Theor Appl Genet (2005) 110: 812–818
DOI 10.1007/s00122-004-1822-6

ORIGINAL PAPER

R. M. Pacheco · J. B. Duarte · R. Vencovsky
J. B. Pinheiro · A. B. Oliveira

Use of supplementary genotypes in AMMI analysis

Received: 20 November 2003 / Accepted: 14 September 2004 / Published online: 3 February 2005

Abstract Improving stability of crop yield in a target production environment is an important breeding objective. It is well known that selection for better stability generally results in lower mean yields and, conversely, that selection for higher mean yields may lead to poorer stability. This paper explores the equivalence between the singular value decomposition used in AMMI analysis and the spectral decomposition used in principal components analysis. This equivalence enables scores of a ''supplementary genotype'' made up of the highest yield value within each environment to be obtained, and these may serve as the ideal check treatment for selection purposes. These scores are used to (1) display this check in a biplot graph, thereby providing a qualitative comparison with the real genotypes related to their interaction with environments; (2) obtain estimates of the squared distances from the projection of each real genotype to the projection of the ''supplementary treatment'', thereby allowing conclusions to be made on the yield stability of each real genotype. This procedure was effective in identifying the most stable soybean cultivars in an example shown for illustration.

Communicated by H.H. Geiger

R. M. Pacheco · J. B. Duarte (✉)
Escola de Agronomia e Engenharia de Alimentos,
Setor de Melhoramento Vegetal,
Universidade Federal de Goiás, Goiânia-Goiás, Brazil
E-mail: jbduarte@uol.com.br

R. Vencovsky · J. B. Pinheiro
Escola Superior de Agronomia Luiz de Queiroz,
Instituto de Genética, Universidade de São Paulo,
São Paulo, Brazil

A. B. Oliveira
Agencia Rural do Estado de Goiás,
Goiânia-Goiás, Brazil

Introduction

Agricultural yield is strongly influenced by environmental conditions that generally lead to wide variations in yield, both among years in a location and among locations in a single year or, even further, between locations and years. Improving yield stability of an agricultural crop throughout a production region is an important objective of genetic breeding programs. Unfortunately, it is not uncommon to have situations where selection based on yield stability causes lower mean yields (Finlay and Wilkinson 1963; Helms 1993) and, conversely, where selection for higher means results in less stability (Simmonds 1991). However, Holland et al. (2002) reported that a selection for wide adaptation to various environments in oat resulted in a significant increase in mean grain yield in the population.

Phenotypic stability parameters of interest include, among others, the Eberhart and Russell (1966) regression coefficient (b_i) and the Lin and Binns (1988) superiority parameter (P_i). The Eberhart and Russell parameter is based on the regression of the phenotypic response (for example, yield) of each genotype on a quantitative index of environmental quality (for example, mean yield of each environment). In this approach, a stable genotype has a unitary regression coefficient, minimum deviations from the regression and a high mean yield. In turn, the Lin and Binns (1988) superiority parameter is the quadratic difference between the yield of a given genotype and the maximum yield obtained within each one of the environments under study. Genotypes with broad adaptation present lower values of this superiority parameter because they have a yield close to the maximum yield within each environment; this in contrast to genotypes less adapted to the set of environments that comprise the cropping region in question.

An alternative form of assessing the phenotypic stability and adaptability involves a multivariate

treatment of the genotype×environment interactions (G×E). Additive main effects and multiplicative interactions (AMMI) analysis is used for partitioning the effects traditionally attributed to this interaction, which represent the residues of the isolation of the main genotype and environment effects estimated by analysis of variance (ANOVA) (Crossa 1990; Gauch 1992; Duarte and Vencovsky 1999). In this case, differences in genotype stability and adaptability to the environments can be assessed qualitatively using graphics named biplots that place the environments according to their principal component scores obtained from the analysis.

The AMMI procedure uses a combination of statistical techniques, namely, ANOVA (univariate) and principal components analysis (PCA) (multivariate). For the purpose of generalization and to prevent against possible interpretive mistakes, Duarte and Vencovsky (1999) recommended treating PCA directly by the mathematical technique called singular value decomposition (SVD). This procedure, derived from matrix algebra, consists of the theoretic base of the biplot graphic representation, which has application in various areas of knowledge (Eckart and Young 1936; Good 1969; Gabriel 1978; Mandel 1982).

