CHOOSING COMPONENTS IN THE ADDITIVE MAIN EFFECT AND MULTIPLICATIVE INTERACTION (AMMI) MODELS

Carlos Tadeu dos Santos Dias[1]*; Wojtek Janusz Krzanowski[2]

[1]USP/ESALQ - Depto. Ciências Exatas, Av. Pádua Dias 11, C.P. 09 - 13418-900 - Piracicaba, SP - Brasil.
[2]School of Mathematical Sciences, Laver Building, North Park Road, Exeter, EX4 4QE, UK.
*Corresponding author <ctsdias@carpa.esalq.usp.br>

ABSTRACT: The additive main effect and multiplicative interaction (AMMI) models allows analysts to detect interactions between rows and columns in a two-way table. However, there are many methods proposed in the literature to determine the number of multiplicative components to include in the AMMI model. These methods typically give different results for any particular data set, so the user needs some guidance as to which methods to use. In this paper we compare four commonly used methods using simulated data based on real experiments, and provide some general recommendations.
Key words: two-way table, principal components, simulation

ESCOLHA DE COMPONENTES NOS MODELOS DE EFEITOS PRINCIPAIS ADITIVOS E INTERAÇÃO MULTIPLICATIVA (AMMI)

RESUMO: Os modelos de efeitos principais aditivos e interação significativa (AMMI) permitem ao analista detectar interações entre linhas e colunas em uma tabela de dupla entrada. Entretanto, existem muitos métodos propostos na literatura para determinar o número de componentes multiplicativos a incluir nos modelos AMMI. Esses métodos fornecem diferentes resultados para um particular conjunto de dados. Assim, o usuário necessita de orientação sobre quais métodos usar. Nesse trabalho investigamos quatro métodos comumente utilizados em um amplo espectro de condições usando dados simulados baseados em dados de ensaios reais. O método de Eastment-Krzanowski apresentou melhor desempenho; cada um dos outros métodos exibiram baixo desempenho em alguma das situações estudadas.
Palavras-chave: tabela de dupla entrada, componentes principais, simulação

INTRODUCTION

There are many tests in the literature for deciding how many components should be retained in additive and multiplicative interaction models. They include those by Gollob (1968), Mandel (1969; 1971), Corsten & van Eijnsbergen (1972), Johnson & Graybill (1972), Hegemann & Johnson (1976), Yochmowitz & Cornell (1978), and Cornelius (1993), among others. Data for analysis are assumed to come in the form of a two-way table. The classical model for such a table is the ANOVA model

$$y_{ij} = \mu + g_i + e_j + \delta_{ij}$$

$i=1,...,g, j=1,...,e,$ where y_{ij} is the observation in the i-th row and j-th column of the table, μ is the overall mean effect, g_i represents the i-th row effect, and e_j the j-th column effect. The terms δ_{ij} can either be seen as residuals or as interaction terms between rows and col-

umns. The above model is called the additive model. However, it is quite possible that the interaction terms δ_{ij} still contain some structure, and therefore one can model them by a set of multiplicative components plus residual error:

$$\delta_{ij} = \sum_{k=1}^{m} \alpha_{ik} \lambda_k \gamma_{jk} + \varepsilon_{ij}$$

The number of components $m \leq$ min $\{g-1, e-1\}$ should be chosen in such a way that the residual ε_{ij} represents white noise. Combining the last two expressions above yields an additive main effect and multiplicative interaction (AMMI) model for the two-way table:

$$y_{ij} = \mu + g_i + e_j + \sum_{k=1}^{m} \alpha_{ik} \lambda_k \gamma_{jk} + \varepsilon_{ij}$$

To estimate the unknown parameters in the model, one usually first uses row/column means for the main effects and then performs a singular

value decomposition of the residual matrix for the interaction parameters. This classical approach corresponds essentially to a least squares fit of the full model.

