Theor Appl Genet (1992) 84:161–172

Using the shifted multiplicative model to search for "separability" in crop cultivar trials*

P. L. Cornelius[1,**], **M. Seyedsadr**[2], **and J. Crossa**[3]

[1] Departments of Agronomy and Statistics, University of Kentucky, Lexington, KY 40546-0091, USA
[2] Gershenson Radiation Oncology Center, School of Medicine, Wayne State University, 4201 St. Antoine, Detroit, MI 48201, USA
[3] International Maize and Wheat Improvement Center (CIMMYT), Apdo. Postal 6-641, 06600 Mexico, D.F., Mexico

Received August 23, 1991; Accepted October 9, 1991
Communicated by A. R. Hallauer

Summary. The shifted multiplicative model (SHMM) is used in an exploratory step-down method for identifying subsets of environments in which genotypic effects are "separable" from environmental effects. Subsets of environments are chosen on the basis of a SHMM analysis of the entire data set. SHMM analyses of the subsets may indicate a need for further subdivision and/or suggest that a different subdivision at the previous stage should be tried. The process continues until SHMM analysis indicates that a SHMM with only one multiplicative term and its "point of concurrence" outside (left or right) of the cluster of data points adequately fits the data in all subsets. The method is first illustrated with a simple example using a small data set from the statistical literature. Then results obtained in an international maize (*Zea mays* L.) yield trial with 20 sites and nine cultivars is presented and discussed.

Key words: Genotype × environment interaction – Shifted multiplicative model – Separability – Concurrent regression model – Crossover interaction – Qualitative interaction

Introduction

The concept of "separability" of genotypic and environmental effects in crop cultivar trials repeated over a series of environments (e.g., locations and/or years) was developed by Gregorius and Namkoong (1986). One possible "operator" which they proposed for mapping of a set of genotypic and environmental effects onto a set of phenotypic response values (formally denoted Ω: $C_\gamma \times C_\varepsilon \to F$) is the multiplicative operator $a(\gamma)\, b(\varepsilon) + c$ where $a(\gamma)$ and $b(\varepsilon)$ are functions of genotypic effects and environmental effects, respectively, and c is a constant. This operator is a special case (in that it has only one multiplicative term) of the shifted multiplicative model (SHMM) for which we have developed some statistical and computational procedures in previous papers (Seyedsadr and Cornelius 1991 b, c). We write SHMM with t multiplicative terms (SHMM_t, say) in the form

$$y_{ij} = \beta + \sum_{k=1}^{t} \lambda_k \alpha_{ik} \gamma_{jk} + e_{ij} \qquad (1)$$

where β is the "shift parameter", α_{i1} and γ_{j1} are "primary effects" of i^{th} environment and j^{th} cultivar, respectively, α_{i2} and γ_{j2} are their "secondary effects", α_{i3} and γ_{j3} their "tertiary effects", etc. The ε_{ij} quantities are random errors, which we shall suppose to be NID$(0, \sigma^2)$. The quantities λ_k $(\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_t > 0)$ are scaling constants that allow us to impose the orthonormality constraints, $\Sigma_i \alpha_{ik}^2 = \Sigma_j \gamma_{jk}^2 = 1$ and $\Sigma_i \alpha_{ik} \alpha_{ik'} = \Sigma_j \gamma_{jk} \gamma_{jk'} = 0$ for $k \neq k'$. If $t = 1$, the right-hand side of (1), apart from the e_{ij} term, is Gregorius and Namkoong's multiplicative operator if we put $c = \beta$ and, for the i^{th} environment and j^{th} genotype, $a(\gamma) = \lambda_1^w \gamma_{j1}$ and $b(\varepsilon) = \lambda_1^x \alpha_{i1}$ where $w + x = 1$.

If SHMM_1 (apart from the e_{ij} term) is plotted against either the geontypic primary effect (γ_{j1} values) or the environmental primary effects (α_{i1} values), the resulting graph has the configuration which Mandel (1961) has characterized as "concurrent" regression lines, so-called because the regression lines all intersect at one point (the "point of concurrence"). If the phenotypic values are located entirely to the left or to the right of the point of concurrence when plotted against environmental primary effects, then genotypic effects are "separable from envi-

* Journal Article No. 91-3-171 of the Kentucky Agricultural Experiment Station published with the approval of the Director
** Correspondence to: P. L. Cornelius

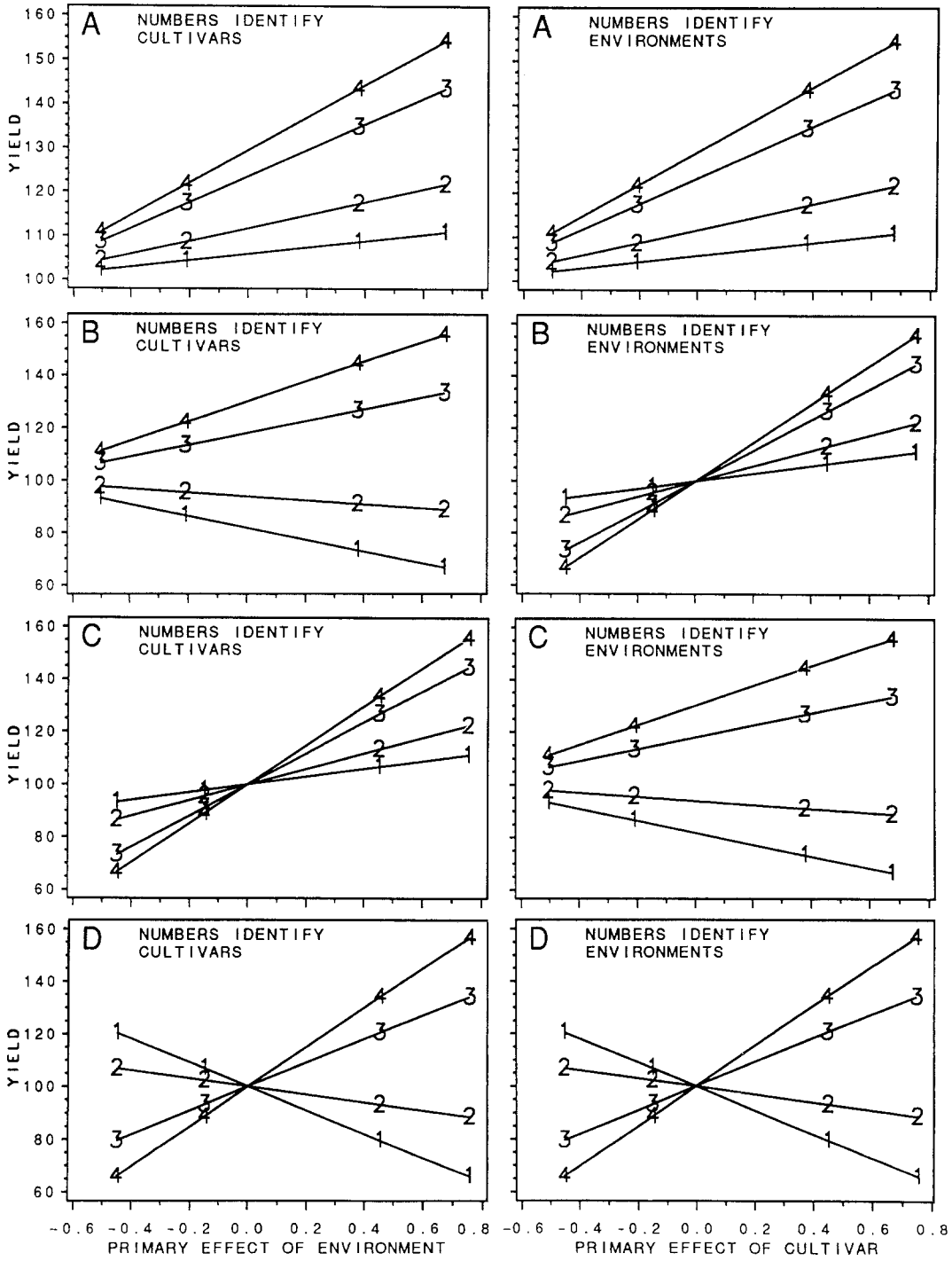

Fig. 1A–D. Four hypothetical $SHMM_1$ response patterns: **A** complete separability; **B** genotypic effects separable from environmental effects; **C** environmental effects separable from genotypic effects; **D** genotypic and environmental effects inseparable