In PCA, the introduction of supplementary individuals or variables (vectors) that have not participated directly in the original analysis is formally described. These vectors are rows and/or columns added to the original matrix of multivariate data (***X***). The objective of adding these variables is the possibility of representing virtually interesting individuals (rows) or variables (columns) which are not part of the real data mass assessed. Thus, by projecting them, for example, on the first principal planes (dispersion graphs of individuals in the space of the variables), the positioning of these individuals in relation to the others, the real individuals, may help in data interpretation (Nagpal 1999). It is necessary to point out that the new graphic representation (with supplementary vectors) should be constructed so as not to affect the original covariance structure of the data, that is, to keep the relative positions of the real individuals and variables.

The purpose of this study is to present the theoretical base of the representation of a supplementary genotype, not in exclusive dispersion graphics of the individuals, but in the so-called biplot (simultaneous graphic representation of the rows and columns of a matrix; that is, in the context of this study, the genotypes and environments) obtained from the AMMI analysis. Thus, defining the construction criteria of this supplementary genotype as being that which reaches the highest yield in each location, the proximity or distance of the real genotypes and environments can be assessed in relation to this reference genotype, which has the characteristics of an ideal control and permits the application of this proposal to real data applied plant breeding programs.

Materials and method

Experimental data

As an illustration, grain yield data were used from a group of 11 competition trials (environments) of soybean cultivars (Table 1) obtained by Oliveira et al. (2003). The genotypes were developed in the soybean breeding program carried out at the Escola de Agronomia e Engenharia de Alimentos, Universidade Federal de Goiás, Brazil. The lines were in the F_{10} generation of selfing (high homozygosis level). Eighteen genotypes were tested, including four control cultivars (FT-2000, Emgopa-315, Msoy-8001 and Conquista). In each environment we used a randomized complete block design with three replications.

The AMMI analysis model

The graphic representation of genotypes and environments by AMMI analysis results from a model of main additive effects and multiplicative interaction. This model is expressed mathematically by:

$$Y_{ij} = \mu + g_i + e_j + \sum_{k=1}^{n} \lambda_k \gamma_{ik} \alpha_{jk} + \rho_{ij} + \varepsilon_{ij}, \text{where:}$$

Y_{ij} is the yield response of genotype i ($i = 1, 2, .., I$) in environment j ($j = 1, 2, ..., J$);

μ is the general mean;

g_i is the main effect associated to the i-th genotype;

e_j is the main effect associated to the j-th environment;

n is the number of principal axes (principal components) necessary to describe the "pattern" of the interaction between the i-th genotype with the j-th environment;

λ_k is the singular value of the k-th principal interaction axis;

γ_{ik} is the i-th element of the singular column vector associated to axis k;

α_{jk} is the j-th element of the singular row vector associated to axis k;

ρ_{ij} is the AMMI residue (interaction "noise"); and

ε_{ij} is the pooled error term.

Note that the classic term $(ge)_{ij}$, traditionally referred as the interaction of the i-th genotype with j-th environment, is modeled here by: $\sum_{k=1}^{n} \lambda_k \gamma_{ik} \alpha_{jk} + \rho_{ij}$.

Table 1 Matrix with the observed yield means (kg/ha) of 18 soybean genotypes in 11 environments. Source: Oliveira et al. (2003)

Genotype	Environments										
	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11
G_1 Conquista	2,047	1,703	2,088	1,195	3,809	2,716	2,567	2,969	2,773	2,750	2,418
G_2 Emgopa 315	1,918	1,865	1,576	1,085	2,380	2,769	2,667	3,033	2,617	2,580	2,352
G_3 FT-2000	1,367	1,176	1,020	803	2,646	2,000	1,950	3,191	2,717	2,972	3,055
G_4 L-7	2,208	2,147	1,585	1,276	2,478	2,264	2,390	2,268	1,986	2,507	1,831
G_5 L-8	1,935	1,794	1,395	1,327	2,488	2,518	1,865	2,187	1,947	1,607	2,067
G_6 L-9	1,494	1,669	2,048	1,437	3,033	2,396	2,721	2,333	2,174	2,372	2,078
G_7 L-11	1,797	1,837	1,731	1,345	2,246	2,571	2,499	2,690	2,355	1,869	2,351
G_8 L-112	1,345	1,503	1,565	993	3,011	2,235	2,359	2,330	1,536	2,180	2,093
G_9 L-113	1,682	2,233	1,518	945	2,524	2,189	2,312	2,231	1,864	2,133	2,046
G_{10} L-12	1,859	2,075	1,837	1,107	2,250	2,271	2,325	2,364	2,074	2,182	2,243
G_{11} L-13	1,991	2,133	1,788	1,183	2,571	2,300	2,513	2,487	2,596	2,233	2,201
G_{12} L-14	2,021	1,981	1,773	1,339	2,515	2,417	2,475	2,673	2,222	2,367	2,114
G_{13} L-149	1,665	1,247	1,522	846	2,404	1,951	2,213	1,679	1,417	2,209	1,947
G_{14} L-16	1,745	1,888	1,845	1,337	2,643	2,443	2,600	2,465	2,187	2,554	2,205
G_{15} L-176	1,619	1,443	1,529	956	2,797	2,047	2,323	1,925	1,621	2,144	2,056
G_{16} L-21	1,519	1,539	1,561	1,009	2,375	2,115	2,188	2,193	2,081	1,946	1,913
G_{17} L-84	1,813	2,073	1,464	1,169	2,292	3,182	1,971	2,361	2,046	2,208	2,006
G_{18} Msoy 8001	1,588	1,571	2,020	903	2,203	2,821	3,184	2,843	2,593	2,396	2,361

