ADAPTABILITY, STABILITY AND GENOTYPE BY ENVIRONMENT INTERACTION USING THE AMMI MODEL FOR MULTIENVIRONMENT TRIALS

Kuang Hongyu[1]

ABSTRACT:(ADAPTABILITY, STABILITY AND GENOTYPE BY ENVIRONMENT INTERACTION USING THE AMMI MODEL FOR MULTIENVIRONMENT TRIALS). The genotype by environment interaction (GEI) has an influence on the selection and recommendation of cultivars. One of the determinant factors in the success of breeding programs that aim to select genotypes for different geographical regions is understanding the interaction between genotypes and environments. The aim of this work is to study the effect of GEI and evaluate the adaptability and stability of productivity (t/ha) of nine maize genotypes using AMMI model (Additive Main effects and Multiplicative Interaction). The AMMI model is one of the most widely used statistical tools in the analysis of multiple-environment trials. The data is provided by the International Maize Improvement Center (CIMMYT) and an experiment concerning 20 genotypes evaluated in 8 countries. The AMMI model identified the best combinations of genotypes and environments with respect to the response variable. Some progenies with greater productivity and stability were identified, although stability is not associated with productivity. The genotypes G4 and G19 were the most recommended in this data set.

Keywords: Adaptability and stability, biplot analysis, multienvironment trials.

[1]Professor Doutor do Departamento de Estatística/Instituto de Ciências Exatas e da Terra. Universidade Federal de Mato Grosso, Av. Fernando Corrêa da Costa, nº 2367, Bairro Boa Esperança. CEP: 78060-900, Cuiabá, MT, Brazil. E-mail: prof.kuang@gmail.com

INTRODUCTION

Genotype by environment interaction (GEI) reflects the different responses of the genotypes to environment conditions, e.g., the best genotypes in an specific environment could not be the best for others. It depends on the considered environment. Therefore, GEI can-not represent all genetic potential environmental conditions, this makes difficult the recommendation of genotypes by the breeder (GAUCH, 2013; HONGYU et al., 2014).

Environmental conditions strongly influence agricultural production, leading to considerable variations in yield. Such influence is discriminated when yield experiments are performed in various locations and in different years (PACHECO ET AL., 2005; AKBARPOUR EL AL., 2014). Such influence is termed genotype-environment interaction (GEI). In the case of multi-environment trial (MET) data, GEI is frequently present. Due to the nature of this kind of data, it is often represented in two-way tables containing genotype means across all of the environments in the study (HONGYU et al., 2014; NEISSE; KIRCH; HONGYU, 2018).

Series environments, so-called Multienvironment trials (METs), are of variety and plant breeding trials conducted at multiple the basis for the development and dissemination of new crop varieties. Multienvironment trials are important to test and identify genotypes for high and stable yield for general and specific adaptation (MALIK; HADASCH; PIEPHO, 2018).

They are frequently analyzed by linear models with effects for genotype, environment, and their interaction. The pattern of GEI in MET data is usually of particular interest because it determines the adaptive pattern of a genotype's response to changing environments. Many models have been suggested for GEI effects in MET data (YAN; KANG, 2003; HONGYU et al., 2014; MALIK; HADASCH; PIEPHO, 2018).

Gauch (2006) suggests that only one of these methodologies should be used and emphasizes the superiority of AMMI analyses in terms of visualization. Using this method, we obtain two biplot graphs. AMMI1 demonstrates the additive effects on the X axis (genotype means and environments) and the first multiplicative interaction axis, called the

Interaction Principal Component Analysis (IPCA), on the Y axis. Genotypes with scores similar to zero are stable. Genotypes to the right have higher productivity than the general mean of the measured trait (PUPIN et al., 2018).

Grouped genotypes show similar adaptation and groups located near to an environment are similarly influenced by that environment. The second graph, AMMI2, plots the IPCA1 and IPCA2. Genotypes and environments located away from the source contribute to interaction and those with IPCA1 and IPCA2 scores close to zero are stable. When the points occupy the same quadrant in the graph the genotype and environment interact positively and when they are located on opposite quadrants, they interact negatively (GAUCH, 2006; PUPIN et al., 2018).

The objective of this work is to study the effect of GEI. By using the AMMI model, it assesses the adaptability and stability of productivity of nine maize genotypes in twenty environments.