       Nonadditive effects are frequently observed in two-way tables, and, as Daniel (1976) points out, the nonadditivity is often associated with just a few rows or columns of the table. As observed by Snee (1982), the interpretation of nonadditivity is less of a problem if replicate observations are present for each of the cells of the table. However when there is only one observation per cell, it is not possible to distinguish between row- or column-related nonhomogeneous variance, and interaction from the observed data alone. In this situation, the use of a model that will detect a variety of different types of nonadditivity in two-way tables can be very helpful.

       For this reason, a primary objective is to determine the optimal number of multiplicative terms in the AMMI model. Also, good prediction of the true trait response in each cell of the two-way table can be achieved by truncating the AMMI model, so criteria for determining the number of components needed to explain the patterns of interactions have been the objects of some research.

       Dias & Krzanowski (2003) described two 'leave-one-out' methods to determine the number m of components to be retained in the AMMI model, and two other methods based on F-tests: the Gollob and Cornelius tests. However, these methods gave different results when applied to several data sets, so some guidance is needed regarding their properties if a choice between them is to be made in practice. In the present paper we conduct a comparison of the methods, to establish conditions under which each method might be expected to perform satisfactorily, by means of a simulation study built round real experimental data.

MATERIAL AND METHODS

Data and Model

       Model and examples are described in terms of a set of g genotypes that have been tested experimentally in e environments. The mean yield for each combination of genotype and environment, obtained from n replications of an experiment (a balanced set of data) can be represented by the array $Y_{(g \times e)} = [y_{ij}]$.

       The AMMI model postulates additive components for the main effects of genotypes (g_i) and environments (e_j), and multiplicative components for the effect of the interaction ($ge)_{ij}$, referred to as genotype $\times$ environment interaction (GEI). Thus, the mean response of genotype i in an environment j is modelled

by:

$$y_{ij} = \mu + g_i + e_j + \sum\nolimits_{k=1}^{m} \alpha_{ik} \lambda_k \gamma_{jk} + \varepsilon_{ij}$$

in which $(ge)_{ij}$ is represented by:

$$\sum\nolimits_{k=1}^{m} \alpha_{ik} \lambda_k \gamma_{jk}$$

under the restrictions:

$$\sum_i g_i = \sum_j e_j = \sum_i (ge)_{ij} = \sum_j (ge)_{ij} = 0.$$

       Estimates of the overall mean (μ) and the main effects (g_i and e_j) are obtained in the context of a simple two-way ANOVA of the array of means $Y_{(g \times e)}$. The residuals from this array then constitute the array of interactions $GE_{(g \times e)} = [(ge)_{ij}]$, and the multiplicative interaction terms are estimated from the singular value decomposition (SVD) of this array. Thus, λ_k is estimated by the k-th singular value of GE, α_{ik} is estimated by i-th element of the left singular vector $\alpha_{k(g \times 1)}$, and γ_{jk} is estimated by j-th element of the right singular vector $\gamma^{\Gamma}_{k(1 \times e)}$ associated with λ_k (Good, 1969; Mandel, 1971; Piepho, 1995). If the interaction contains one or more terms ($m \geq 1$), then it is assumed that the errors have a homogeneous variance after the nonadditivity effects have been taken into account.

       Snee (1982) points out that nonadditivity in a two-way table, with one observation per cell, may be due either to row- or column-related nonhomogeneous variance, or to interaction between the row and column factors. He discusses one example involving the effect of four varieties of wheat in 13 locations, and shows that the variety-related nonhomogeneous variance detected by several authors is more usefully interpreted as a level-dependent interaction due to the differences among the yields of different varieties of wheat being larger in those locations that have higher yields. Also, he comments that experience has indicated that frequently only one multiplicative term is needed to describe the nonadditivity.