The model is fitted sequentially, combining the ANOVA and the PCA. Thus the residuals from the fitting of the main effects by ANOVA are modeled, in a second step, by PCA. In this stage, the scores or coordinates of the genotypes and environments are produced on the principal interaction axes, conventionally known as $IPCA_k$ ($k = 1, 2,..., n$), that permit their representation together in a biplot graph.

To avoid possible interpretative mistakes, resulting from the exchange of sign of the genotype and environment scores, by PCA, some authors recommend treating the second stage directly by SVD (Gauch 1992; Duarte and Vencovsky 1999).

Equivalence between SVD and PCA

SVD enables an approximation of a matrix to be obtained by another of a lower rank to understand the structure behind the data, that is, the law of matrix formation. Mathematically, the exact partitioning of a real matrix $B_{(rxc)}$ is given by:$B = \sum_{k=1}^{p} \lambda_k u_k v'_k$, where λ_k is its k-th singular value [$k = 1, 2,..., n,..., p$; with $p \le \min(r, c)$ being the rank of B]; u_k and v'_k are the singular associated vectors, column and row, respectively. An approximation of rank n to B is determined by:$\tilde{B} = \sum_{k=1}^{n} \lambda_k u_k v'_k$.

Arranging the p singular values in decreasing order on a diagonal matrix $S_{(p)}$, and the respective singular vectors u_k and v'_k in two matrices (orthogonal) $U_{(rxp)}$ and $V'_{(pxc)}$, respectively, the SVD expression can be described as: $B = USV'$ (having U, sequentially the p vectors u_k in their columns, and V' the p vectors v'_k in their rows). In this case, the approximated matrix that captures the subjacent pattern of the B data is determined by:$\tilde{B} = \tilde{U}\ \tilde{S}\ \tilde{V}'$; where$\tilde{U}, \tilde{S}$ and $\tilde{V}'$ contain only the n first term of their corresponding matrices. From this, Gabriel (1971) derived the so-called row (G) and column (H) markers of the matrix B, that is, $G = \tilde{U}\ \tilde{S}^{1/2}$ and $H = \tilde{S}^{1/2}\tilde{V}'$, which can be represented together in an n-dimensional graphic ($n \le 3$) called a biplot.

In the context of the AMMI analysis, the matrix to be partitioned and understood in its formation law will be the G×E interaction matrix. Therefore,$B_{(I\times J)} = (\hat{ge})_{ij}$, where its elements$(\hat{ge})_{ij}$ correspond to the least squares estimates of the term $(ge)_{ij}$ in the conventional model. Under the conditions of a fixed effects model, these estimates are:$(\hat{ge})_{ij} = Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..}$, where $\bar{Y}_{i.}$, $\bar{Y}_{.j}$ and $\bar{Y}_{..}$ denote the means of genotype i, of environment j and the general mean, respectively. Since it is a deviation matrix with nil sums on the rows and column, the rank of B is reduced to $p = \min\{I-1, J-1\}$.

The correspondence established in the literature between the AMMI analysis and PCA can be demonstrated from two symmetrical matrices derived from B: BB' and $B'B$. The spectral decomposition, which is the particular case of the SVD of symmetrical matrices and the theoretical foundation of PCA for these matrices, is given by: $BB'_{(I)} = U\Lambda U' = \sum_k \lambda_k^2 u_k u'_k$ and $B'B_{(J)} = V\Lambda V' = \sum_k \lambda_k^2 v_k v'_k$, where Λ is the diagonal matrix with the not nil eigenvalues λ_k^2 that are common to BB' and $B'B$ (λ_k^2 is exactly equal to the square of the respective singular value of B), and U and V, already defined, are here the respective matrices of the eigenvectors of BB' and $B'B$.

