UDC (UDK) 502/504:575.22

Naser SABAGHNIA [1]

MULTIVARIATE STATISTICAL ANALYSIS OF GENOTYPE × ENVIRONMENT INTERACTION IN MULTI-ENVIRONMENT TRIALS OF BREEDING PROGRAMS

SUMMARY

In final stages of plant breeding programs, a large number of new improved genotypes are tested over a wide range of test environments and the underlying statistics used to model this system may be rather complicated. Usually, the presence of the genotype × environment (GE) interaction effect complicates the selection of the most favorable genotypes for a target test environment. There are several statistical methods available to analyze results of multi-environment trials including a range of univariate and multivariate procedures. Univariate methods have inadequate capacity to fully explain the GE interaction structure because they attempt to define the GE interaction by one or two parameters but the multiplicative GE interaction is far too complex to be summarized by only some limited parameters. In contrast, multivariate statistical methods explore multi-directionality aspects of the GE interaction and try to extract more information. The most common multivariate statistical methods are cluster analysis (CA), principal components analysis (PCA), principal coordinates analysis (PCOA), factor analysis (FA), the additive main effect and multiplicative interaction (AMMI), shifted multiplicative model (SHMM), site regression biplot (GGE). This paper reviews these multivariate statistical methods for analyzing a multi-environment trial dataset. Several AMMI stability parameters were discussed and three of these important models (AMMI, GGE and SHMM) are compared.

Keywords: adaptation, biplot, stability analysis, yield

INTRODUCTION

Plant adaptation is the process by which genotypes become more suited to thrive in a given test environment and the term refers to the association between a plant and its environment It can therefore be used to explain process and condition. The common breeding target is to develop genotypes with high yield and stable performance over a range of production environments (Allard and Bradshaw, 1964). Stability is yield variability over environment and genotype adaptability is a term used to describe yield variability across locations averaged over years. Plant breeders are concerned with both stability and adaptability when making selections from breeding lines and they should be closely related if

[1] Naser SABAGHNIA, Ph.D (e-mail: sabaghnia@yahoo.com), Department of Agronomy and Plant Breeding, Faculty of Agriculture, University of Maragheh, Maragheh, Iran.

the genotype × environment (GE) interaction is caused by unpredictable environmental variables (Annicchiarico, 1997). In choosing genotypes, a breeder would mostly be interested in their relative stability at a specific farm location. Where genetic differences in performance correspond to factors related with particular locations, they can be exploited by the development of regional breeding or selection programs if sufficient resources exist (Kang, 2002). The major objective of all crop-breeding programs is to develop pest and disease resistant genotypes as these genetically resistant genotypes have many benefits.

The phenotype of a plant is the result of its genotype and the environment in which it develops but these effects may not be independently identifiable, hence consideration of the GE interaction in plant breeding. Some genotypes perform well in a wide range of test environments, but others require specific environmental conditions to show their genetic potential (Crossa, 1990). Most agronomists are concerned with the production of particular genotypes over time and place. Genetic improvement of crops involves modification of a genotype to produce a more appropriate expression for a particular environment but that may change over time, either in the short or longer term (Gauch, 2006). In practice, breeding for crop productivity and adaptation depends on the manipulation of both genetics and environment. Therefore, tasks for genetic adaptation are only one aspect of plant breeding, and may not be the most appropriate means to resolve the primary limit to productivity or adaptation (Yan et al., 2000).

Several statistical procedures have been developed to describe the GE interaction and facilitate genotype recommendations in breeding programs (Ferreira et al., 2006). Common methods are broadly categorized in terms of parametric (univariate and multivariate) or nonparametric strategies. The univariate parametric strategy includes variance components-based methods (Wricke, 1962; Shukla, 1972) and joint linear regression methods (Eberhart and Russell, 1966; Hernandez et al., 1993), but the nonparametric strategy includes the rank values of genotypes (Huehn, 1979), and the multivariate strategy incorporates several statistical methods (Williams, 1952; Gauch, 1988). Multivariate methods have some advantages including deletion of noise from the data pattern, summarizing the dataset, and revelation of data structure (Crossa, 1990). In contrast with conventional statistical strategies, the function of multivariate analysis is to elucidate the internal structure of data from which hypotheses can be produced and tested by statistical procedures (Gauch, 1996).