Methods for selecting number of interaction components

       Dias & Krzanowski (2003) present two methods based on a full 'leave-one-out' procedure that optimizes the cross-validation process (i.e. maximises the number of data points left in the set at each iteration without incurring bias due to resubstitution), and two methods based on F-tests. The Eastment & Krzanowski (1982) and Gabriel (2002) cross-validation methods assume that we wish to predict the elements x_{ij} of the matrix X by means of the model:

$x_{ij} = \sum_{k=1}^{m} u_{ik} d_k v_{jk} + \varepsilon_{ij}$. The methods are those outlined by Krzanowski (1987) and Gabriel (2002) respectively, in which we predict the value $\hat{x}_{ij}^m$ of x_{ij} ($i=1,...,g; j=1,...,e$) for each possible choice of m (the number of components), and measure the discrepancy between actual and predicted values as

$$PRESS(m) = \frac{1}{ge} \sum_{i=1}^{g} \sum_{j=1}^{e} (\hat{x}_{ij}^m - x_{ij})^2$$

However, to avoid bias in the prediction, the data point x_{ij} must not be used in the calculation of $\hat{x}_{ij}^m$ for each i and j. Hence, appeal to some form of cross-validation is indicated, and the two approaches differ in the way that they handle this. However, both assume that the SVD of X can be written as $X=UDV^T$. The Eastment-Krzanowski method utilizes the following statistic to determine the number of components in the model:

$$W_m = \frac{PRESS(m-1) - PRESS(m)}{D_m} \div \frac{PRESS(m)}{D_r}$$

where D_m is the number of degrees of freedom required to fit the m-th component and D_r is the number of degrees of freedom remaining after fitting the m-th component. The standard cross-validation procedure is to subdivide X into a number of groups, delete each group in turn from the data, evaluate the parameters of the predictor from the remaining data, and predict the deleted values (Wold, 1976; 1978). The most precise prediction results when each deleted group is as small as possible, and in the present instance that means a single element of X (Krzanowski, 1987). Denoting by $X^{(-i)}$ the result of deleting the ith row of X and mean-centering the columns, and by $X_{(-j)}$ the result of deleting the jth column of X and mean centering the rows, following the scheme given by Eastment & Krzanowski (1982), one can write

$X^{(-i)} = \bar{U}\,\bar{D}\,\bar{V}^T$ with $\bar{U} = (\bar{u}_{st})$, $\bar{V} = (\bar{v}_{st})$, and

$\bar{D} = \text{diag}(\bar{d}_1,...,\bar{d}_p)$

and

$X_{(-j)} = \tilde{U}\,\tilde{D}\,\tilde{V}^T$ with $\tilde{U} = (\tilde{u}_{st})$, $\tilde{V} = (\tilde{v}_{st})$, and

$\tilde{D} = \text{diag}(\tilde{d}_1,...,\tilde{d}_{p-1})$

Now consider the predictor

$$\hat{x}_{ij}^m = \sum_{t=1}^{m} (\tilde{u}_{it}\sqrt{\tilde{d}_t})(\bar{v}_{tj}\sqrt{\bar{d}_t})$$

Each element on the right-hand side of this equation is obtained from SVD of X, mean-centered after omitting either the i-th row or the j-th column. Thus, the value x_{ij} has nowhere been used in calculating the prediction, and maximum use has been made of the other elements of X. Calculations here are exact, so there is no problem with convergence as opposed to expectation maximization approaches that have also been applied to AMMI, but are not guaranteed to converge.

On the other hand, the Gabriel method takes a mixture of regression and lower-rank approximation and utilizes $PRESS(m)$ and $PRECORR(m) = Corr(x_{ij}, \hat{x}_{ij}^m \mid \forall\ i,j)$ to achieve the same goal. The proposed algorithm for cross-validation of lower rank approximations is as follows.