In the case of PCA, the scores of the individuals (identified by the rows of B) for the representation in a biplot graph are obtained by spectral decomposition applied to the matrix of covariance of the variables. The rule that weights each B vector-row of the original matrix for the principal component axes ($IPCA_k$) is, therefore, given by the eigenvectors of $B'B$, that is, the V matrix. Consequently, the scores of the individuals for

the n first axes is determined by $G_{(r\times n)} = B\tilde{V}$. Similarly, although not commonly described in the literature, the rule that weights each B vector-column for the principal axes is given by the eigenvectors of the BB' (matrix U), which is associated to the individual's covariance. Thus, the matrix with the scores of the variables for the n first principal axes is obtained by: $H_{(n\times c)} = \tilde{U}'B$. For a biplot representation, however, the two decompositions (double PCA) cannot be done independently, but rather in an integrated manner.

This warning with respect to the use of a doubly-integrated PCA results from the fact that the independent analyses can produce row and column scores corresponding to different graphic representations (generally symmetrical, that is, with scores showing changed signs). Although this does not interfere in the dispersion of the individuals, nor in the correlations of the variables, it certainly implies mistaken interpretations about association among individuals and variables in the biplot. The solution to the problem can be obtained by calculating the BB' eigenvectors from the $B'B$ eigenvectors, or vice-versa, instead of obtaining them independently. Thus, being that v_k is the k-th $B'B$ eigenvector associated to the eigenvalue λ_k^2, the k-th BB' eigenvector is given by $u_k = (1/\lambda_k)Bv_k$. Otherwise, knowing the u_k eigenvector, v_k can be obtained by $v_k = (1/\lambda_k)B'u_k$ (Duarte and Vencovsky 1999; Souza 1988).

Apart from an easily solved difference in scale it is possible to show that the two approaches permit the same relative individual and variable scores to be obtained for the biplot representation. The first has the advantage that these scores are naturally integrated, because they are obtained from a single decomposition (SVD); the other involves two applications of spectral decomposition, thus the PCA needs to be doubly-integrated. On the other hand, as previously presented, only the second approach supplies a rule to directly weight a row or column vector observed to produce its respective scores. Thus, if the objective is to apply the same rule to represent graphically, in the same biplot, an extra individual (a supplementary genotype here), the expressions derived from the last approach offer advantages.

Supplementary genotype representation

First, based on the proposal by Lin and Binns (1988), where the standard is the maximum response for each location and the superiority of a cultivar is measured by the quadratic distance measured from this response, the maximum yields in each location were aligned as belonging to a supplementary genotype of interest (ideal control). Thus, if Y_{ij} is the yield of the i-th genotype in the j-th environment and Y_{sj} is the maximum response among all of the genotypes in the j-th environment, the set of the values Y_{sj} ($j = 1, 2... J$) produces the vector associated to the supplementary genotype s.

It should be emphasized that the position of this "ideal control" in a biplot graph must be calculated so as not to exercise any influence on the graphic coordinate of the genotypes (individuals) and environments (variables) assessed. The procedure to do this is based on the same rule available for producing the scores of the real genotypes assessed in this paper, that is, $G_{(I\times n)} = B\tilde{V}$, where B is the G×E interaction matrix and $\tilde{V}$ is the orthogonal matrix with the n first eigenvectors of $B'B$ placed in its columns. Thus, the coordinate or score of a specific genotype i on a principal axis k ($k = 1, 2... n$) is determined by: $\text{IPCA}_{ik} = b_i'v_k = \sum_{j=1}^{J} b_{ij}v_{jk}$, where b_i is the row vector of interest in B with elements b_{ij}. This procedure was adapted from Droesbeke and Fine (1995).

In the AMMI analysis, the scores of the real genotypes for their placing on a biplot graph are obtained by $G = \tilde{U}\,\tilde{S}^{1/2}$, where $\tilde{U}$ is a matrix that has sequentially the n first vectors u_k in its columns and each vector has as many rows as the number of real genotypes. Thus, the coordinates of a supplementary genotype (u_{sk}) can be estimated in the same way. Given that $u_k = (1/\lambda_k)\ Bv_k$, each of its elements is given by $u_{ik} = (1/\lambda_k b_i')v_k$. Thus, in the AMMI analysis, we have: $u_{sk} = (1/\lambda_k)(\hat{ge})_s'v_k$; where λ_k is the singular value obtained with the real genotypes and $(\hat{ge})_s'$ is the vector $1\times J$ with the estimates $(\hat{ge})_{sj}$ associated to the supplementary genotype (G_s), with $(\hat{ge})_{sj} = Y_{sj} - \bar{Y}_{s.} - \bar{Y}_{.j} + \bar{Y}_{..}$. Then the coordinate of the genotype will be determined by $\text{IPCA}_{sk} = \sqrt{\lambda_k}u_{sk}$, which is equivalent to $(1/\sqrt{\lambda_k})(\hat{ge})_s'v_k$, being that the referred difference in scale between the two approaches by PCA and SVD corresponds exactly to the value $1/\sqrt{\lambda_k}$.