ronmental effects" by the Gregorius and Namkoong definition. If such a configuration is obtained when phenotypic values are plotted against genotypic primary effects, then environmental effects are "separable from genotypic effects". Figure 1 displays four hypothetical cases. In Case A there is complete separability. Cases B and C each show one type of separability, but not the other. Neither type of separability is found in Case D. Note that Case A, complete separability, can also have a graphical configuration with the projected point of concurrence to the upper-right of the graphs rather than to the lower-left. Graphs displaying separability in Cases B and C can also be such that the point of concurrence is to the right rather than to the left of the graph.

We will be particularly interested in Cases A and B as these are response patterns in which there are no genotypic rank-change interaction, and if the yield trial data do not fit such a pattern, then we will search for subsets of environments in which such a response pattern is obtained. This will require that $SHMM_1$ give a satisfactory fit to the data and that the estimated primary effects of environments are all of the same sign. On the plot of regression lines for genotypes against environmental primary effects, the point of concurrence is to the left of the graph if the environmental effects are all positive, but to the right if they are all negative. Since the products $\alpha_{i1}\,\gamma_{j1}$ are invariant under a simultaneous sign change, the polarity of the horizontal axes on the graphs is arbitrary apart from simultaneous reversal in each matching pair of graphs. We shall adopt the convention that the polarity will be chosen such that environmental primary effects are positively correlated with the environmental means, $\bar{y}_{i.}$. Note that this will make it impossible for the point of concurrence to be located either to the upper-left or the lower-right of the graph.

For brevity, we will shorten expressions such as "separability of genotypic effects from environmental effects" to simply "separability of genotypic effects" or "genotypic separability".

Theory of SHMM

Seyedsadr and Cornelius (1991 c) derived the least squares estimate of the shift parameter $\hat{\beta}$ as

$$\hat{\beta} = y_{..} - \sum_{k=1}^{t} \hat{\lambda}_k \hat{\bar{\alpha}}_k \hat{\bar{\gamma}}_k$$

where $\hat{\bar{\alpha}}_k = r^{-1} \sum_{i=1}^{r} \hat{\alpha}_{ik}$ and $\hat{\bar{\gamma}}_k = c^{-1} \sum_{j=1}^{c} \hat{\gamma}_{jk}$; r and c are the row and column dimensions of the two-way table of data. A well-known result in matrix algebra is that any matrix $\boldsymbol{Z}$ can be expressed as $\boldsymbol{Z} = \hat{\boldsymbol{A}}\hat{\boldsymbol{\Lambda}}\hat{\boldsymbol{G}}'$, where $\hat{\boldsymbol{\Lambda}}$ is a diagonal matrix with positive numbers, $\hat{\lambda}_k$, on the diagonal, and $\hat{\boldsymbol{A}}$ and $\hat{\boldsymbol{G}}$ are such that $\hat{\boldsymbol{A}}'\hat{\boldsymbol{A}} = \hat{\boldsymbol{G}}'\hat{\boldsymbol{G}} = \boldsymbol{I}_m$ where $m = \text{rank}(\boldsymbol{Z})$. Denoting the k^{th} column of $\hat{\boldsymbol{A}}$ and $\hat{\boldsymbol{G}}$, respectively, by $\hat{\boldsymbol{\alpha}}_k$ and $\hat{\boldsymbol{\gamma}}_k$, matrix $\boldsymbol{Z}$ can be expressed as $\boldsymbol{Z} = \hat{\boldsymbol{A}}\hat{\boldsymbol{\Lambda}}\hat{\boldsymbol{G}}' = \sum_{k=1}^{m} \hat{\lambda}_k \hat{\boldsymbol{\alpha}}_k \hat{\boldsymbol{\gamma}}_k'$ and an individual element of $\boldsymbol{Z}$ as $z_{ij} = \sum_{k=1}^{m} \hat{\lambda}_k \hat{\alpha}_{ik} \hat{\gamma}_{jk}$. The quantities $\hat{\lambda}_k$, $\hat{\alpha}_{ik}$, $\hat{\gamma}_{jk}$ for $k = 1, \ldots, t$ are the least squares estimates of our model parameters λ_k, α_{ik} and γ_{jk} if we define $z_{ij} = y_{ij} - \hat{\beta}$ and the terms in $\sum_k \hat{\lambda}_k \hat{\alpha}_{ik} \hat{\gamma}_{jk}$ are arranged in decreasing order with respect to the $\hat{\lambda}_k$ values. In matrix algebra, the decomposition $\boldsymbol{Z} = \hat{\boldsymbol{A}}\hat{\boldsymbol{\Lambda}}\hat{\boldsymbol{G}}'$ is known as the "singular value decomposition" (SVD) of $\boldsymbol{Z}$. The $\hat{\lambda}_k$ are known as "singular values" and the vectors $\hat{\boldsymbol{\alpha}}_k$ anf $\hat{\boldsymbol{\gamma}}_k$ as "left singular vectors" and "right singular vectors", respectively. Briefly stated, we say that the least squares estimates of the λ_k, α_{ik} and γ_{jk} for $k = 1, \ldots, t$ are given by the "first t components" of the SVD of $\boldsymbol{Z} = \boldsymbol{Y} - \hat{\beta}\boldsymbol{J}$ where $\boldsymbol{J}$ is a $r \times c$ matrix of ones. Since $\boldsymbol{Z}$ depends on $\hat{\beta}$, and $\hat{\beta}$ depends on the SVD of $\boldsymbol{Z}$, the least squares fitting of SHMM requires an iterative algorithm. Changing the number of multiplicative terms changes $\hat{\beta}$ and, consequently, changes the estimates of all λ_k, α_{ik} and γ_{jk}. Computational algorithms are described by Cornelius and Seyedsadr (1991 c), and it is expected that the first author's Fortran program, EIGAOV, can be made available in the near future.

By computer simulation, Seyedsadr (1987) estimated the expected values and sampling standard deviations of the sum of squares owing to the sequential addition of multiplicative terms to the model when the data are two-way tables of standard normal deviates with row and column dimensions ($r \times c$) ranging from 3×2 up to 99×19. These results are analogous to results obtained by Mandel (1971) for models in which the estimates of the linear terms in the model do not depend on estimates of the multiplicative terms. [The additive main effects and multiplicative interaction (AMMI) model (Gauch 1988; Gauch and Zobel 1988) is one such model.] Results for the first three multiplicative terms were reported by Seyedsadr and Cornelius (1991 c) and functions which approximate these results for the first five terms as functions of r and c were obtained by Seyedsadr and Cornelius (1991 a) using regression techniques. It is suggested that sequential sums of squares (S_k, say, for the k^{th} multiplicative term) be presented in an analysis of variance (anova) format using the computer simulation results for the expectations, or the functions which approximate these results, as the degrees of freedom (df) for the purpose of computing mean squares.