For given (GEI) matrix X, use the partition

$$X = \begin{bmatrix} x_{11} & \mathbf{x}_{1.}^T \\ \mathbf{x}_{.1} & X_{\backslash 11} \end{bmatrix}$$

and approximate the submatrix $X_{\backslash 11}$ by its rank m fit using the SVD

$$X_{\backslash 11} = \sum_{k=1}^{m} \mathbf{u}_{(k)} \mathbf{d}_{(k)} \mathbf{v}_{(k)}^T = UDV^T,$$

where $U = [u_1,..., u_m]$, $V = [v_1,..., v_m]$, and $D = \text{diag}[d_1,..., d_m]$.
Then predict x_{11} by

$$\hat{x}_{11} = \mathbf{x}_{1.}^T V D^{-1} U^T \mathbf{x}_{.1}$$

and obtain the cross-validation residual $e_{11} = x_{11} - \hat{x}_{11}$.

Similarly obtain the cross-validation fitted values $\hat{x}_{ij}$ and residuals $e_{ij} = x_{ij} - \hat{x}_{ij}$ for all other elements x_{ij}, $i=1,...,g$, $j=1,...,m$; $(i,j) \neq (1,1)$. Each will require a different partition of X. These residuals and fitted values can be summarized by $PRESS(m)$ and $PRECORR(m)$ respectively.

With either method, choice of m can be based on some suitable function of $PRESS(m)$. However, the features of this statistic differ for the two methods. Gabriel's approach yields values that first decrease and then (usually) increase with m. He therefore suggests that the optimum value of m is the one that yields the minimum of the function. The Eastment-Krzanowski approach produces (generally) a set of values that is monotonically decreasing with m. The optimum value for m is the highest value of m at which the statistic W is greater than unity.

The Gollob (1968) approximate *F*-test assumes that $n\hat{\lambda}_k^2/\sigma^2$ is distributed as a chi-square variable, and he suggests using the statistic

$$F = \hat{\lambda}_m^2 / (f_1 \text{MS(Error mean)})$$

against an *F*-distribution with $f_1=g+e-1-(2m)$ and *ge(n-1)* degrees of freedom, or *g(e-1)(n-1)* degrees of freedom if blocks are present, to test the *m*-th multiplicative term of the model for significance. By contrast, Cornelius et al. (1992) uses

$$F_{R,m} = (\text{SS(GEI)} - \sum_{k=1}^{m} \hat{\lambda}_k^2) / (f_2 \text{MS(Error mean)})$$

against an *F*-distribution with $f_2=(g-1-m)(e-1-m)$ and *ge(n-1)* degrees of freedom, or *g(e-1)(n-1)* degrees of freedom if blocks are present, to test significance of lack of fit of the *m*-term model.

Piepho (1995) investigated the robustness (to the assumptions of homogeneity and normality of the errors) of some alternative tests to select an AMMI model. He comments that *F* tests applied in accordance with Gollob's (1968) criterion are liberal in that they select too many multiplicative terms. Of the four methods he studied, including that of Gollob (1968), the test proposed by Cornelius et al. (1992) was the most robust. The author thus recommends that preliminary evaluations should be conducted to verify the validity of the assumptions if one of the other tests is to be used. In our study, the tests of Gollob and Cornelius were performed at α = 0.05. We note that other possible tests of both *PRESS* and *F* form exist (e.g. the *PRESS* test in Cornelius et al., 1996, and the *F* test based on the Goodman-Haberman theorem described by Cornelius, 1993), but we have not included them in the comparisons to keep the investigation to a manageable size.

Simulation Study

We studied by simulation the behaviour of the above methods of selecting the number of interaction components in AMMI models. However, to make the simulations realistic, we based them around real data sets for which the model is appropriate, according to the following scheme.