MATERIAL AND METHODS

The data used in the analysis has been provided by the International Maize Improvement Center (CIMMYT) since 2012 in their online Research Data Repository. Various traits were measured from 7 different genotypes in 8 environments across 5 countries. The experiment had an RCB design with 2 replications, but only the averages over replications are available from the publication, as well as the mean square error.

The environments are drought-prone and typically receive less than 500 mm of rainfall during the cropping cycle. Table 1 and 2 presents further information on the number of countries, genotypes, environments and the number of locations in each.

Table 1 – Relation of environments and locations.

Environments	Countries	Locations
E1	Costa Rica	Santa Lucia
E2	Colombia	La Catalina Pereira
E3	Bolivia	San Julian
E4	Bolivia	Muyupampa
E5	Ghana	Ejura
E6	Ghana	Kwadaso
E7	Myanmar	Kim PonTaung
E8	Myanmar	Tatkone

Table 2 – Maize genotype codes and their pedigrees.

Genotypes	Pedigree
G1	(CLQRCYQ59)/(CLQS89YQ04)
G2	(CLQRCYQ64)/(CLQRCYQ65)
G3	(CLQRCYQ59)/(CML161)
G4	(CLQRCYQ59)/(CLQRCYQ49)
G5	(CML161)/(CML165)
G6	(CL02720/CLRCY017)
G7	(CML451/CL02450)

All of the analyses in this study concern only the yield weight, which was the only measured trait with replicates. Replication is a key characteristic for the AMMI model, making it possible to perform analysis of variance (ANOVA). The yield weight (t/ha) was primarily analyzed with simple ANOVA and conjoint analysis to assess the genotypic and environmental main effects as well as the GEI effects. Once GEI was evaluated as a significant effect present in the data, adaptability and phenotypic stability analyses were performed using the AMMI and GGE models. All of the analyses presented in this study were performed using R statistical software, version 3.5.1 (R DEVELOPMENT CORE TEAM, 2018).

Analysis of variance

There are three numbers from the ANOVA which provide a preliminary indication as to whether AMMI analysis will be worthwhile. These are the sum of squares (SS) for genotypes (G), GEI signal (GES), and GEI noise (GEN). The SS values for G and GEI are direct outputs from ANOVA (GAUCH, 2013).

To estimate the SS for GEN, simply multiply the error mean square (from replication) by the number of degrees of freedom (*DF*) for GEI (GAUCH, 1992). Then obtain GES by subtracting GEN from GEI. AMMI analysis is appropriate for datasets that have substantial G and substantial GES. Especially when the SS for GES is at least as large as that for G, as happens frequently, AMMI analysis will probably be worthwhile (GAUCH, 2013). On the other hand, occasionally GEI is buried in noise, with the SS for GEN approximately equal to that for GEI. In that case, GEI should be ignored, so AMMI analysis is inappropriate (GAUCH, 2013; HONGYU et al., 2014).

AMMI Model

AMMI is a model family rather than a single model. Consequently, model diagnosis is required to determine which member of this model family is the best for a given dataset and research purpose (GAUCH, 2013; HONGYU et al., 2014). The AMMI model combines two methods: analysis of variance and singular value decomposition in a unique model, additive components for the main effects of genotypes (g_i), environments (e_j) and multiplicative components for the interaction effect (*ge*)$_{ij}$ (DIAS; KRZANOWSKI, 2003; ARCINIEGAS-ALARCN et al., 2011; GAUCH et al.,2011; GAUCH, 2013; HONGYU et al., 2014). Therefore, the model equation for the *i*-th genotype in the *j*-th environment in *r* blocks (repetitions) is (GAUCH, 1992):

$$Y_{ijr} = \mu + g_i + e_j + \sum_{k=1}^{n} \lambda_k \gamma_{ik} \alpha_{jk} + \rho_{ij} + b_{(r)j} + \varepsilon_{ijr} \tag{1}$$

where Y_{ijr} is the phenotypic trait (e.g. yield) of genotype *i* in environment *j* for replicate *r*, μ is the grand mean, g_i are the genotype main effects as deviations from μ, e_j are the environment main effects as deviations from μ, λ_k is the singular value for the Interaction Principal Component (IPC) axis *k*, α_{jk} and γ_{ik} are the genotype and environment IPC cores (i.e. the left and right singular vectors) for axis *k*, $b_{(r)j}$ is the effect of the block *r* within the environment *j*, *r* is the number of blocks, ρ_{ij} is the residual containing all multiplicative terms not included in the model (1); *n* is the number of axes or principal components (PC) retained by the model, and ε_{ijr} is the experimental error, assumed independent with identical distribution, $\varepsilon_{ijr} \sim N(0, \sigma^2/r)$.

