaction	Case			
	1		2	
	Method 1	Method 2	Method 1	Method 2
Heterogeneous variances	20.0	0.0	5.6	0.0
Imperfect correlation	0.0	20.0	14.4	20.0
Total	20.0	20.0	20.0	20.0

[a] Method 1 is based on heterogeneity among environments in the scaling of differences among genotypes; Method 2 is based on heterogeneity among genotypes in the scaling of differences among environments. The remaining interaction is ascribed to a deviation from a perfect positive correlation of genotype rankings among environments (Method 1) or of environmental rankings among genotypes (Method 2)

due to imperfect correlations of environmental rankings among genotypes, i.e., environmental rankings of the two genotypes are completely reversed. However, Method 2 also gives the same result for Case 2, even though the genotypes change in ranks within environments. This existed because the responsiveness of the genotypes to environments was still the same. In contrast, Method 1 gives different results for the two cases. In the first case, all of the interaction is due to heterogeneous variances, and a correct conclusion is that no re-rankings of genotypes across environments occurred. In the second case, Method 1 gives a mixed signal, with the majority (72%) of the interaction being attributed to imperfect correlations. This result occurred because genotypic ranks changed in all but Environment 3. These examples clearly illustrate that, dependent on the purpose of the research, correct partitioning and interpretation must be used.

Partitioning interaction into two types within genotypes

As discussed previously, with genotypes fixed and environments random, the problem is to identify 'stable' genotypes. Method 2 is appropriate in this situation but, for more than two genotypes, simple partitioning of the interaction into two types is not very useful since recommendations among the fixed genotypes cannot be made. Further partitioning of the interaction into that due to genotype and type of interaction within each genotype is needed.

Many procedures for the analysis of fixed genotypes with respect to their stability have been developed (Kang 1990; Lin et al. 1986; Lin and Butler 1990). Graphical methods of classifying cultivar responsiveness to environments were presented by Mooers (1921). Yates and Cochran (1938), Finlay and Wilkinson (1963), and Moll et al. (1978) developed statistical methods to facilitate those comparisons. Their procedure is to use the mean performance of all genotypes in that environment, called the environmental or site value, as a quantitative measure of the relative productivity of that environment. The linear regression coefficient of observed means, for any given genotype, on site values is used as a measure of that genotype's sensitivity, and the overall mean indicates general adaptability. Although the genotype-by-environment interaction is also partitioned into parts with their procedure, that associated with linear regression and residual, the utility of those partitions is to determine if the interaction can be characterized by linear associations. If deviations from the linear regression are significant, the relationship between yield and site value is nonlinear for some genotypes. In that case, the comparison of linear regression coefficients may give erroneous conclusions. Partitioning of the interaction in that manner is not related to the partitions presented in this paper. Note also that analysis of either case given in Table 4 by the regression method of Finlay and Wilkinson (1963) yields indeterminate results because all site values are the same.

Lin et al. (1986) noted at least three concepts of stability. Of these, they recommended either of the following definitions. (1) A genotype is considered stable if its response to environmental effects is parallel to the mean response of all genotypes in the trial, i.e., low $SS(G \times E)_i$. (2) A genotype is stable if the among-environment variance is small, i.e., $S_i^2/(e-1)$ is small.

The interaction sum of squares due to the ith genotype, $SS(G \times E)_i$, measures stability based on the first definition. Although formulated differently, this estimate is equivalent to the statistic given by Wricke (1962). However, such statistics are only pertinent relative to other genotypes included in the study.

Similarly, the sum of squares, $SS(G \times E)_{ii'}$, is an effective measure of similarity between pairs. If $SS(G \times E)_{ii'} = 0$, the responses of the genotypes parallel each other. As $SS(G \times E)_{ii'}$ becomes larger, the responses become more dissimilar. This indicator is equivalent to the coefficient given by Abou-El-Fittouh et al. (1969).

The decomposition of $SS(G \times E)_{ii'}$ into that due to $SS(HV)_{ii'}$ and $SS(IC)_{ii'}$ is convenient for multi-criteria clustering procedures. Lefkovitch (1985) suggested a similar, though nonorthogonal, decomposition into that due to pattern distance, $2(1-r_{ii})$, and Frechet distance $(\bar{Y}_{i..} - \bar{Y}_{i'..}) + (S_i - S_{i'})$ as defined by Dowson and Landau (1982). Lefkovitch (1985) used these measures to from clusters of common genotypes.