Suppose that $\mathbf{Y} = [y_{ij}]$ is the two-way table for a chosen data set, and that we fit the AMMI model

$$y_{ij} = \mu + g_i + e_j + \sum_{k=1}^{m} \alpha_{ik} \lambda_k \gamma_{jk} + \varepsilon_{ij}$$

for a given value of *m*. Let an over-bar denote average, with '.' indicating the subscripts over which summation takes place. Therefore, $\hat{\mu} = \bar{y}_{..}$, $\hat{g}_i = \bar{y}_{i.} - \bar{y}_{..}$, $\hat{e}_j = \bar{y}_{.j} - \bar{y}_{..}$ and $\alpha_{ik}\lambda_k\gamma_{jk}$ is estimated from the *k*-th

component of the SVD of $\mathbf{GE} = \bar{y}_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}$. The model assumes $\varepsilon_{ij} \sim N(0,\sigma^2)$, and σ^2 is estimated from the error sum of squares in the ANOVA of the data set.

To generate a data set in the simulation study for a chosen value of *m*, we simply form the values

$$\hat{y}_{ij} = \hat{\mu} + \hat{g}_i + \hat{e}_j + \sum_{k=1}^{m} \hat{\alpha}_{ik} \hat{\lambda}_k \hat{\gamma}_{jk} + \hat{\varepsilon}_{ij}$$

where the $\hat{\varepsilon}_{ij}$ are independently drawn from a $N(0,\hat{\sigma}^2)$ distribution. By these means it is possible to generate repeated data sets which mimic the patterns of the original data, but which have built in to them a specified number of interaction components (governed by the initial choice of *m*). Also, by varying some of the parameter values, it is possible to force particular structures on the data. For example, by changing the values of $\hat{\lambda}_i$ we can force some components to have a greater or smaller contribution to the interaction. We used both of these devices in the simulations reported below.

Wheat Yield Data

Our main study uses as a basis the wheat yield data, Trial 3, given by Cornelius & Crossa (1999) who describe five multi-environment (MET) international cultivar trials, all set in randomized block design. Trial 3 concerns maize (*Zea mays* L.), with 9 cultivars, 20 sites and n = 4.

The first series of simulation experiments was conducted to investigate the sensitivity of the methods to increasing numbers of interaction components in the data. We set the number of interaction components *m* successively equal to *0, 1, 2,...,9* and then used all estimated parameters from the original data, and the scheme outlined above, to produce the simulated data. We generated 100 repetitions for each of these 10 experiments, and used the four methods described above to determine in each case how many interaction components should be included in the model. Table 1 reports the value chosen most frequently for each method and each experiment.

We then conducted three further series of experiments, constraining the emphasis of components in each case as suggested in the previous section. In the first series of experiments, we arranged for each of the nine interaction components to have approximately equal weight. To do this we arranged that $\hat{\lambda}$ values had the following pattern: $\hat{\lambda}_1^2 = c$, $\hat{\lambda}_2^2 = 0.99c$, $\hat{\lambda}_3^2 = 0.98c$, ..., $\hat{\lambda}_9^2 = 0.92c$, but with c chosen such that $\sum_{i=1}^{9} \hat{\lambda}_i^2$ equalled the original GEI sum of squares and $\hat{\lambda}_1^2 \geq \hat{\lambda}_2^2 \geq ... \geq \hat{\lambda}_9^2$. Once again 100 repetitions were provided for each experiment, and modal

Table 1 - Modal number of AMMI multiplicative terms chosen by the Eastment-Krzanowski, Gabriel, Gollob and Cornelius methods and the original eigenvalues squared, with the 'true' singular values put equal to the original empirical singular values; 100 repetitions of simulated data based on Trial 3 results in Cornelius & Crossa (1999).

m	Eastment-Krzanowski	Gabriel	Gollob	Cornelius	Original singular values squared
0	1	1	1	1	140.32
1	2	2	2	1	37.71
2	2	3	2	2	26.05
3	2	4	3	2	21.53
4	2	3	4	3	12.37
5	1	3	5	4	6.90
6	1	3	5	4	2.51
7	1	3	5	4	2.29
8	1	2	5	-	0
Original data set	2	7	2	2	

choices of number of interaction components are presented for each method and each experiment in Table 2.