Successive IPCs are denoted by IPC1, IPC2, IPC3, and so on. The number of these components is one less than the minimum of the number of genotypes and number of environments. The member of the AMMI model family retaining zero components is denoted by AMMI0, and the successive members which retain one or more components are denoted by AMMI1, AMMI2, AMMI3, and so on, up to the full model retaining all components denoted by AMMIF (GAUCH, 1992; GAUCH, 2013). When the fitted values of the full model automatically equal the raw data Y_{ijr} exactly, the residual term disappears. However, the reduced models leave a residual ε_{ijr} (GAUCH, 2013; HONGYU et al., 2014).

In the first phase of the main effects in the additive part (mean, genotypes and environments), they are fitted by a traditional AOV to the matrix of means $[\mathbf{Y}(g \times e)]$ in the r blocks. The result is the nonadditivity residual which means the GEI, $(ge)ij$. This interaction is the multiplicative part of the model. In the second phase, the interaction is analysed by singular value decomposition (SVD) of the matrix $\{\mathbf{GE}(g \times e) = [(ge)ij]\}$ or by principal component analysis (PCA) (GAUCH, 1992; DIAS AND KRZANOWSKI, 2006; GAUCH, 2013; HONGYU ET AL., 2014).

The **GE** matrix is the interaction matrix between genotypes and environments (residual of main effects), where each element $(ge)ij$ of **GE** is given by (GAUCH, 1992)

$$(ge)_{ij} = Y_{ij} - \overline{Y}_{i.} - \overline{Y}_{.j} + \overline{Y}_{..}$$

where Y_{ij} is the mean of the repetitions of genotype i in environment j, with $i = 1, 2, \ldots, g$ and $j = 1, 2, \ldots, e$; $\overline{Y}_{i.}$ is the mean of genotype i, $\overline{Y}_{.j}$ is the mean of environment j and $\overline{Y}_{..}$ is the global mean.

The results were analysed by graphical representation in biplot, a technique which is useful in principal component analysis. The graphic is used to represent simultaneously the rows and columns of a data matrix. It indicates the existing groups in the observations and in this way shows dispersions and correlations between the variables (GAUCH, 1992).

RESULTS AND DISCUSSION

The results of the joint analysis of variance for productivity of maize (t/ha) show that the genotype, environment and interaction effects are significant ($p < 0.01$). That indicates that the genotypes present different behavior in the environments. It justifies a study of the behavior of the genotypes in order to identify the magnitude of interaction with the environments (GAUCH, 1992; HONGYU et al., 2014).

Table 3 presents the ANOVA for AMMI 6. The sources listed in this table correspond to the terms in Eq. (1) for the AMMI model applied to a yield trial with an RCB experimental design. In the AMMI analysis the square sum of the interaction ($SS_{GEI} = 146.99$) corresponds to the eigenvalues ($\sum_{k=1}^{n} \lambda_k^2$) (Table 3). It can be inflated by the presence of noise (inexplicable variation) in the response variable. It is necessary to make an adjustment of the interaction by singular value decomposition of the GEI matrix. The SS for G is 191.01 and for the irrelevant E it is 941.89 (Table 3). This matrix has rank $p = 6$, therefore it can be decomposed into six principal components that correspond to the partial square sum in the analysis of variance.

Model diagnosis for AMMI is routinely done using cross-validation of experiments with replication. However, this is not an option for this wheat trial because the replicated data were not published (GAUCH, 2013). Observing the decomposition of the genotype-by-environment interaction through the AMMI model (Table 3), the first and second principal components (PCs) are significant ($p < 0.05$) by Gollob (1968) F test and explain 52% and 28.2% respectively of the variation of the SS_{GEI}. These three principal components sum to 80.2% of the SS_{GEI}, which is considered a pattern response present in the SS_{GEI} with 22 degrees of freedom (DF).