Partitioning of the interaction sum of squares by genotype and type of interaction is desirable because the $SS(IC)_i$ statistic is a coefficient of stability. Because the statistic incorporates both environmental correlations between genotypes, $r_{ii'}$, and their environmental variances, $\sigma_i^2 = S_i^2/(e-1)$, the statistic includes aspects of both concepts of stability recommended by Lin et al. (1986). If either $r_{ii'} = 1$ or $\sigma_i^2 = 0$ then $SS(IC)_i = 0$ and the genotype is considered stable. As $r_{ii'}$ declines or σ_i^2 increases, $SS(IC)_i$ will increase. This result is particularly true for negative values of $r_{ii'}$. Thus, $SS(IC)_i$ is a more informative indicator of nonstability than $SS(G \times E)_i$ because the latter is a measure of nonparallelism due to any

Table 6. Performance of six genotypes reared in five random environments

Genotype	Environment				
	1	2	3	4	5
1	3.00	4.25	5.50	6.75	8.00
2	2.00	3.60	5.00	6.50	8.00
3	1.00	2.75	4.50	6.25	8.00
4	6.00	5.75	5.50	5.25	5.00
5	5.00	5.00	5.00	5.00	5.00
6	4.00	4.25	4.50	4.75	5.00

Table 7. Sums of squares for partitioning of interaction sum of squares of performance data given in Table 6 into types within crosses and common genotypes of crosses

Genotype		Type of interaction				Total		Environmental variance
		Heterogeneous variances		Imperfect correlations				
i	i'	$SS(HV)_{ii'}$	$SS(HV)_i$	$SS(IC)_{ii'}$	$SS(IC)_i$	$SS(G \times E)_{ii'}$	$SS(G \times E)_i$	σ_i^2
1	2	0.052		0.000		0.052		
1	3	0.208		0.000		0.208		
1	4	0.833		1.042		1.875		
1	5	1.302		0.000		1.302		
1	6	0.833	3.228	0.000	1.042	0.833	4.270	3.906
2	1	0.052		0.000		0.052		
2	3	0.052		0.000		0.052		
2	4	1.302		1.250		2.552		
2	5	1.875		0.000		1.875		
2	6	1.302	4.583	0.000	1.250	1.302	5.833	5.625
3	1	0.208		0.000		0.208		
3	2	0.052		0.000		0.052		
3	4	1.875		1.458		3.333		
3	5	2.552		0.000		2.552		
3	6	1.875	6.562	0.000	1.458	1.875	8.020	7.656
4	1	0.833		1.042		1.875		
4	2	1.302		1.250		2.552		
4	3	1.875		1.458		3.333		
4	5	0.052		0.000		0.052		
4	6	0.000	4.062	0.208	3.958	0.208	8.020	0.156
5	1	1.302		0.000		1.302		
5	2	1.875		0.000		1.875		
5	3	2.552		0.000		2.552		
5	4	0.052		0.000		0.052		
5	6	0.052	5.833	0.000	0.000	0.052	5.833	0.000
6	1	0.833		0.000		0.833		
6	2	1.302		0.000		1.302		
6	3	1.875		0.000		1.875		
6	4	0.000		0.208		0.208		
6	5	0.052	4.062	0.000	0.208	0.052	4.270	0.156
Total			28.330		7.916		36.246	

cause while the former is particularly associated with a genotype that responds in a different direction to common environmental effects (negative $r_{ii'}$) or which is highly sensitive to environmental effects (large σ_i^2). See Freeman (1973) or Lin et al. (1986) for other methods of sensitivity analysis.

Numerical example. Hypothetical data, given in Table 6, will be used to demonstrate the utility of this partitioning. The example was devised such that genotypes 1, 2, and 3 were equally well adapted to good environments but varied somewhat in their response to poor environments, with genotype 3 being the most diminished by poor environments. Genotypes 4, 5, and 6 were moderately adapted to all environments. Genotype 4 was the only line that was better adapted to poor than good environments. Thus, genotypes 3 and 4 were the most different in their responses.