In the second series of experiments, we inflated the first three $\hat{\lambda}_i^2$ to have high values and reduced the remaining $\hat{\lambda}_i^2$, but in proportion to their actual values so that $\sum_{i=1}^{9} \hat{\lambda}_i^2$ was again equal to the original sum of squares of GEI. In the third series of experiments, the first three $\hat{\lambda}_i^2$ values were again inflated while the remaining $\hat{\lambda}_i^2$ values were reduced, but the latter were now set to be approximately constant. Results from these two series are given in Tables 3 and 4 respectively.

As a final study we used the data of Trial 5 from Cornelius & Crossa (1999) as the basis for our simulations. This was a simple trial, and considered by the original author to be relatively free of interaction effects. Here we have 8 cultivars and 59 sites. We therefore only conducted the unconstrained form of experiments (i.e. as in Table 1), and show the results in Table 5.

DISCUSSION

We admit at the outset that our investigations are by no means exhaustive (a comparison of just four methods on a relatively small set of conditions), and are not designed in such a way that 'optimal' methods can be determined with respect to some well-defined mathematical objective function. Instead, we start from the view that there may often be conflict in practice among the methods, and that what is required by the practitioner is some reassurance that a chosen method is likely to perform 'reasonably'. So our conclusions in this discussion are essentially qualitative in nature, measured against what we believe to be reasonable interpretations of each studied situation.

Table 2 - Modal number of AMMI multiplicative terms chosen by the Eastment-Krzanowski, Gabriel, Gollob and Cornelius methods; 100 repetitions of simulated data based on Trial 3 results in Cornelius & Crossa (1999), 'true' singular values nearly equal.

M	Eastment-Krzanowski	Gabriel	Gollob	Cornelius
0	1	1	1	1
1	2	2	2	1
2	3	3	2	2
3	3	4	3	3
4	1	5	4	4
5	1	6	5	5
6	1	7	6	6
7	1	1	7	7
8	1	1	8	-

Table 3 - Modal number of AMMI multiplicative terms chosen by the Eastment-Krzanowski, Gabriel, Gollob and Cornelius methods; 100 repetitions of simulated data based on Trial 3 results in Cornelius & Crossa (1999), three large singular values and the rest small, but proportional to the empirical values from analysis of the original data.

M	Eastment-Krzanowski	Gabriel	Gollob	Cornelius
0	1	1	1	1
1	2	2	2	1
2	3	3	2	2
3	3	4	3	3
4	3	4	3	3
5	3	4	3	3
6	3	4	3	3
7	3	4	3	3
8	3	4	3	-

The four methods of selecting the number of multiplicative interaction components have indeed yielded different results on the simulated data sets used here. Moreover, the results for these tests give conflicting recommendations on a particular *m* value for a data set. This feature can be seen clearly in Tables 1-5.

In general terms, the two methods based on *F*-tests are conservative and are prone to select a fairly high value of *m*. The Eastment-Krzanowski method, on the other hand, is clearly the most parsimonious: it never selects a value of *m* greater than 3, and usually recommends either $m = 1$ or $m = 2$. We note that this accords closely with the view of Snee (1982) as regards an adequate description of nonadditivity.

Table 4 - Modal number of AMMI multiplicative terms chosen by the Eastment-Krzanowski, Gabriel, Gollob and Cornelius methods; 100 repetitions of simulated data based on Trial 3 results in Cornelius & Crossa (1999), three large singular values and the rest small and all equal.