Test of statistical significance

Since we will be concerned with the number of multiplicative terms necessary to describe the variation in a particular data set and also whether $SHMM_1$ will be considered an adequate model for a given subset of environments, we will now describe some statistical tests which can be used to address these questions.

Since the S_k are not independent and are not distributed as chi-square random variables, the usual linear model theory leading to F-statistics for inference purposes does not hold. For the case where no estimate of σ^2 independent of the cell means is available, Seyedsadr and Cornelius (1991 b) obtained $\Lambda_t = S_t / \sum_{k=t}^{p} S_k$, where $p = \min(r-1, c)$, provided $c \leq r$ (if $c > r$, interchange r and c), as an appropriate statistic for judging the statistical significance of the t^{th} multiplicative term. This leads to sequential statistical tests that may be used to determine how many multiplicative terms the model actually should contain. Critical values of Λ_t were obtained from the aforementioned simulation study for two-way tables with dimensions up to 99×19. Moreover, a method of moments approximation of the distribution of $[(p-t+1)\,\Lambda_t - 1]/(p-t)$ by a Beta distribution was described, and values of Beta parameters, a_t and b_t, were tabulated. Given the Beta parameters, approximate probabilities for the sequential tests may be computed using any software that will compute the incomplete Beta function. We refer to this test as the Seyedsadr-Cornelius (SC) test.

Goodman and Haberman (1990) proved that the residual sum of squares in an AMMI model with t multiplicative terms is asymptotically distributed as $\sigma^2 \chi_\nu^2$ where $\nu = (r-t-1)(c-t-1)$ as the λ_k become large {more precisely, as $\min[(\lambda_k - \lambda_{k+1}),\ k = 1, \ldots, t]/\sigma \to \infty$, such that λ_1/λ_t remains finite}. This implies that Mandel's (1971) allocation of df to the S_k for AMMI models are an overallocation for multiplicative terms that are clearly non-null. This may have only a small effect on the distribution of the ratio Λ_t. However, Marasinghe (1985) and Schott (1986) proved that the asymptotic distribution of Λ_t, under the hypothesis H_0: $\lambda_t = 0$, as $\lambda_k/\sigma \to \infty$ *for* $k = 1, \ldots, t-1$ is the same as that of Λ_1 in a smaller problem, namely, with $r-t+1$ rows and $c-t+1$ columns, for which Johnson and Graybill (1972) have described a method for determining statistical significance. This theoretical result from Marasinghe and Schott also follows from Goodman and Haberman's asymptotic expression for the residuals.

It seems clear from the Goodman and Haberman result that the asymptotic distribution of the residual sum of squares in $SHMM_t$ as the λ_k for $k=1, \ldots, t$ become large will be $\sigma^2 \chi^2_\nu$ where $\nu=(r-t)(c-t)-1$ provided that we introduce one additional constraint, namely, that $\sum_{k=1}^{t} \lambda_k \bar{\alpha}_k \bar{\gamma}_k$ remains finite. The further constraint is necessary to prevent β from approaching $\pm\infty$, a consequence of which would be $SHMM_t$ approaching $AMMI_{t-1}$ (Seyedsadr and Cornelius 1991 c). The implication for the *df* to be assigned to the S_k is similar to the case of AMMI, but an analogous result for the asymptotic distribution of Λ_t under $H_0: \lambda_t = 0$ obtained for AMMI by Marasinghe (1985) and Schott (1986) is not so easily proved for SHMM, if indeed it holds at all. (The difficulty here is that whereas the t-term model residuals are a function only of the $(t-1)$-term residuals for AMMI models, this is not true for SHMM). Nevertheless, until a more exact result is available, we believe that a good approximate asymptotic test can be done by comparing Λ_t to the distribution of Λ_1 in a problem with $r-t+1$ rows and $c-t+1$ columns using the SC Beta approximation. We will refer to this test as the Seyedsadr-Cornelius/Schott-Marasinghe (SC/SM) test. Except for the first multiplicative term where the two tests are identical, the SC/SM tests are generally more conservative than the SC tests. The Goodman-Haberman, Schott and Marasinghe theorems suggest that the SC tests for terms other that the first are likely to be somewhat liberal. Simulation results of Marasinghe and Schott indicate that the asymptotic tests for multiplicative terms in AMMI will be somewhat conservative except when the true λ_k values for the earlier terms are very large. It is reasonable to suppose that the same will be true for SHMM. Thus, we may surmise that the correct P value for Λ_t is something intermediate between the values obtained by SC and SC/SM tests.

If an estimate of within-cell error is available, then test criteria which use this estimate may be devised. Suppose y_{ij} in model (1) is a cell mean of n replicates. Then the variance of e_{ij} is $\sigma^2=\sigma_p^2/n$ where σ_p^2 is the "plot error", i.e., within-cell error on an individual plot basis. Let $\mu_1=E(S_t/\sigma^2)$ and $\mu'_2=V(S_t/\sigma^2)$. Further, let s^2 and f denote the estimate of σ_p^2 and its *df*, respectively. The following are two F approximations which may be used as approximate tests of the significance of the t^{th} multiplicative term.

1) Let $q_1=\mu'_2+\mu_1^2+(f-4)\mu_1$, $q_2=(f-2)\mu'_2+2\mu_1^2$, $b=q_1\mu_1/q_2$, $a=1+(f-2)q_1/q_2$. Then

$$F_1=\frac{nS_t/b}{fs^2/a}$$

is approximately distributed as F with $2b$ and $2a$ *df*.

2) $F_2=nS_t/\mu_1 s^2$ is approximately distributed as F with f_1 and f *df*, where $f_1=2\mu_1^2/\mu'_2$.

Both of these F approximations are obtained indirectly. The first uses the fact that if aX/b is distributed as F with $2b$ and $2a$ *df*, then $X+1$ has the same distribution as $1/B$ where B is a Beta random variable with parameters a and b. The values for a and b follow from putting $X=nS_t/fs^2$ and putting the first and second moments of $X+1$ equal to the corresponding moments of $1/B$. This method is previously unpublished, but is one which the first author has used for many years for testing statistical significance of multiplicative terms in AMMI models. The second F approximation employs essentially the same device as was used by Johnson (1976) to derive a critical value for a simultaneous test procedure for interaction contrasts. Specifically, we find u and f_1 such that uS_t/σ^2 is approximately (by the method of moments) distributed as chi-square with f_1 *df*. Our software computes only F_1 since it was in use prior to Johnson (1976), and subsequent investigation (Cornelius, unpublished) has suggested that the F_1 and F_2 tests will control Type I error rates for AMMI models about equally well. Moreover, the F_1 and F_2 test are asymptotically equivalent as $f\to\infty$. Boik (1985) studied the accuracy of the F_2 approximation for computing probabilities for testing the significance of the first multiplicative term in AMMI models and concluded that it was accurate enough for most applications. It is reasonable to expect this also to be true for SHMM. In applying the tests, μ_1 is put equal to the expectation and μ'_2 to the square of the standard deviation of S_t obtained in the simulation study of Seyedsadr and Cornelius (1991 c) (or Mandel 1971, if one is testing AMMI components). Here again, one may invoke our previous reasoning based on theorems of Goodman and Haberman (1990). If so, then μ_1 and μ'_2 are obtained using the expectation and standard deviation of S_1 in a problem of reduced dimension, namely, $r-t+1$ rows and $c-t+1$ columns. We will denote F_1 and F_2 with μ_1 and μ'_2 so obtained as F_{GH1} and F_{GH2}.