Table 3 – Conjoint analysis of variance of the productivity of maize trials and decomposition of the sum of squares of (GEI)

Source	DF	SS	MS	p-value
Environment (E)	7	941.89	134.55	< 0.01
Block/ Environment	16	0.00	0.00	> 0.05
Genotype (G)	6	191.01	31.83	< 0.01
GEI	42	146.99	3.50	< 0.01
PC1	12	76.46	6.37	< 0.01
PC2	10	41.52	4.15	0.03
PC3	8	21.01	2.63	0.24
PC4	6	4.61	0.77	0.88
PC5	4	3.10	0.78	0.81
PC6	2	0.30	0.15	0.92
Residual	96	189.89	1.98	-
Total		1469.8	-	-
General mean	4.08			
CV(%)	34.46			

The last stage of the AMMI analysis is the graphical representation of genotypes and environments in the biplot (GABRIEL, 1971) and identification of mega-environment. It is necessary to determine the position in the interaction of singular axes. From the matrix **U**, **S** and **V** resulting from the singular value decomposition (SVD) of the **GE** matrix, we obtain the results of interest (GAUCH, 2013; HONGYU et al., 2014).

In this case, it is possible to build up the biplot. The biplots are: i) AMMI1 - Means vs PC1 (Figure 1); ii) AMMI2 - PC1 vs PC2 (Figure 2). The biplot graphics are used to analyse the dispersion of genotypes, environments and the interaction between them. The AMMI1 biplot contains the variation of the principal additive effects of genotypes and environments. This is shown in the horizontal axis, while the variation of the multiplicative effects of the GEI is shown in the vertical axis. In the biplot, AMMI2 is visualized by the multiplicative effects of the GEI contained in the first two PCs.
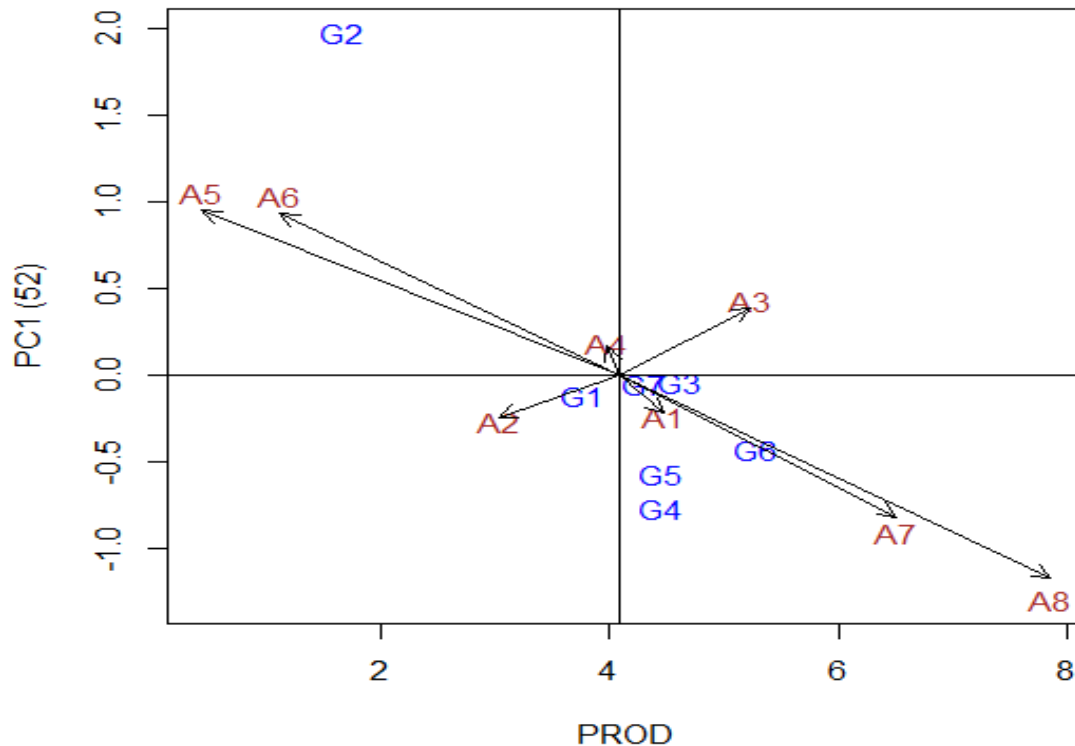
Figure 1. Biplot AMMI1 (Means vs PC1) for the data of productivity of maize (t/ha) with seven genotypes (G) and eight environments (E)