Partitioning for each genotype and pair is given in Table 7. From the marginal values, $SS(G \times E)_i$, genotypes 3 and 4 were identified as least stable, and genotypes 1 and 6 were identified as most stable. However, these comparisons failed to show that genotype 4 responded in the opposite direction to all others and that genotype 5 was in fact the most stable since it was invariant. Comparisons based on environmental variability, σ_i^2, identified genotype 5 as the most stable but also gave identical low rankings for genotypes 4 and 6. In contrast, comparisons based on $SS(IC)_i$ were more informative. Genotype 4 was identified as most different in its response while genotype 5 was identified as most stable. These results illustrate the inadequacies of other single

alternative definitions of stability as discussed by Lin et al. (1986).

Biological examples

An example appropriate for method 1. An example of an experiment with genotypes random and environments fixed was reported by Muir (1986). In that experiment, genotype-by-production environment interaction in poultry was examined. Commercial breeders of poultry make selections based on production of birds housed in single-bird cages, while producers house improved birds in multiple-bird cages. Due to competitive effects that exist in multiple-bird cages, but not in single-bird cages, the potential exists for a strong genotype-by-environment interaction to develop. Because the inference is only to the two known environmental effects, single- versus multiple-bird (nine-bird) cages, environments are fixed. Genotypes consisted of sire families that were split and reared in each environment. Since differences between genotypes were random, the inference space was to all genotypes of that breed, and the effect is considered random. In that paper, a genotype-by-cage environment interaction was reported for the trait 'days survival', which is the number of days a bird survives from day of housing. However, whether the interaction was due to a re-ranking of genotypes or simply due to a change in the variance was not determined. This interaction was re-analyzed and partitioned by Method 1 (Table 8).

The outcome shows that 63% of the interaction was due to imperfect correlations. Thus re-ranking of genotypes occurred in the two environments. Although a test of significance is not possible, the data strongly suggest that breeders should evaluate performance in multiple-bird cages to improve performance of that trait in production environments.

An example appropriate for method 2. An example of an experiment with fixed genotypes and random environments was provided by the classical experiment of Yates and Cochran (1938). In that report, five barley cultivars were grown at six experiment stations over 2 years. At each location and year, three samples were taken. While locations and years were both random environmental effects, the results were summed, for simplicity, over samples and years (Table 9). A significant location-by-cultivar interaction was detected. Analysis by Methods 1 and 2, for comparison, are given in Table 10. Analysis by Method 2 indicates that approximately equal proportions of the interaction sum of squares are due to differential sensitivity and imperfect correlations. In contrast, results from Method 1 indicates that the majority of the interaction (78%) is due to imperfect correlations. These results infer that, if possible, major environmental effects should be identified. Thereby, the dominant part of the environmental effect becomes fixed, and specialized cultivars could be developed for each.

However, if specialized features of each location cannot be identified, the next best procedure is to characterize each cultivar as to stability and adaptability. Partitioning by each genotype, as shown in Table 11, is needed. From comparison of $SS(G \times E)_i$, cvs 'Trebi' and 'Peatland' were least stable based on the first definition of stability. In contrast, comparisons based on $SS(IC)_i$ show that 'Peatland' was the most stable while 'Trebi' remained the least stable. Examination of the data verify these conclusions. 'Peatland' was best adapted to poor environments and did not show corresponding improvements with better environments. Thus, 'Peatland' was stable in the sense of not being sensitive to changing environments.

Table 8. Partitioned sum of squares, based on Method 1, of genotype-by-environment interaction for poultry data summarized by Muir (1986)[a]

Type of interaction	Sum of squares
Heterogeneous variances	159.9
Imperfect correlation	269.1
Total	429.0

[a] Method 1 is based on the variance among genotypes within the *j*th environment

Table 9. Yields of five barley genotypes in six environments (locations) summed over 2 years and three plots[a]

Genotype	Environment					
	1	2	3	4	5	6
Manchuria	161.7	247.0	185.4	218.7	165.3	154.6
Svansota	187.7	257.5	182.4	183.3	138.9	143.8
Velvet	200.1	262.9	194.9	220.2	165.8	146.3
Trebi	196.9	339.2	271.2	266.3	151.2	193.6
Peatland	182.5	253.8	219.2	200.5	184.4	190.1

[a] Summarized from Yates and Cochran (1938)

Table 10. Sums of squares from partitioning of the genotype-by-environment interaction for barley data into types of interaction[a]

Type of interaction	Method 1	Method 2
Heterogeneous variances	318.0	771.8
Imperfect correlation	1,159.7	705.8
Total	1,477.7	1,477.7