The asymptotic chi-squaredness of the SHMM residual sum of squares among cells allows us to (conservatively) test this residual against s^2 by a conventional F test with $f_2=(r-t)(c-t)-1$ and f *df*. The F statistic for this is

$$F_R=n\left[\sum\sum(y_{ij}-\bar{y}_{..})^2-\sum_{k=1}^{t}S_k\right]\Big/ f_2 s^2 .$$

Significance of the F_R test implies that the t-term model is an inadequate model, but this test does not have high power for detecting the need for another multiplicative term.

A simple example

A wheat (*Triticum aestivum* L.) yield data set with four cultivars in 13 locations previously analyzed by Snee (1982) (and several other authors cited by Snee) provides a simple illustration of the use of SHMM to identify subsets of environments in which genotypic effects are separable. The data, expressed as bushels per acre by previous authors, have been converted to g m^{-2} (Table 1) rounded to the nearest integer. As was done by Snee, we have arranged the cultivars and locations in rank order, each with respect to their means over levels of the other factor. The SHMM anova (Table 2) indicates that two multiplicative terms are required in order for SHMM to adequately describe these data.

In plotting the fitted $SHMM_2$ model (Fig. 2), we show the shift parameter plus primary effects as a set of concurrent regression lines and the $SHMM_2$-estimated yields as a scatter of points around these regression lines. The raw means are not shown. The entire data set lacks genotypic separability, as the point of concurrence occurs within the scatter of data and broken line graphs connecting the $SHMM_2$ estimated yields for each cultivar will cross over at several points. This will almost always be true when secondary effects are significant. However, it is possible to have significant secondary effects and yet have genotypic separability. Such a situation would imply that the appropriate operator is not the multiplicative operator.

Table 1. Yields ($g\,m^{-2}$) of four wheat cultivars tested in 13 locations and $SHMM_2$ parameter estimates

Location	Cultivar					Estimated effects ($SHMM_2$)	
	1	4	2	3	Mean	Primary ($\hat{\alpha}_{i1}$)	Secondary ($\hat{\alpha}_{i2}$)
7	371	269	261	142	261	0.493	−0.065
13	277	296	219	197	247	0.393	−0.563
8	372	169	230	125	224	0.415	0.381
10	311	197	212	151	218	0.359	0.042
5	304	202	197	136	210	0.342	0.040
1	293	131	162	131	179	0.268	0.307
2	272	160	146	112	172	0.248	0.182
6	174	182	172	157	171	0.178	−0.270
9	133	151	146	157	147	0.100	−0.254
4	132	123	125	120	125	0.062	−0.051
11	100	152	105	133	122	0.035	−0.287
3	122	108	95	112	109	0.024	−0.004
12	51	149	32	138	92	−0.061	−0.424
Mean	224	176	162	139	175	0.220	−0.074
Estimated effects ($SHMM_2$)							
Primary ($\hat{\gamma}_{j1}$)	0.794		0.411	0.415		0.169	$\hat{\beta} = 102.0$
Secondary ($\hat{\gamma}_{j2}$)	0.471		−0.743	0.027		−0.474	$\hat{\lambda}_1 = 720.5$
							$\hat{\lambda}_2 = 183.7$

Table 2. SHMM anova of the wheat data

Source of variation	*df* [a]	Sum of squares	Mean square	*P* value [b]	
				SC test	SC/SM
Primary effects	23.74	244 895	10 313	<0.0001	<0.0001
Secondary effects	14.61	29 124	1 994	0.0007	0.0036
Tertiary effects	8.57	7 103	829	0.0592	0.0976
Remainder	4.08	1 476	362	–	–
$SHMM_2$ residuals	12.65	8 579	678	–	–

[a] *df* are computed by formulas developed by Seyedsadr and Cornelius (1991 a) to approximate the expectations of the *SS* when the data are standard normal deviates

[b] For description of the SC and SC/SM tests, see text

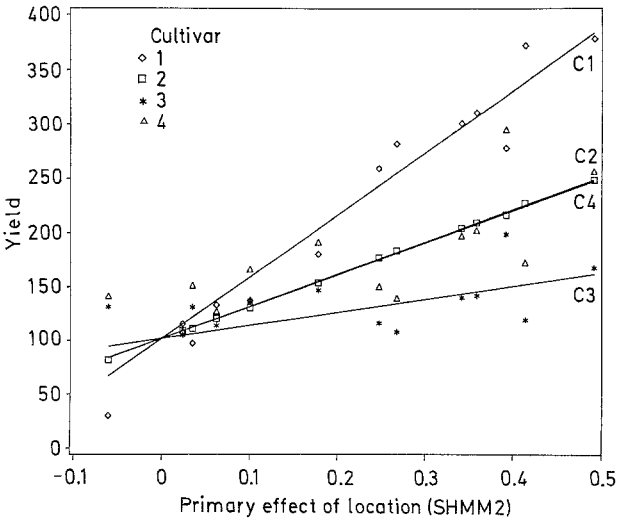

Fig. 2. $SHMM_2$ fitted to the wheat yield data. Plotted points are the $SHMM_2$ estimated yields. *Straight lines* show shift parameter plus $SHMM_2$ estimates of primary effects

Search for subsets with genotypic separability

For this example, we do not have an estimate of within-cell error available, so the SC and SC/SM tests are the only significance tests among those we have described which are available to use in our search for subsets of locations that have genotypic separability. In addition to these tests, however, we recommend also observing the magnitude of anova mean squares owing to secondary, tertiary, ..., etc., effects, as well as the $SHMM_1$ residual mean squares (i.e., secondary, tertiary, etc., pooled) obtained in analyses of subsets as compared to the residual mean square of $s_R^2 = 678$ (Table 2) obtained when the chosen parsimonious model ($SHMM_2$ in our example) was fitted to the entire data set. What we would like to see is that such mean squares are as small as 678, or at least not much greater; else it implies that the subsetting is giving us estimated yields, the fit to the data of which is actually worse than the overall model, $SHMM_2$ in our case, which was deemed necessary to describe the entire data set. To put it another way, the variation captured by secondary effects in the overall model should be successfully recovered as primary effects in subsets and the differing patterns in those primary effects from one subset to another. We would like to be more precise concerning the comparison of mean squares from SHMM anovas of subsets to the residual mean square from the SHMM anova of the entire data set, but critical values for such comparisons are an unsolved mathematical problem.

The $SHMM_2$ secondary effects ($\hat{\alpha}_{i2}$) of locations are included in Table 1. It is reasonable to expect that those which are near zero would form a subset for which $SHMM_1$ would be adequate. If $\hat{\beta}$ and $\hat{\gamma}_k$ were known parameters rather than estimates, then the variance of $\hat{\lambda}\hat{\alpha}_{ik}$ would be σ^2. It is clear from Goodman and Haberman (1990) that the correct variance of $\hat{\lambda}\hat{\alpha}_{ik}$ would be a much more complicated expression. For simplicity, let us regard $\hat{\lambda}\hat{\alpha}_{ik}$ as being "near zero" if it differs from zero by less than one standard deviation, i.e., if $|\hat{\alpha}_{ik}| < \sigma/\hat{\lambda}_k$. Substituting the $SHMM_2$ residual mean square as an estimate of σ^2, the i^{th} environment has "near zero" secondary effect if $|\hat{\alpha}_{i2}| < \sqrt{678}/183.7 = 0.142$. (If the pooled within cell error estimate s^2 was available, we would use $s/\sqrt{n}$ to estimate σ.) Thus, as a first attempt, we will try the following groups.