M	Eastment-Krzanowski	Gabriel	Gollob	Cornelius
0	1	1	1	1
1	2	2	2	1
2	3	3	2	2
3	3	3	3	2
4	3	3	4	2
5	1	3	4	3
6	1	3	4	3
7	1	3	4	3
8	1	3	4	-

Passing to particular results, consider the first experiment. The number of interaction components chosen by each method from a single analysis of the original data (as reported by Dias & Krzanowski, 2003) is given at the foot of Table 1: all methods suggest two components, except Gabriel's method which suggests seven components. Table 1 then shows the modal number of components chosen by each method as progressively more interaction terms are included in the simulated data. Since there appear to be only about two interaction components in the original data, and since the components are added according to their ranking in order of decreasing values (last column of Table 1), we would expect the addition of the third and succeeding components to have little effect. The four methods show different patterns of behaviour. The Eastment-Krzanowski method is very stable, choosing an *m* value of 2 for the original data compared with 1 or 2 for all simulations. Gabriel's and Gollob's methods are very volatile, the former contrasting $m = 7$ for the original data with values 1, 2, 3 or 4 for the simulations, and the latter having all values between one and five in the simulations, but $m = 2$ for the original data. Cornelius' method is less volatile but still exhibits a range of values from 1 to 4.

Consider next the simulations with constrained interaction components. In Tables 3 and 4, where the interaction components have been constrained by heavily weighting the first three components, it is probably reasonable to say that no method should choose a value greater than 3. The methods of Eastment-Krzanowski and Cornelius accord with this, while those of Gabriel and Gollob recommend a value of 4 for a substantial proportion of cases. However, in both cases we would expect the highest value reached to then stay con-

Table 5 - Modal number of AMMI multiplicative terms chosen by the Eastment-Krzanowski, Gabriel, Gollob and Cornelius methods and the original eigenvalues squared; 100 repetitions of simulated data based on Trial 5 results in Cornelius & Crossa (1999).

m	Eastment-Krzanowski	Gabriel	Gollob	Cornelius	Original singular values squared
0	1	1	1	1	65.53
1	1	2	2	1	48.04
2	1	3	2	2	43.56
3	1	4	3	3	37.79
4	1	5	4	3	23.85
5	1	1	5	4	22.78
6	1	1	5	5	21.61
7	1	1	6	6	0
Original data set	1	1	2	3	

Abbreviations: AMMI, additive main effects and multiplicative interaction model; ANOVA, analysis of variance; GEI, genotype x environment interaction; MET, multi-environment trials; MS, mean square, PRESS, predictive sum of squares; PRECORR, predictive correlation; SS, sum of squares; SVD, singular value decomposition.

stant as more components are added. Here the Eastment-Krzanowski method fails in Table 4, where m drops back down from 3 to 1 from $M = 5$ onwards.

In Table 2 the components have a fairly even weighting, so perhaps we would expect there to be a monotonically increasing value of m as M rises. All methods apart from that of Eastment-Krzanowski do indeed exhibit such an increasing set of values, whereas this method again drops back from 3 to 1 at $M = 4$ onwards.

Finally, Table 5 reports a case where terms are successively added in the simulations as m increases, but where the originator of the data believed there to be little interaction. The very even distribution and low values of the squared eigenvalues (last column of Table 5) support this view. The Eastment-Krzanowski method suggests just one component (the minimum value possible) for all M values, while all other methods exhibit an upward trend in number of components chosen as m increases, and select very large values for at least some of these cases. Tables 1-5 do not present a modal number of AMMI multiplicative terms for the Cornelius method from m equal to eight because of negative degrees of freedom.

In summary, therefore, it is a rather mixed picture. The Eastment-Krzanowski method is stable and behaves appropriately in the presence of a relatively small number of 'important' components, but generally underestimates when there is a larger number. The Gabriel method is very volatile and tends to choose slightly too many components in many situations. Of the F-test methods, the one by Cornelius exhibits appropriate behaviour in the presence of 'important' components, but is less stable than Eastment-Krzanowski. Finally, the Gollob version is broadly similar to the Cornelius method but with slightly worse stability, and is likely to choose more components in some situations. So, finally, the Eastment-Krzanowski method appears to be the preferable cross-validatory one while the Cornelius method would be the recommended F-test method. The former is preferable to the latter if parsimony is a major consideration, but the latter is preferable if the presence of a large number of interaction components is suspected. Given the different performances of the two methods, however, both should perhaps be used before making a decision in a practical application. These conclusions support those arrived at by Dias & Krzanowski (2003) from the analysis of real data.