According to the values of the two first principal components (CP1 and CP2) or by Fig. 1, G6, G3 G5 and G4 are the genotypes with best answers and more productive in the environmental conditions prevailing during crop development. The genotypes are more stable in G1, G7 and G3 (Figure 1 and 2). These can be grown in all the locations where the study was carried out. Among them, the genotype G3 display productivity above the general mean and are between the two groups (productive and stable). They indicate that this genotype is associated with adaptability and stability. However, the genotypes with high mean productivity can-not be stable. The case of the genotype G6 shows a specific adaptability to the environment A7 and A8; the genotype G2 to the environment A5 and A6; G7 to the environment A4 (Figure 1 and 2).
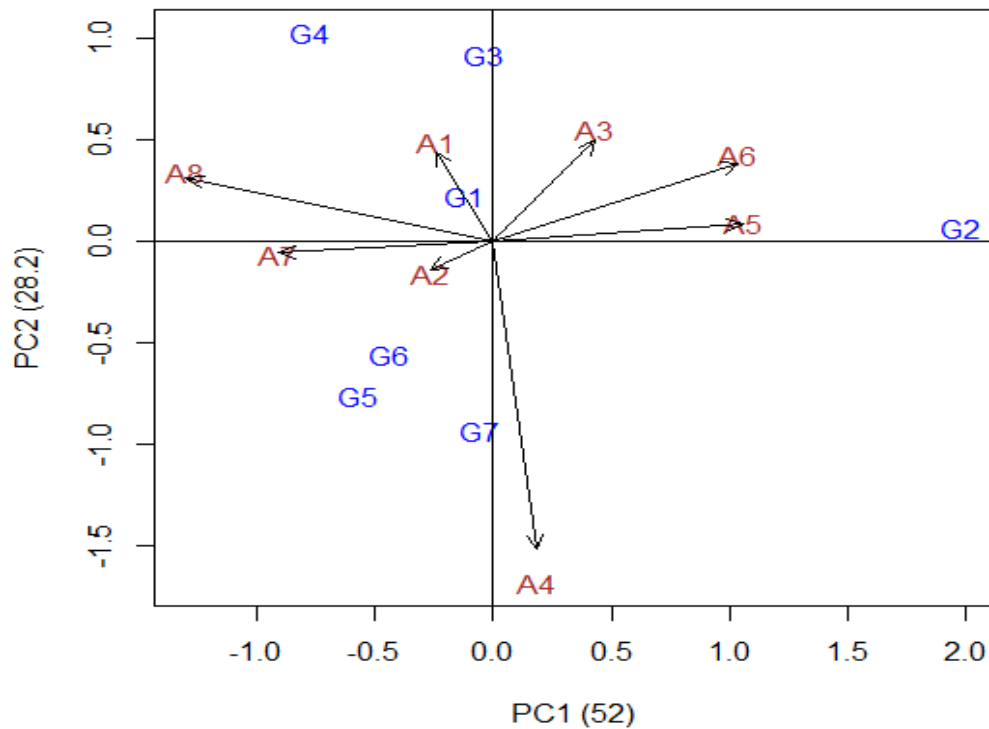
Figure 2. AMMI2 biplot (PC1 vs PC2) for data on the productivity of maize (t/ha) with seven genotypes (G) and eight environments (E)

As regards the environments, poor contributions were presented by the environments E5, E6 and E7. The environments more discrepant were E4, E5, E6, E7 and E8. E8 and E7 give the highest mean of productivity and E5 and E6 the lowest mean (Figure 1). These can be considered examples of favorable and unfavorable environments, respectively (Figure 1 and 2). The environments E5 and E6 are very similar, like the environments E7 and E8 as well. The environments E5, E6 are very different in relation to the characteristics of E7 and E8. This information can be observed in Figure 1. It is possible to replace one of these locals by other more representative of the region, where the cultivars will be recommended. The AMMI analysis can be used efficiently in the identification of superior environmental conditions for agricultural exploitation (selection of growing locations) and superior mean performance genotypes (YAN, 2010; GAUCH, 2013).