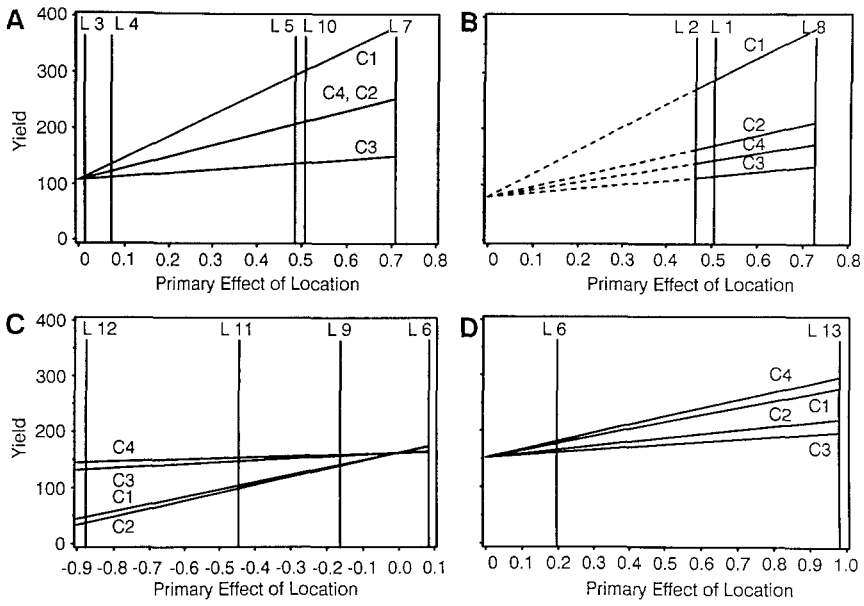

Fig. 3A–D. Plots of $SHMM_1$ independently fitted to data from subsets of locations: **A** Group 1 (locations 3, 4, 5, 7, 10); **B** Group 2 (locations 1, 2, 8); **C** Group 3a (locations 6, 9, 11, 12); **D** Group 3c (locations 6 and 13)

Table 3. SHMM anovas of location groups 1, 2 and 3

Source of variation	*df* [a]	Sum of squares	Mean square	*P* value [b]	
				SC test	SC/SM
	Group 1 (Locations 3, 4, 5, 7, 10)				
Primary effects	11.75	120 416	10 249	<0.0001	<0.0001
Secondary effects	5.24	1 173	224	0.4401	0.4817
Tertiary effects	1.82	386	212	0.1455	0.1517
Remainder	0.193	2	9	–	–
$SHMM_1$ residuals	7.25	1 561	215	–	–
	Group 2 (Locations 1, 2, 8)				
Primary effects	8.15	71 624	8 786	0.0013	0.0013
Secondary effects	2.60	1 061	408	0.9849	0.9878
Remainder	0.245	659	2 687	–	–
$SHMM_1$ residuals	2.85	1 721	604	–	–
	Group 3 (Locations 6, 9, 11, 12, 13)				
Primary effects	11.75	64 889	5 523	0.0037	0.0037
Secondary effects	5.24	9 576	1 829	0.0033	0.0055
Tertiary effects	1.82	384	211	0.3350	0.3444
Remainder	0.193	9	49	–	–
$SHMM_2$ residuals	2.02	343	195	–	–
$SHMM_1$ residuals	7.25	9 970	1 375	–	–

[a] *df* are computed by formulas developed by Seyedsadr and Cornelius (1991a) to approximate the expectations of the *SS* when the data are standard normal deviates

[b] For description of the SC and SC/SM tests, see text

Group 1: Locations 3, 4, 5, 7, 10 (secondary effects near zero);
Group 2: Locations 1, 2, 8 (secondary effects positive);
Group 3: Locations 6, 9, 11, 12, 13 (secondary effects negative).

It is clear from the anova of Group 1 (Table 3) that $SHMM_1$ gives an adequate fit for this group. In Group 2, no evidence for significant secondary effects was found, and the mean square for this was 408. The large mean square for remainder after secondary effects of 2687 (which is due only to tertiary effects since $SHMM_3$ will fit the data on three locations exactly) is somewhat disconcerting, but the sum of squares is only 659 and, when pooled with the secondary effects, gives a $SHMM_1$ residual mean square of 604, less than our target value of $s_R^2 = 678$. Apparently, we can also regard $SHMM_1$ as satisfactory for this group.

$SHMM_1$ fitted to Group 1 gives a graph (Fig. 3A) very similar to the configuration of primary effects in the portion of Fig. 1 that contains these locations (to the right of the point of concurrence), which should be expected since this is the group that had secondary effects close to zero in the overall $SHMM_2$. The fitted $SHMM_1$ in Group 2 (Fig. 3B) has the same rank order of cultivars as in Group 1. The two groups differ primarily in the positioning of the regression lines for cultivars 2 and 4 relative to one another and to the high-yielding cultivar 1 and the low-yielding cultivar 3. Cultivars 2 and 4 are essentially alike in location Group 1, but cultivar 2 seems to be distinctly the better of the two in Group 2. The equivalence of rank order of cultivars in these two location groups suggests that they collectively form one group in which there is genotypic separability, but a mathematical representation of it as such requires some operator other than the multiplicative operator.

In Group 3, the SC and SC/SM tests are significant for scondary effects, and the mean square for these effects is 1829, considerably greater than $s_R^2 = 678$. Its $SHMM_1$ residual mean square is also quite unacceptable. Clearly,

Table 4. SHMM anovas of location subgroups from group 3

Source of variation	*df* [a]	Sum of squares	Mean square	*P* value [b]	
				SC test	SC/SM
	Group 3A (Locations 6, 9, 11, 12)				
Primary effects	10.03	25 772	2 570	0.0001	0.0001
Secondary effects	3.96	556	140	0.9418	0.9528
Remainder	1.01	359	356	–	–
$SHMM_1$ residuals	4.97	915	184	–	–
	Group 3B (Locations 9, 11, 12)				
Primary effects	8.15	18 294	2 244	0.0007	0.0007
Secoandary effects	2.60	246	94	0.9617	0.9676
Remainder	0.245	120	488	–	–
$SHMM_1$ residuals	2.85	365	128	–	–
	Group 3C (Locations 6, 13)				
Primary effects	6.00	18 396	3 068	0.0202	0.0202
Remainder	1.00	67	67	–	–

[a] *df* are computed by formulas developed by Seyedsadr and Cornelius (1991a) to approximate the expectations of the *SS* when the data are standard normal deviates

[b] For description of the SC and SC/SM tests, see text

Group 3 needs further subdivision. We will consider two approaches:

1) Trim the group of locations based on the original secondary effect estimates obtained when $SHMM_2$ was fitted to the entire data set;
2) choose a subdivision based on estimated effects when $SHMM_2$ is fitted to Group 3 alone.