Genotype adaptability and stability

The criterion of vector construction associated to the supplementary genotype here considered, which is composed of the highest yield obtained in each one of the locations, simulates a "control genotype" of optimum yield adaptation, relative to the set assessed, and one of high agronomic stability, in the sense that this virtual genotype explores the positive interactions with environments. Thus in a biplot, the distance between the points corresponding to an assessed genotype and the supplementary genotype provides a measure of similarity between these genotypes according to this concept. A quadratic distance can be easily calculated by: $d_{is}^2 = \sum_{k=1}^{n} (\text{IPCA}_{ik} - \text{IPCA}_{sk})^2$.

To execute the analyses, a computer routine in Statistical Analysis System/Interactive Matrix Language (SAS/IML) language was developed, based on a program made available by Duarte and Vencovsky (1999).

Results and discussion

As presented by Oliveira et al. (2003), the model selected to describe the set of data used here was the AMMI2 (Table 2). This means that two principal interaction axes

Table 2 Joint ANOVA and AMMI decomposition of the G×E interaction effects, on a treatment mean basis, for soybean grain yield (2001/2002). Source: Oliveira et al. (2003)

Sources of variation	*df*	SS	MS	*P* > *F*
Interaction G×E	151	12,891,459.05	853.90	0.0006
AMMI1	26	4,715,627.11	181,370.27	< 0.0001
Residual	125	8,175,831.94	65,406.66	0.0273
AMMI2	24	3,140,704.47	130,862.69	0.0001
Residual	101	5,035,127.47	49,852.75	0.4231
Pooled error	216	–	48,375.58	–

were needed to describe the pattern behind the main effects residues (matrix *B*). Consequently, the graph that would normally be constructed to represent the genotypes and environments jointly is the AMMI2 biplot, whose coordinates are shown in Table 3. For illustration in this study, bearing in mind its methodological character, both the AMMI1 biplot (means×IPCA1) as the biplot AMMI2 (IPCA1×IPCA2) are presented (Figs. 1 and 2, respectively).

It is observed that with the approach presented in this paper, one or more supplementary genotypes can be represented in AMMI biplot displays and that this representation can help in the interpretation of data obtained from variety trials. In this case, the supplementary genotype of theoretical interest should be one with maximum yield in each location, according Lin and Binns (1988). In both figures it is easy to notice the proximity between the G_1 genotype (Conquista cultivar) and the supplementary genotype of interest (G_s), here conceived as being the one that reaches the highest mean yield in each location sampled by the experiments (which has positive interactions with environments). This cultivar was emphasized by Oliveira et al. (2003) in their study as having the most favorable yield performance when compared to the other genotypes assessed.

The interpretation of a biplot for G×E interaction is made based on the magnitude and on the signs of the scores of the genotypes and the environments for the interaction axes considered. Values of close to zero characterize genotypes and environments that have a low contribution to the interaction, being considered stable. In the AMMI1 biplot (Fig. 1) the points around the line that marks the points zero on the IPCA1 axis correspond to the most stable genotypes and environments. These include G_6, G_7, G_8, G_{12}, G_{14}, G_{15} and G_{16} and environments A_5 and A_7. In the AMMI2 biplot (Fig. 2) stable genotypes and environments are those whose points are situated close to the origin, that is, with scores close to zero for the IPCA1 and IPCA2 axes. The G_{14} and G_{16} genotypes and the A_7 environment are in this situation. The stability of the genotypes is an indication of their adaptive amplitudes, that is, stable genotypes are widely adapted to the environments being tested. However, for cultivar recommendation purposes, stable genotypes should also have desirable characteristics such as high yield. Given the definition of a supplementary genotype