Multivariate statistical methods are appropriate for analyzing two-way layouts of genotypes and environments in multi-environment trials. The response of a special genotype in various test environments may be conceived as a pattern in multi-dimensional space, with the coordinates of an individual axis being that of yield or another trait. Cluster analysis (Abou-El-Fittouh et al., 1969), principal components analysis (Freeman and Dowker, 1973), principal coordinates analysis (Mungomery et al., 1974), factor analysis (Peterson and Pfeiffer, 1989), the additive main effect and multiplicative interaction (Zobel et al., 1988), shifted multiplicative model (Cornelius et al., 1992), site regression biplot (Yan et al.,

2000) are the most common multivariate statistical methods used for investigation of the GE interaction and yield stability analyses. Many studies have used multivariate stability statistics to analyze the GE interaction in agricultural trials. There is increasing global interest in using these statistics by plant breeders due to potential high returns relative to stability parameters. This review combines theoretical considerations and empirical studies to provide a comprehensive perspective. This discussion should enhance plant breeders' understanding of multivariate analysis of the GE interaction.

CLUSTER ANALYSIS

Cluster analysis based on differences in genotypes' responses across test environments is the most commonly used multivariate method. There are two major types of the multivariate method that have been used to extract patterns of the GE interaction, classification and ordination techniques. Abou-El-Fittouh et al. (1969) proposed cluster analysis as a technique to classify test environments for cotton. Cluster analysis involves grouping similar entities in clusters and is effective for summarizing redundancy in data. A number of studies have been done to classify test environments or genotypes using cluster analysis in wheat (Fox and Rosielle, 1982), barley (van Oosterom et al., 1993) and soybean (Hanson, 1994). Identifying those genotypes with similar responses to environmental changes but different from genotypes in other groups can be intellectually satisfying, profitable, or sometimes both. The cluster analysis does not detect a particular statistical method but it often doesn't need to make any assumptions about data distribution. ANOVA and joint linear regression models are used for analyzing two-way data but they do not identify the level which is responsible. To meet these targets, several cluster methods have been suggested, some of which classify individuals for similarity according to the one-way method (Edwards and Cavalli-Sforza, 1964; Callinski and Corsten, 1985); and others classify individuals for similarity of interactions based on the two-way method (Lin and Thompson, 1975; Lin and Butler, 1990).

There are two major procedures for grouping genotypes according to their response to environmental changes; the first was proposed by Abou-El-Fittouh et al. (1969) in which genotype is a vector of n attributes indicated by m environments using the distance coefficient. Similar to this method, Mungomery et al. (1974) has used squared distance as a similarity index for clustering. In the second method, Lin and Thompson (1975) used the deviation MS from the linear regression model of the GE interaction (Finlay and Wilkinson, 1963) as a dissimilarity index for clustering. As an alternative procedure in the first method, Lin (1982) used the GE interaction mean square as a dissimilarity index for genotype classification through a slight adjustment of the distance coefficient of Abou-El-Fittouh et al. (1969) procedure. The dissimilarity index of Lin and Thompson (1975) benefits both genotype and GE interaction effects and Lin and Butler (1990) introduced a new dissimilarity index according to regression analysis that benefits only genotype as the main effect. Also, Lin and Butler

(1990) suggested a new dissimilarity index based on the mean square of only the GE interaction in contrast to the dissimilarity index of Lin and Thompson (1975) that uses both effects of genotype and GE interaction effects in ANOVA.

An important aspect of cluster analysis is having a well-defined stopping criterion or cutoff point. It is very important to determine the right cut-off point to decrease the risk of Error type II. A cutoff point can be determined if the dissimilarity index has some relationship with the deviation mean square from a regression model or the GE interaction MS in ANOVA (Lin, 1982). For dissimilarity indices of Lin and Thompson (1975), Lin (1982) and Lin and Butler (1990,) some F-tests for stopping the clustering procedure are defined. Lin and Butler (1990) present a detailed illustration of clustering and computation of the dissimilarity index. Formulas for the dissimilarity indexes in each method and their degrees of freedom are given in Table 1.