Pursuing the first approach, let us delete the location with the most extreme secondary effect, i.e., location 13 ($\hat{\alpha}_{i2} = -0.563$), giving Group 3a: locations 6, 9, 11, 12. The anova of this subgroup (Table 4) indicates that $SHMM_1$ gives a good fit to the data, but the fitted $SHMM_1$ (Fig. 3C) shows location 6 to be to the right and locations 9, 11 and 12 to the left of the point of concurrence. Thus, the fitted $SHMM_1$ displays crossover interaction. Note, however, that all cultivar differences in location 6 are small and possibly insignificant, giving this location a potential capability of being "played as a wild card" in selecting location subsets.

Pursuing the second approach, the $SHMM_2$ primary and secondary effects in Group 3 are:

Location	$\hat{\alpha}_{i1}$	$\hat{\alpha}_{i2}$
13	0.752	−0.636
6	0.135	−0.096
9	−0.088	−0.041
11	−0.282	−0.286
12	−0.573	−0.709

The secondary effects are all negative, but the primary effects include both positive and negative values. Let us split this group into:

Group 3b: locations 9, 11, 12 (primary effects negative);
Group 3c: locations 6, 13 (primary effects positive).

Group 3b is the same as would be obtained by removing location 6 from Group 3a previously analyzed. Anovas of Groups 3b and c (Table 4) clearly indicate that $SHMM_1$ is a satisfactory model for each of these subgroups. Graphical display of the fitted $SHMM_1$ for Group 3b (not shown) gives a configuration very similar to the portion of Fig. 3C (Group 3a) to the left of the point of concurrence. For Group 3c the fitted $SHMM_1$ is shown as Fig. 3D. Locations 6 and 13 actually have the same observed ranking of cultivars in the raw data; thus, it is reasonable that they should group together.

A more complicated example

Data from an international maize (*Zea mays* L.) cultivar trial (EVT 16B) (Table 5) with 20 experimental sites and nine cultivars provide an example with a more complicated pattern of interaction. To facilitate comparison with $SHMM_1$-estimated yields in subsets given later (Table 7), we have arranged the data in Table 5 by subsets of sites that were eventually obtained. Data are means of four replications, and the pooled error mean square is 602 847 (150 712 on a cell-mean basis) with 480 *df*. Analysis of these data revealed that at least three multiplicative terms would be necessary to adequately model the response pattern in the entire data set (F_{GH1} and F_R tests indicated a need for three multiplicative terms; F_1 showed significance for six terms). We chose to use $SHMM_3$ as a starting point for subsetting the data. The $SHMM_3$ residual mean square among cells was 224 366. The steps by which we obtained subsets of sites with genotypic separability are summarized in Table 6. The first subdivision was made on the basis of $SHMM_3$ tertiary effects, the intention being to obtain subsets in which $SHMM_2$ would give an adequate fit. This was not entirely successful, but subsequent subdivisions did isolate groups with genotypic separability.

By the *SC* and F_1 tests, Group 1 still required $SHMM_3$ for an adequate model. The fitted $SHMM_3$ had secondary effects of sites that were all of the same sign, but some primary effects were positive and others negative. It seems expedient in such a circumstance to split the group into subsets based on sign of the primary effects. Doing so led immediately to one subset of three sites (Group 1a, see Table 6), which was satisfactory. $SHMM_2$ was adequate for the remaining sites in Group 1 and dichotomous subdivision on the basis of secondary effects successfully split them into two satisfactory subgroups.

Table 5. Yields (kg ha^{-1}) of nine maize cultivars at 20 sites

Site	Cultivar									Mean
	1	2	3	4	5	6	7	8	9	
7	2382	2515	3529	2998	3556	3949	3537	3088	3061	3229
10	3100	2972	2785	2843	2688	3024	2889	3353	2274	2936
19	1632	2282	3059	2233	3073	3011	3211	2634	2735	2652
6	6437	6036	6459	6678	6882	6916	6745	4986	5610	6305
14	5849	5932	5886	6439	6359	6380	5820	5522	6282	6052
4	4566	4963	5136	6030	5831	5980	4342	4442	5781	5230
5	4380	5201	4178	5672	5414	5591	4277	4476	5407	4955
9	4647	4714	5448	4864	5588	5603	4318	4001	5553	4971
13	6721	5627	6294	7332	7174	7262	5544	4117	6920	6332
16	5010	5196	5455	6351	6070	5730	5013	4551	5278	5406
1	3622	3426	3446	3720	3165	4116	3354	4529	3136	3613
3	5554	4937	5117	4542	6173	5205	5389	5248	3780	5105
15	4601	4126	4537	6331	6328	5961	4346	4321	4889	5049
17	4415	4211	4749	5161	5454	5807	3862	5243	4989	4877
20	4587	4396	5018	4988	5776	5088	4056	4806	4822	4838
8	6011	5278	4731	2516	2732	2983	4206	4484	3309	4028
2	3728	3919	4082	4539	4079	4878	4767	3393	4500	4209
18	3344	4515	4295	5618	4498	5333	5276	2940	5244	4552
11	4433	4349	4526	7117	5995	6150	5052	3713	6430	5307
12	6873	7571	7727	8385	8106	7637	7444	5816	8091	7516
Mean	4617	4603	4823	5218	5247	5330	4672	4283	4930	4858

Group 2 also still required $SHMM_3$ for an adequate model according to the F_1 test and $SHMM_2$ according to F_{GH1}. An attempt to subdivide it on the basis of $SHMM_3$ tertiary effects gave a three-site group (2b), which was definitely unsatisfactory (second effects significant by both F_1 and F_{GH1} tests), and a two-site group (2a), which might have been considered marginally satisfactory. [The $SHMM_1$ residual mean square was 265 025, which looks uncomfortably large compared to the $SHMM_3$ residual mean square from the entire data set and the pooled error mean square (on a cell-mean basis), but none of the previously described tests for the secondary effects against the pooled error were significant. The SC and SC/SM tests of secondary effects do not exist for a subgroup of only two sites.] However, a second attempt to subdivide Group 2, this time using $SHMM_3$ secondary effects as the criterion, led to a satisfactory group of three sites (2d). The other two sites (1 and 3) in Group 2 are unsatisfactory as a group of two and will be left as ungrouped sites.

In Group 3, $SHMM_2$ appeared to be satisfactory, but $SHMM_3$ apparently was needed for a subset of four sites (3b) obtained after one subdivision, which had split one site off from the other four. Two successive subdivisions, each of which split one site away from the remaining sites, finally gave the satisfactory two-site group 3f. At this point it seemed reasonable to investigate whether we could find any satisfactory grouping of the three sites (8, 11 and 12) from Group 3, which had been separated out one at a time in previous steps. Doing so led to Group 3g (sites 11 and 12) as a marginally satisfactory group of two. The $SHMM_1$ residual mean square in this subset was 267 321, really no better than Group 2a, which was discarded in favor of an alternative subgrouping of Group 2. However, it appears doubtful if any better subgrouping of Group 3 can be found than Groups 3f, 3g, and site 8 as an ungrouped site, except possibly to let sites 11 and 12 also be ungrouped.

Thus, the final groupings consisted of one group of five sites (1d), two groups of three sites each (1a and 2d), three groups of two sites each (1c, 3f, 3g) and three sites (1, 3, 8) which were not found to satisfactorily group with other sites. $SHMM_1$ predicted yields in the location subsets are shown in Table 7.