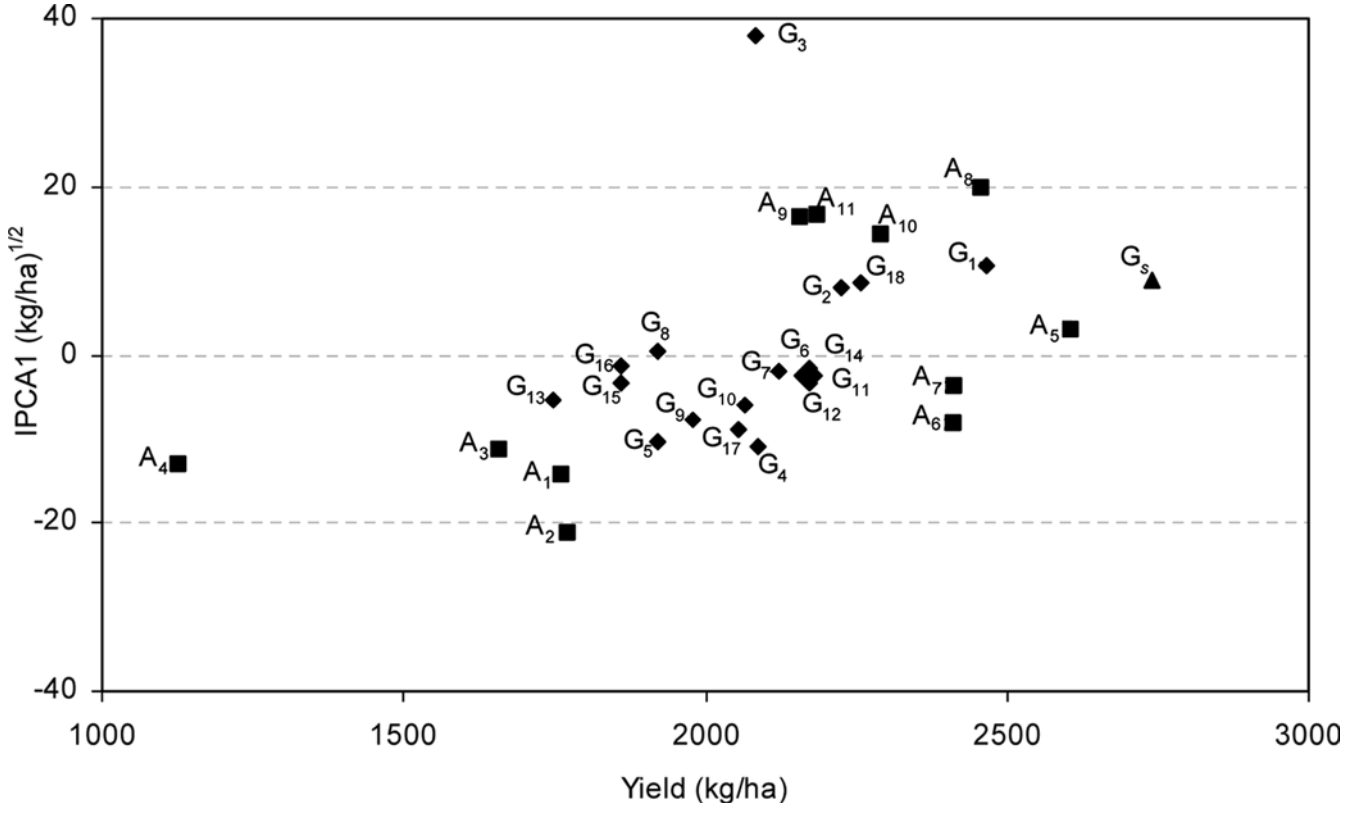

Fig. 1 Graphic display of genotypes [assessed (G_1–G_{18}) and supplementary (G_s)] and environments (A_1–A_{11}) in the AMMI1 biplot

Table 3 Mean yield (kg/ha) and scores of the 18 genotypes assessed (G_1–G_{18}), the supplementary genotype (G_s) and the 11 environments considered (A_1–A_{11}) for the two first axes of the biplot representation (IPCA1, and IPCA2)

Genotypes/environments	Mean yield (kg/ha)	IPCA1	IPCA2
G_1	2,457.73	10.68	16.02
G_2	2,258.36	8.60	−13.31
G_3	2,081.55	38.02	−2.08
G_4	2,085.45	−11.02	−0.64
G_5	1,920.91	−10.38	−4.62
G_6	2,159.55	−2.37	12.69
G_7	2,117.36	−1.75	−12.44
G_8	1,922.73	0.51	15.76
G_9	1,970.64	−7.77	−0.12
G_{10}	2,053.36	−6.06	−6.07
G_{11}	2,181.45	−2.56	−5.34
G_{12}	2,172.45	−3.45	−3.99
G_{13}	1,736.36	−5.39	12.22
G_{14}	2,173.82	−1.48	2.14
G_{15}	1,860.00	−3.44	14.51
G_{15}	1,858.09	−1.34	−0.15
G_{17}	2,053.18	−8.73	−12.96
G_{18}	2,225.73	7.93	−11.61
G_s	2,739.27	8.80	9.68
A_1	1,756.28	−13.98	−3.47
A_2	1,770.94	−21.01	−10.25
A_3	1,659.17	−11.30	7.27
A_4	1,125.28	−12.83	1.57
A_5	2,592.5	3.15	32.59
A_6	2,400.28	−7.90	−12.25
A_7	2,395.67	−3.68	2.58
A_8	2,456.78	19.95	−11.38
A_9	2,155.89	16.33	−13.56
A_{10}	2,289.39	14.44	8.14
A_{11}	2,185.39	16.82	−1.24