CONCLUSIONS

This study revealed that genotypes, environments and GEI were significant for grain yield. The genotypes therefore performed differently with respect to yield in each of the eleven test environments and their relative performance varied from one environment to another. AMMI analysis showed that environment effects accounted for a larger proportion of the total variation in the sum of squares for grain yield than genotype effects and $G \times E$ effects.

Among SVD-based statistical analyses, AMMI is a unique analysis that completely and always separates G, E, and GEI as required for most agricultural research purposes. Furthermore, it separates signal from noise as well as any other method for the purpose of gaining accuracy. Most stability parameters were consistent with AMMI parameters in detecting the stable and non-stable genotypes with some exceptions based on the concept of stability for each of the stability parameters in the current study.

REFERENCES

ARCINIEGAS-ALARCN, S.; GARCIA-PEÑA, M.; DIAS, C.T.S.; KRZANOWSKI, W. J. An alternative methodology for imputing missing data in trials with genotype by-environment interaction. **Biometrical Letters**. 47: 1–14. 2010.

DIAS, C.T.S.; KRZANOWSKI, W.J. Choosing components in the additive main effect and multiplicative interaction (AMMI) models. **Scientia Agricola**. 63: 169–175. 2006.

DIAS, C.T.S.; KRZANOWSKI, W.J. Model selection and cross validation in additive main effect and multiplicative interaction models. **Crop Science**. 43: 865–873. 2003.

HONGYU, K.; GARCIA-PEÑA, M.; ARAÚJO, L. B.; DIAS, C. T. S. **Statistical analysis of yield trials by AMMI analysis of genotype x environment interaction.** Biometrical Letters, Poznan, v.51, p.89-102, 2014.

GABRIEL, K.R. The biplot graphic display of matrices with application to principal component analysis. **Biometrika**. v. 58, p. 453-467, 1971.

GAUCH, H.G. **Statistical analysis of regional yield trials: AMMI analysis of factorial designs**. Elsevier, Amsterdam. 1992.

GAUCH, H. G. (Ed.) **Genotype-by-environment interaction**. New York: CRC Press, p.85-122. 1996.

GAUCH, H. G. **Statistical Analysis of Yield Trials by AMMI and GGE**, Maio 18, 2006.

GAUCH H.G. A Simple Protocol for AMMI Analysis of Yield Trials. **Crop Science**: (in press). 2013.

GAUCH H.G., RODRIGUES P.C., MUNKVOLD J.D., HEFFNER E.L., SORRELLS M. Two New Strategies for Detecting and Understanding QTL × Environment Interactions. In **Crop Science**. 51: 96–113. 2011.

GOLLOB, H.F. A statistical model which combines feature of factor analytic and analysis of variance techniques. **Psychometrika** 33: 73–115. 1968.

MALIK, W.A. HADASCH, J.F. PIEPHO, H.P. Nonparametric Resampling Methods for Testing Multiplicative Terms in AMMI and GGE Models for Multienvironment Trials. **crop science**, vol. 58, march–april. 2018.

NEISSE, A.C.; KIRCH, J.L.; HONGYU, K. AMMI and GGE Biplot for genotype × environment interaction: a medoid–based hierarchical cluster analysis approach for high–dimensional data. **Biometrical Letters**. 2018.

PACHECO, R.M.; DUARTE, J.B.; VENCOVSKY, R.; PINHEIRO, J.B.; OLIVEIRA, A.B. Use of supplementary genotypes in AMMI analysis. **Theor Appl Genet** 110: 812-818. 2005.

PUPIN, S.; SILVA, P.H.M.; PIOTTO, F.A.; MIRANDA, A.C.; ZARUMA, D.U.G.; SEBBENN, A.M.; MORAES, M.L.T. Genotype x Environment interaction, stability, and adaptability in progenies of Eucalyptus urophylla S.T. BLAKE using the AMMI model. **Genetica**. 67, 51 − 56. 2018.

R DEVELOPMENT CORE TEAM. R: a language and environment for statistical computing. **R Foundation for Statistical Computing**, Vienna, Austria, 2018. URL https://www.R-project.org/.

YAN, W. Optimal Use of Biplots in Analysis of Multi-Location Variety Test Data. **Acta Agronomica Sinica**, 36 (11): 1805–1819. 2010.

YAN, W.; KANG, M.S. **GGE biplot analysis: a graphical tool for breeders, geneticists, and agronomists**. Flórida: Boca Raton, 2003. 286p.