Validation of the EVT 16B example

To validate the procedure, we investigated the extent to which genotypic crossover (i.e., rank change) interactions still exist in the raw cell means within the groups of sites finally chosen. The $SHMM_1$-estimated yields within these groups, of course, predict no crossover interactions whatever, and any which existed in the original data have been "smoothed-out", i.e., regarded as noise, in fitting

Table 6. Summary of steps by which subsets of environments with genotypic separability were identified in the EVT16B data set

Subgroups	Sites	Criterion	Is $SHMM_1$ adequate?
		First subdivision ($SHMM_3$) [a]	
1	{4, 5, 6, 7, 9, 10, 13, 14, 16, 19}	Tertiary effects near zero	No
2	{1, 3, 15, 17, 20}	Tertiary effects negative	No
3	{2, 8, 11, 12, 18}	Tertiary effects positive	No
		Subdividing group 1 ($SHMM_3$)	
1a	{7, 10, 19}	Primary effects negative	Yes
1b	{4, 5, 6, 9, 13, 14, 16}	Primary effects positive	No
		Subdividing group 1b ($SHMM_2$)	
1c	{6, 14}	Secondary effects negative	Yes
1d	{4, 5, 9, 13, 16}	Secondary effects positive	Yes
		Subdividing group 2 ($SHMM_3$)	
2a	{1, 17}	Tertiary effects positive	No (?)
2b	{3, 15, 20}	Tertiary effects negative	No
		Subdividing group 2 (second attempt, $SHMM_3$)	
2c	{1, 3}	Secondary effects positive	No
2d	{15, 17, 20}	Secondary effects negative	Yes
		Subdividing group 3 ($SHMM_2$)	
3a	{8}	Secondary effects large, positive	–
3b	{2, 11, 12, 18}	Secondary effects small, negative	No
		Subdividing group 3b ($SHMM_3$)	
3c	{12}	Primary effects negative	–
3d	{2, 11, 18}	Primary effects positive	No
		Subdividing group 3d ($SHMM_2$)	
3e	{11}	Secondary effects negative	–
3f	{2, 18}	Secondary effects positive	Yes
		Subdividing {8, 11, 12} (regrouped, $SHMM_2$)	
3a	{8}	Secondary effects large	–
3g	{11, 12}	Secondary effects small	Yes (?)

[a] Model shown in parenthesis is the model needed to adequately model the group to be subdivided

$SHMM_1$ to the subsets. A total of $r(r-1)\,c(c-1)/4$ interaction contrasts of the form $\bar{y}_{ij.}-\bar{y}_{i'j.}-\bar{y}_{ij'.}+\bar{y}_{i'j'.}$, i.e., 2×2 contrasts, for $i<i'$ and $j<j'$ exist in a $r\times c$ table. Any of these is statistically significant at $P<\alpha$ by an ordinary t-test if its absolute value exceeds $t_{\alpha(f)}\sqrt{4s^2/n}$ where f, s^2 and n are as previously defined. For the EVT16B data this critical value is 1526 for $\alpha=0.05$. Out of the 6840 2×2 interaction contrasts in the EVT16B data, 2728 (39.88%) of them are crossover interactions (i.e., actually show genotypic rank change in their observed pattern) and 1042 (15.23% of all 2×2 contrasts; 38.20% of the crossover interactions) exceed the critical value 1526.

In the subsets finally chosen, there are 684 2×2 contrasts, 158 (23.10%) of them being crossover interactions, of which 18 (2.63% of all 2×2 contrasts in subsets; 11.39% of the crossover interactions in subsets) exceed the critical value 1526. Clearly, the percentage of interactions in subsets which are crossover interactions is significantly less than for the overall data set, and the percentage of those interactions which are statistically significant is dramatically less. Moreover, if a random 20×9 table with no true genotypic differences apart from that which derives from random error was randomly subpartitioned into subsets of rows of the sizes obtained here, we would expect 5% of the 2×2 interactions to be "significant" by a 0.05-level t-test. We would further expect roughly half of those interactions to be crossover and half to be non-crossover interactions. Thus, 2.63% of the interactions appearing as "significant" crossover interactions is close to the expected 2.5%.

Still more dramatic is a comparison of the magnitude of the crossover interactions. For those with sites in different subsets ("between subsets") they range from −5372 to 6180 with a mean absolute value of 1503. Thus, their mean absolute value was almost as large as the critical value 1526. The within-subset crossover interactions, i.e., those which "slipped through the cracks" in our procedure, ranged from −2700 to 2092 with a mean absolute value of only 878. For those judged significant, the mean absolute value was 2463 between subsets and 1862 within subsets.

Two other interesting observations are that 8 of the 18 "significant" crossover interactions in subsets involve site 5 in Group 1d, and 45 of the 158 crossover interactions in

Table 7. Yields estimated from $SHMM_1$ fitted to subsets of locations

Site	Cultivar								
	1	2	3	4	5	6	7	8	9
	Site group 1a								
7	2738	2882	3179	2903	3179	3219	3229	3058	3045
10	2871	2976	3192	2992	3192	3222	3229	3104	3095
19	1717	2162	3077	2228	3076	3201	3230	2703	2664
Mean	2442	2673	3149	2708	3149	3214	3229	2955	2935
Raw mean	2521	2589	3125	2691	3105	3328	3212	3025	2857
Rank [a]	9	8	3	7	4	2	1	5	6
	Site group 1c								
6	6364	6027	6395	6756	6917	6954	6633	4957	5745
14	6071	5961	6082	6200	6253	6265	6160	5610	5869
Mean	6218	5994	6238	6478	6485	6609	6396	5284	5807
Raw mean	6143	5984	6173	6559	6620	6648	6282	5254	5946
Rank [a]	6	7	5	3	2	1	4	9	8
	Site group 1d								
4	5072	4985	5199	5878	5818	5832	4669	4210	5605
5	4812	4742	4914	5462	5414	5425	4487	4117	5242
9	4832	4761	4936	5494	5445	5456	4501	4125	5269
13	6031	5881	6248	7413	7310	7334	5339	4552	6944
16	5173	5079	5309	6039	5974	5989	4739	4246	5745
Mean	5184	5099	5321	6057	5992	6007	4747	4250	5814
Raw mean	5064	5140	5302	6050	6015	6033	4699	4317	5788
Rank [a]	6	7	5	1	3	2	8	9	4
	Ungrouped sites from group 2								
1	3622	3426	3446	3720	3165	4116	3354	4529	3136
Rank	4	6	5	3	8	2	7	1	9
3	5554	4937	5117	4542	6173	5205	5389	5248	3780
Rank	2	7	6	8	1	5	3	4	9
	Site group 2d								
15	4519	4093	4738	5986	6334	6037	3991	4731	4987
17	4558	4291	4694	5474	5691	5506	4228	4690	4850
20	4571	4361	4679	5292	5463	5317	4312	4675	4801
Mean	4549	4248	4704	5584	5829	5620	4177	4699	4879
Raw mean	4534	4244	4768	5493	5853	5619	4088	4790	4900
Rank [a]	7	8	5	3	1	2	9	6	4
	Ungrouped site from group 3								
8	6011	5278	4731	2516	2732	2983	4206	4484	3309
Rank	1	2	3	9	8	7	5	4	6
	Site group 3f								
2	3617	4093	4078	4713	4160	4668	4621	3382	4552
18	3402	4325	4297	5528	4456	5441	5351	2946	5217
Mean	3509	4209	4187	5120	4308	5054	4986	3164	4885
Raw mean	3536	4167	4188	5078	4288	5105	5022	3167	4872
Rank [a]	8	6	7	1	5	2	3	9	4

Table 7 (continued)

Site	Cultivar								
	1	2	3	4	5	6	7	8	9
	Site group 3 g								
11	4387	4637	4836	7006	6069	5971	5089	3393	6377
12	6948	7103	7226	8565	7986	7926	7382	6335	8177
Mean	5667	5870	6031	7786	7026	6949	6236	4864	7277
Raw mean	5653	5960	6127	7751	7050	6893	6248	4764	7260
Rank [a]	8	7	6	1	3	4	5	9	2

[a] Rank order (high to low) of $SHMM_1$ estimated yields

subsets involve Site 10 in Group 1 a. Only 3 of these crossover interactions involving Site 10 exceed the critical value 1526. Site 10, with many of its cultivar differences small, seems to be behaving somewhat like a "wild card" in Group 1 a.