that we have here considered, the smaller the quadratic distance between its projection and the projection of a specific assessed genotype, for each one of the interaction axes, the greater the similarity of the G×E interaction behavior of this genotype with the behavior shown by the supplementary one. Thus, regardless of which AMMI model is chosen, the quadratic distance can be estimated from the projection of the supplementary genotype to the real genotype. If a model with more than three axes is used, even if a biplot display cannot be designed, the information on the interaction pattern associated to the genotypes tested can be obtained by the accumulated quadratic distance of the axes considered.

Table 4 presents the quadratic distances among the projections of each genotype assessed and the supplementary genotype, on the first two principal components axes. It can be observed that in IPCA1 the G_2 and G_{18} genotypes are even closer to the supplementary genotype than the G_1 genotype (cv. Conquista). Thus, for genotypic and environmental factors that mostly account for the first axis of interaction, these are the genotypes that show the most positive interactions with environments in the trial conditions. However, considering that in this case the appropriate biplot, based on an statistical test, is the AMMI2, it is shown that with respect to G×E interaction, the G_1 genotype is indeed the one that shows the smaller distance from the Lin and Binns (1988) ideal genotype. Then, the Conquista variety could be identified as stable under this point of view, that is, the genotype that most approximated the supplementary (virtual) genotype or the ideal control.

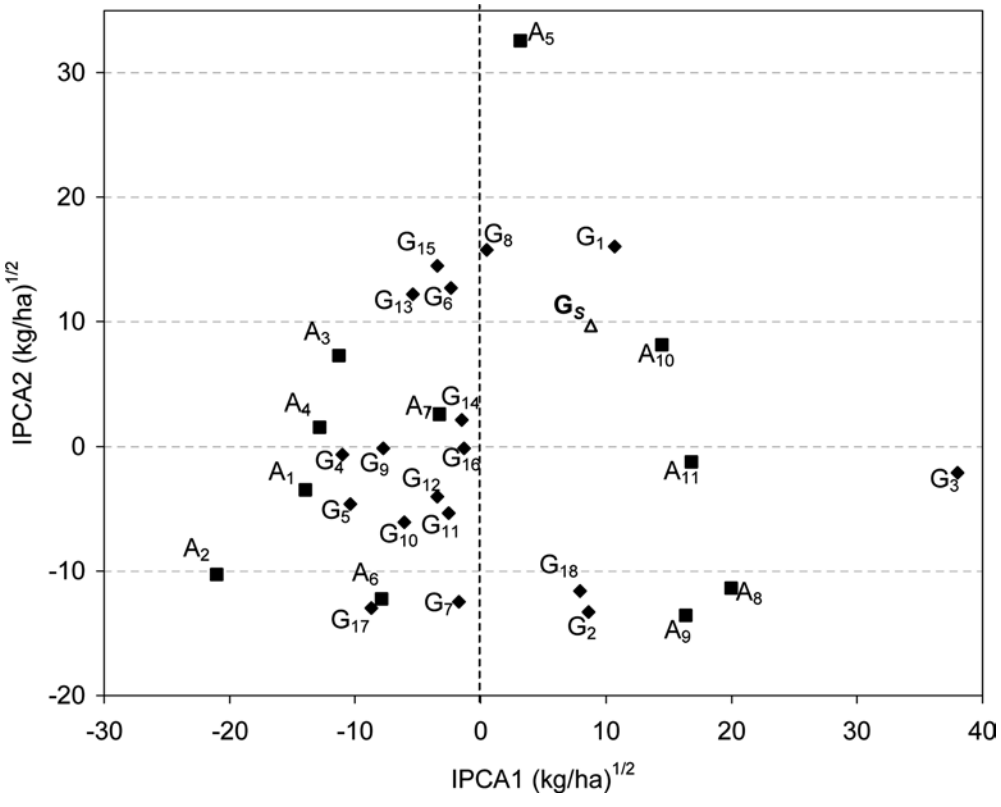

Fig. 2 Graph display of genotypes [assessed (G_1–G_{18}) and supplementary (G_s)] and environments (A_1–A_{11}) in the AMMI2 biplot