Table 1: The four possible methods for cluster analysis based on regression and ANOVA models

Method	Strategy	Source†	Distance measure	v_1‡	v_2¶	Reference
1	Regression	G & GE	$d_{(1,2,\ldots,r)} = [SSD_{(1,2,\ldots,r)} - \sum_{i=1}^{r} SSD_i]/[2(r-1)]$	$2(r-1)$	$(m-1)(n-2)$	Lin and Thompson (1975)
2	ANOVA	GE	$d_{(1,2,\ldots,r)} = SS(GE)_i/[(n-1)(r-1)]$	$(n-1)(r-1)$	$m(rep-1)(n-2)$	Lin (1982)
3	ANOVA	G & GE	$d_{(1,2,\ldots,r)} = [SS(GE)_i + SS(G)_i]/[n(r-1)]$	$n(r-1)$	$m(rep-1)(n-2)$	Lin and Butler (1990)
4	Regression	GE	$d_{(1,2,\ldots,r)} = [\sum_{i=1}^{r} SSR_i - SSR_{(1,2,\ldots,r)}]/(r-1)$	$(r-1)$	$(m-1)(n-2)$	Lin and Butler (1990)

†Grouping according to similarity of which sources
‡ Degrees of freedom for fraction of F-test
¶ Degrees of freedom for denominator of F-test

For formulas 1 and 4; SSR_i, SSD_j indicate the sums of squares (SS) due to the regression and the SS of deviation from the regression for genotype *i*. Also, $SSR_{(1,2,\ldots,r)}$, $SSD_{(1,2,\ldots,r)}$ show the corresponding SS from the linear regression for genotypes 1, 2, . . . , r and $r \leq m$. For formulas 2 and 3; SSG_i, $SSGE_j$ indicate the sums of squares (SS) due to genotype and the SS of the GE interaction for genotype *i*. Also, *m* is the number of genotypes, *n* is the number of environments, *r* is the number of genotypes in a newly formed cluster and *rep* is the number of experiment replications. A FORTRAN-77 program, known as Sl16 (Lin et al. 1992) is available for different methods of cluster analysis.

There are various other ways of scaling and standardizing data including environment-centered, environment-standardized, environment heritability-weighted and environment-ranked methods as well as these mentioned clustering procedures (Delacy et al., 1996). There are also numerous clustering methods that can be considered as inadequate as sometimes give different results when given the same dataset. The basic similarity of all clustering methods is that they use some similarity or distance measurements to classify items into groups. These measurements used for clustering either genotype or test environment are given in Table 2.

Table 2. Similarity and distance measures between two special genotypes

Similarity measure¶		Origin
Uni-criterion approach	Euclidean distance	
	$d^2(A)_{ii'} = \sum_{j=1}^{q} (X_{ij} - X_{i'j})^2$	Hanson, 1970
	$d^2(B)_{ii'} = \sum_{j=1}^{q} \left[(X_{ij} - \bar{X}_{i.}) - (X_{i'j} - \bar{X}_{i'.}) \right]^2 / q$	Abou-El-Fittouh et al. 1969
	Standardized distances	
	$d_s^2(A)_{ii'} = \sum_{j=1}^{q} \left(\frac{X_{ij} - \bar{X}_{i.}}{S_i} - \frac{X_{i'j} - \bar{X}_{i'}}{S_{i'}} \right)^2$	Fox and Rosielle, 1982
	$d_s^2(B)_{ii'} = \sum_{j=1}^{q} \left(\frac{X_{ij} - \bar{X}_{i.} - \bar{X}_{.j} + \bar{X}_{..}}{W_i} - \frac{X_{i'j} - \bar{X}_{i'} - \bar{X}_{.j} + \bar{X}_{..}}{W_{i'}} \right)^2$	Mungomery et al. 1974
	$d^2(B)_{ii'} = n \sum_{i \neq i} (S_i - S_{i'})^2 / 2g$	Muir et al. 1992
	$d^2(A)_{ii'} = S(X_{ij}, X_{i'j})$ if $i \neq i'$ and $d^2(A)_{ii'} = 0$ if $i = i'$	Abou-El-Fittouh et al. 1969
Correlation coefficient	$r(A)_{ii'} = \frac{\sum_j (X_{ij} - \bar{X}_{i.})(X_{ij} - \bar{X}_{i'.})}{\{\sum_j (X_{ij} - \bar{X}_{i.})^2 \sum_j (X_{i'j} - \bar{X}_{i'.})^2\}^{0.5}}$	Guitard, 1960
	$r(A)_{ii'}^{\varnothing} = \frac{\sum_j \delta_{ij} \delta_{i'j}}{\sum_j (\delta_{ij}^2 \sum_j \delta_{i'j}^2)^{0.5}}$	Perkins and Jinks, 1968b
	$SS(IC)_i = n \sum_{i \neq i} (1 - r_{ii'})(S_i S_{i'}) / g$	Muir et al. 1992
	$r(B)ii' = \frac{\sum_j (X_{ij} - \bar{X}_{i.} - \bar{X}_{.j} + \bar{X}_{..})(X_{i'j} - \bar{X}_{i'.} - \bar{X}_{.j}}{\{\sum_j (X_{ij} - \bar{X}_{i.} - \bar{X}_{.j} + \bar{X}_{..})^2 \sum_j (X_{i'j} - \bar{X}_{i'.} - \bar{X}_{.}}$	Hhbgood, 1977