Baker (1988, 1990) used the method of Azzalini and Cox (1984) to evaluate 2×2 subtables for the presence of crossover interactions. By the Azzalini-Cox test, significant crossover interaction is indicated if the absolute value of the difference between the two cultivars exceeds a critical value in each of the two environments and the cultivar differences in the two environments are of opposite sign. The critical values may differ if different estimates of the standard errors are used. Using the pooled error as the error estimate for all environments, the 0.05 level critical value for the Azzalini-Cox test is 1586. There were 43 significant crossover interactions by this test, none of which involves sites from the same subset.

The Azzalini-Cox test is clearly intended to have an experiment-wise error rate of α, and its power to detect real crossover interactions will decrease as the dimensions ($r \times c$) of the two-way table of environments and cultivars increases. If we, instead, substitute a 0.05-level t-test (comparison-wise error rate of $\alpha = 0.05$), the critical value for a cultivar difference within a site is 1079. Using this critical value, 187 significant crossover interactions were found. Again, none of them were within a subset.

We can also define an "interaction-wise" Type I error rate, i.e., error rate per 2×2 table tested for crossover interaction. If the critical value for a simple cultivar difference is $C = t_\theta \sqrt{2s^2/n}$ where t_θ is a value of Student's t such that $P[t > t_\theta] = \theta$, then the probability of declaring a crossover interaction to exist when there is, in fact, no cultivar difference in either environment (such is the null hypothesis in the Azzalini-Cox test), is $2\theta^2$ [expression (3) of Azzalini and Cox]. Putting $2\theta^2 = \alpha$ gives $\theta = (\alpha/2)^{1/2}$. For $\alpha = 0.05$, we have $\theta = 0.1581$, $t_{0.1581} = 1.0033$ for 480 df, and $C = 551$. Using this critical value, 646 significant crossover interactions were found. Only 9 of them were within subsets, all of them in subset 1 d (7 of these involve site 9 and, of these 7, 4 involved cultivars 3 and 4). The mean absolute value of the 2×2 interaction contrasts was 1682 for these 9 within-site subtables and was 2600 for the 637 between-site subtables showing significance. The number of possible 2×2 subtables within subset 1 d is 360. Thus, 9 significant crossover interactions is only half the number expected in this subset under the null hypothesis ($0.05 \times 360 = 18$), and roughly one-fourth of the number expected within all subsets ($0.05 \times 684 = 34.2$). We previously mentioned that the selected subsets contained 18 interactions with crossover pattern, which were significant by 0.05-level t-test of 2×2 interaction contrasts. Of these, 12 failed to show simultaneous significance of both within-site cultivar differences even by this liberal Azzalini-Cox-type test with a 0.05 interaction-wise error rate.

Concluding comments

This analysis indicated that the shifted multiplicative model, with the aid of appropriate statistical procedures, can be quite effectively used to identify subsets of environments in which genotypic effects are separable. The method described is "exploratory" in that the investigator explores the data as he makes decisions at each step. At some steps, if a decision made does not seem to lead to a useful result, an alternative decision may be explored. It is a "step-down" method in that it begins with the entire data set and seeks to subdivide it into subsets. These properties are as opposed to clustering methods where the computer does everything in a "step-up" direction, building first small clusters and proceeding to amalgamate the small clusters into larger clusters truncating the process when the next amalgamation fails to produce a cluster that can be adequately modeled by $SHMM_1$.

Obviously, the subsets identified by the exploratory method described here are not unique since the result finally obtained depends on decisions made at earlier stages, and particularly on what subdivision was chosen at the initial stage. In practice, alternative decisions might be tried at the initial or later stages to observe what end

results are ultimately obtained. In the process, the investigator will learn much about the structure of his data.

In further work, we will seek to develop clustering algorithms that can be implemented on a computer without the need for decisions by the investigator at intermediate steps.

References

Azzalini A, Cox DR (1984) Two new tests associated with analysis of variance. J R Statist Soc B 46:335–343

Baker RJ (1988) Tests for crossover genotype-environmental interactions. Can J Plant Sci 68:405–410

Baker RJ (1990) Crossover genotype-environmental interaction in spring wheat. In: Kang MS (ed) Genotype-by-environment interaction and plant breeding. Department of Agronomy, Louisiana Agric Exp Stn, LSU Agric Center, Baton Rouge, La.

Boik RJ (1985) A new approximation to the distribution function of the studentized maximum root. Commun in Stat B Simul Comput 14:759–767

Gauch HG Jr (1988) Model selection and validation for yield trials with interaction. Biometrics 44:705–715

Gauch HG Jr, Zobel RW (1988) Predictive and postdictive success of statistical analyses of yield trials. Theor Appl Genet 76:1–40

Gregorius H-R, Namkoong G (1986) Joint analysis of genotypic and environmental effects. Theor Appl Genet 12:413–422

Goodman LA, Haberman ST (1990) The analysis of nonadditivity in two-way analysis of variance. J Am Stat Assoc 85:139–145

Johnson DE (1976) Some new multiple comparison procedures for the two-way AOV model with interaction. Biometrics 32:929–934

Johnson DE, Graybill FA (1972) An analysis of a two-way model with interaction and no replication. J Am Stat Assoc 67:862–868

Mandel J (1961) Nonadditivity in two-way analysis of variance. J Am Stat Assoc 56:878–888

Mandel J (1971) A new analysis of variance model for nonadditive data. Technometrics 13:1–8

Marasinghe MG (1985) Asymptotic tests and Monte Carlo studies associated with the multiplicative interaction model. Commun Stat A Theory Methods 14:2219–2231

Schott JR (1986) A note on the critical values used in stepwise tests for multiplicative components of interaction. Commun Stat A Theory Methods 15:1561–1570

Seyedsadr SM (1987) Statistical and computational procedures for estimation and hypothesis testing with respect to the shifted multiplicative model. Ph.D thesis, University of Kentucky

Seyedsadr M, Cornelius PL (1991 a) Functions approximating the expectations and standard deviations of sequential sums of squates in the shifted multiplicative model for a two-way table. University of Kentucky. Department of Statistics Technical Report 322

Seyedsadr M, Cornelius PL (1991 b) Hypothesis testing for components of the shifted multiplicative model for a nonadditive two-way table. University of Kentucky. Department of Statistics Technical Report 315

Seyedsadr M, Cornelius PL (1991 c) Shifted multiplicative models for nonadditive two-way tables. Commun Stat B Simul Comput 2/(3): (in press)

Snee RD (1982) Nonadditivity in a two-way classification: Is it interaction or nonhomogeneous variance? J Am Stat Assoc 77:515–518