Table 4 Quadratic distances (d_{is}^2) between each assessed genotype and the supplementary genotype (G_s) when projected on the AMMI1 and AMMI2 biplot

Genotypes	AMMI1	AMMI2
G_1	3.55	43.73
G_2	0.04	528.68
G_3	854.12	992.48
G_4	392.90	499.57
G_5	367.73	572.25
G_6	124.73	133.75
G_7	111.19	600.74
G_8	68.65	105.51
G_9	274.58	370.80
G_{10}	220.93	469.00
G_{11}	128.98	354.64
G_{12}	150.11	337.11
G_{13}	201.43	207.84
G_{14}	105.76	162.73
G_{15}	149.75	173.02
G_{15}	102.75	199.47
G_{17}	307.19	819.93
G_{18}	0.76	454.31

Additionally, although it has not been the focus of this study, the score of a particular variable *j* (e.g. environment of interest) in each principal axis can be obtained similarly, whether it is an observed or supplementary variable. From the general equation derived from SVD, $H = \tilde{S}^{1/2}\tilde{V}'$, that results in$\mathrm{IPCA}_{kj} = \sqrt{\lambda_k} v_{kj}$. This value is the equivalent to the expression $\mathrm{IPCA}_{kj} = 1\sqrt{\lambda_k} u'_k (\hat{ge})_j$, where$(\hat{ge})_j$ is the vector $I\times 1$ with the estimates$(\hat{ge})_{ij}$ associated to the variable (environment) of interest. Hence, once a criterion to construct a supplementary environment of interest is defined through additional studies, it is possible also to generate scores for its representation on a biplot graph. Estimates of the distance of the real environments to this desirable virtual environment can be obtained similarly by: $d_{js}^2 = \sum_{k=1}^{n} (\mathrm{IPCA}_{kj} - \mathrm{IPCA}_{ks})^2$,where $\mathrm{IPCA}_{\mathrm{ks}}$ is the supplementary environment score in the *k*-th principal axis.

References

Crossa J (1990) Statistical analyses of multilocation trials. Adv Agron 44:55–85

Droesbeke JJ, Fine J (1995) Analisis en Componentes Principales. In: II Seminário de Capacitacion de docentes. Universidade Central de Venezuela, Université Libre de Bruxelles

Duarte JB, Vencovsky R (1999) Interação genótipos×ambientes: Uma Introdução à Análise "AMMI" Série Monografias, n 9. Ribeirão Preto, Sociedade Brasileira de Genética, 60p

Eberhart SA, Russell G (1966) Stability parameter for comparing varieties. Crop Sci 6:36–40

Eckart C, Young G (1936) The approximation of one matrix by another of lower rank. Psychometrika 1:36–40

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

Gabriel KR (1971) The biplot graphic of matrices with application to principal components analysis. Biometrika 58:453–467

Gabriel KR (1978) Least squares approximation of matrices by additive and multiplicative models. J Roy Stat Soc Ser B 40:186–196

Gauch HG (1992) Statistical analysis of regional yield trials: AMMI analysis of factorial designs. Elsevier, Amsterdam

Good IJ (1969) Some applications of the singular decomposition of a matrix. Technometrics 11:823–831

Helms TC (1993) Selection for yield and stability among oat lines. Crop Sci 33:423–426

Holland JB, Bjonstad A, Frey KJ, Gullord M, Wesenberg DM (2002) Recurrent selection for broad adaptation affects stability of oat. Euphytica 126:265–274

Lin CS, Binns MR (1988) A superiority measure of cultivar performance for cultivar×location data. Can J Plant Sci 68:193–198

Mandel J (1982) Use of the singular value decomposition in regression analysis. Am Stat 36:15–24

Nagpal PS (1999) Guide to advanced data analysis: using IDAMS software. UNESCO. Division of information and informatics. Cited 24 July 2003. Available at <http://www.unesco.org/webworld/idams/advguide>

Oliveira AB, Duarte JB, Pinheiro JB (2003) Emprego da análise AMMI na avaliação da estabilidade produtiva em soja. Pesq Agropec Bras 38:357–364

Simmonds NW (1991) Selection for local adaptation in a plant-breeding programme. Theor Appl Genet 82:363–367

Souza J (1988) Análise em componentes principais e suas aplicações: métodos estatísticos nas ciências psicossoicais. v.II. Brasilia. Thesaurus