ACKNOWLEDGEMENTS

REFERENCES

CORNELIUS, P.L. Statistical tests and retention of terms in the additive main effects and multiplicative interaction model for cultivar trials. **Crop Science**, v.33, p.1186-1193, 1993.

CORNELIUS, P.L.; CROSSA, J Prediction assessment of shrinkage estimators of multiplicative model for multi-environment cultivar trials. **Crop Science**, v.39, p.998-1009, 1999.

CORNELIUS, P.L.; SEYEDSADR, M.; CROSSA, J. Using the shifted multiplicative model to search for "separability" in crop cultivar trials. **Theoretical and Applied Genetics**, v.84, p.161-172, 1992.

CORNELIUS, P.L.; CROSSA, J. SEYEDSADR, M. Statistical tests and estimators of multiplicative models for genotype-by-environment interaction. In: KANG, M.S.; GAUCH, H.G. (Ed.) **Genotype-by-environment interaction.** Boca Raton: CRC Press, 1996. p.199-234.

CORSTEN, L.C.A.; VAN EIJNSBERGEN, C.A. Multiplicative effects in two-way analysis of variance. **Statistica Neerlandica**, v.26, p.61-68, 1972.

DANIEL, C. **Application of statistics to industrial experimentation**. New York: John Wiley, 1976.

DIAS, C.T.S.; KRZANOWSKI, W.J. Model selection and cross-validation in additive main effect and multiplicative interaction (AMMI) models. **Crop Science**, v.43, p.865-873, 2003.

EASTMENT, H.T.; KRZANOWSKI, W.J. Cross-validatory choice of the number of components from a principal component analysis. **Technometrics**, v.24, p.73-77, 1982.

GABRIEL, K.R. Le biplot – outil d'exploration de données multidimensionelles. **Journal de la Societe Francaise de Statistique**, v.143, p.5-55, 2002.

GOLLOB, H.F. A statistical model which combines features of factor analytic and analysis of variance techniques. **Psychometrika**, v.33, p.73-115, 1968.

GOOD, I.J. Some applications of the singular decomposition of a matrix. **Technometrics**, v.11, p.823-831, 1969.

HEGEMANN, V.; JOHNSON, D.E. On analyzing two-way AOV with interaction. **Technometrics**, v.18, p.273-281, 1976.

JOHNSON, D.E.; GRAYBILL, F.A. An analysis of a two-way model with interaction and no replication. **Journal of the American Statistical Association**, v.67, p.862-868, 1972.

KRZANOWSKI, W.J. Cross-validation in principal component analysis. **Biometrics**, v.43, p.575-584, 1987.

MANDEL, J. The partitioning of interaction in analysis of variance. **Journal of Research of the National Bureau of Standards B. Mathematical Sciences**, v.73, p.309-328, 1969.

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

PIEPHO, H.P. Robustness of statistical test for multiplicative terms in additive main effects and multiplicative interaction model for cultivar trial. **Theoretical and Applied Genetics**, v.90, p.438-443, 1995.

SNEE, R.D. Nonadditivity in a two-way classification: is it interaction or nonhomogeneous variance? **Journal of American Statistical Association**, v.77, p.515-519, 1982.

WOLD, S. Pattern recognition by means of disjoint principal component models. **Pattern Recognition,** v.8, p.127-139, 1976.

WOLD, S. Cross-validatory estimation of the number of components in factor and principal component models. **Technometrics**, v.20, p.397-405, 1978.

YOCHMOWITZ, M.G.; CORNELL, R.G. Stepwise tests for multiplicative components of interaction. **Technometrics**, v.20, p.79-84, 1978.