[a] Method 1 is based on the variance among genotypes within the *j*th environment. Method 2 is based on the variance among environments within the *i*th genotype

Table 11. Sum of squares for partitions of the genotype-by-environment interaction within genotype into types of interaction by Method 2 for barley data

Genotype	Type of interaction		Total	Environmental variances
	Heterogeneous variances $SS(HV)_i$	Imperfect correlations $SS(IC)_i$	$SS(G \times E)_i$	σ_i^2
Manchuria	94.6	128.9	223.5	1,350
Svansota	77.4	136.8	214.2	1,810
Velvet	79.1	139.7	218.8	1,686
Trebi	339.5	183.5	523.0	4,665
Peatland	181.2	117.0	298.2	751
Total	771.8	705.9	1,477.7	

Comparisons based on environmental variance, σ_i^2, gave similar rankings as $SS(IC)_i$. Because all of the cultivars responded in a similar direction to alternative environmental factors, differences in correlations between pairs is not a major factor, and hence the main influence on $SS(IC)_i$ is due to differences in σ_i^2.

'Trebi' was best adapted to good environments, but also showed the greatest sensitivity, thus 'Trebi' was considered unstable. However, if environments were a random sample of all possible environments in which the cultivars could be grown, the cultivars would have been planted in proportion to the expected environmental exposure. In that case, the best all-around cultivar is simply the one that gives the highest average yield; in this case, 'Trebi'. However, the highest average yield may not be as important as having a 'guaranteed' yield, i.e., a stable cultivar. This property is particularly true in lesser developed countries.

References

Abou-El-Fittouh HA, Rawlings JO, Miller PA (1969) Classification of environments to control genotype by environment interactions with an application to cotton. Crop Sci 9:135–140

Casella G, Berger RL (1990) Statistical inference, Wadsworth and Brooks/Cole, Pacific Grove, Calif.

Cockerham CC (1963) Estimation of genetic variances. In: Hanson WD, Robinson HF (eds) Statistical genetics and plant breeding. NAS-NRC, Washington, D.C., pp 53–94

Dickerson GE (1962) Implications of genetic-environmental interaction in animal breeding. Anim Prod 4:47–63

Dowson DC, Landau BV (1982) The Fréchet distance between multivariate normal distributions. J Multivar Anal 12:450–455

Eisen EJ, Saxton AM (1983) Genotype by environment interactions and genetic correlations involving two environmental factors. Theor Appl Genet 67:75–86

Finlay KW, Wilkinson GN (1963) The analysis of adaptation in a plant-breeding programme. Aust J Agric Res 14:742–754

Freeman GH (1973) Statistical methods for the analysis of genotype-environment interactions. Heredity 31:339–354

Kang MS (1990) (ed) Genotype-by-environment interaction and plant breeding. Department of Agronomy, Louisiana Agricultural Experiment Station, Louisiana State University Agricultural Center, Baton Rouge, LA 70803-2110

Lefkovitch LP (1985) Multi-criteria clustering in genotype-environment interaction problems. Theor Appl Genet 70:585–589

Lin CS, Butler G (1990) Cluster analyses for analyzing two-way classification data. Agron J 82:344–348

Lin CS, Binns MR, Lefkovitch LP (1986) Stability analysis: Where do we stand? Crop Sci 26:894–900

Muir WM (1986) Relative efficiency of selection for performance of birds housed in colony cages based on production in single bird cages. Poult Sci 64:2239–2247

Moll RH, Cockerham CC, Stuber CW, Williams WP (1978) Selection responses, genetic-environmental interactions, and heterosis with recurrent selection for yield in maize. Crop Sci 18:641–645

Mooers CA (1921) The agronomic placement of varieties. J Am Soc Agron 13:337–352

Robertson A (1959) The sampling variance of the genetic correlation coefficient. Biometrics 15:469–485

Wricke G (1962) Über eine Methode zur Erfassung der ökologischen Streubreite in Feldversuchen. Z Pflanzenzücht 47:92–96

Yamada Y (1962) Genotype by environment interaction and genetic correlation of the same trait under different environments. Jpn J Genet 37:498–509

Yamada Y, Itoh Y, Sugimoto I (1988) Parametric relationships between genotype × environment interaction and genetic correlation when two environments are involved. Theor Appl Genet 76:850–854

Yates F, Cochran WG (1938) The analysis of groups of experiments. J Agric Sci 28:556–580