¶ summarized from Lin et al. (1986)

PRINCIPAL COMPONENTS ANALYSIS

Principal components analysis (PCA) is a multivariate statistical method to identify data patterns as well as similarities and dissimilarities among variables based on ordination techniques of multivariate methods. According to Jolliffe (2002), the initial explanation of the PCA technique was given by Pearson (1901

cited in Jolliffe, 2002) and Hotelling (1933 cited in Jolliffe, 2002). Gower (1966) discussed links between PCA and various other statistical methods and provided a number of important geometric insights. The PCA method has been applied in a wide range of areas such as agriculture and genetics; it would be easy to add more to these fields. It is used to find optimal ways of combining variables into a small number of subsets, and the main applications of this method can be determined by analysis of multiple indicators, the measurement of complex constructs, scale construction and data reduction. In other words, the PCA procedure is appropriate when obtained measures on a number of variables are changed to a smaller number of artificial variables or principal components. The PCA method may then be used to establish predictors or criterion variables in subsequent analyses.

Freeman and Dowker (1973) used PCA to interpret causes behind the GE interaction as a good tool, but Perkins (1972) found that PCA was not useful for studying adaptation. Hirosaki et al., (1975 cited in Crossa, 1990) reported that PCA was more efficient than the linear regression model in explicating performance of genotypes and Polignano et al., (1989) combined PCA with cluster analysis as an effective way of forming subgroups of faba bean. Under some conditions, the PCA is a generalization of the joint linear regression model (Williams, 1952). Mandel (1971) analyzed a two-way layout by applying the AMMI model; ANOVA for main effects and PCA for the interaction between main effects. Kempton (1984) used AMMI model for summarizing the pattern of genotype responses across environments. The display of genotypes and environments along the first two PCA axes for the interaction is known as a biplot (Gabriel, 1971). The PCA as an ordination technique that may have limitations such as reeducation dimensionality of data, distortions may sometimes occur, and if the magnitude of variance accounted for by the first two PCA axes is small, genotypes or environments that are far apart may be indicated by points that are close together (Gower, 1967). Furthermore, low correlation among variables prevents the occurrence of only a few dimensions from accounting for most of the variation and components may do not have any clear association with environmental factors. Finally, contrary to ANOVA, PCA assumes a complete multiplicative model without any explanation of the genotype and environment effect (Zobel et al. 1988).

The joint linear regression model uses only two stability statistics, the regression coefficient and deviation MS, to explain the structure of response of a genotype across environments therefore most of the information is wasted. The PCA can overcome this difficulty by giving more statistics, the scores on the PCA axes, to describe the response pattern of a genotype (Eisemann et al., 1990). The PCA confounds the additive (main effects) pattern of data with the non-additive (GE interaction) and nonlinear relationship in the data prevents the PCA from efficiently describing the real relationships between variables. However, in recent decades it seems that the PCA, as a proper statistical procedure has not

been used individually but is mostly used in an AMMI model layout in combination with ANOVA.

PRINCIPAL COORDINATES ANALYSIS

Principal coordinates analysis (PCOA) is a generalization of PCA and measures similarities between genotypes. It is a method used to explore and to visualize similarities or dissimilarities of a dataset. This method assumes that the original variables had a Euclidean space and that their similarities are modeled by Euclidean distance (Gower, 1966; Westcott, 1987). The PCOA starts with a similarity or distance matrix and assigns for each individual, a position in a low-dimensional space. The main target of the PCOA is to transform the data from one series of coordinate axes to another. Like PCA, this analysis maintains most of the initial configuration of the dataset in the first axes so some original information is lost. The PCOA can effectively reduce the pattern of a two-way dataset of multi-environment trials' dimensions in a subspace of fewer dimensions (Ibanmez et al., 2001). The mentioned two-way structure can also be conceptualized as environment points in genotype dimensions. The limitations of the PCOA as an ordination approach of multivariate methods are similar to PCA. Furthermore, the nonlinear relationships prevent effective explanation of real relationships between genotypes (Gower, 1971).

The PCOA (Westcott, 1987) was used for yield stability analysis by some authors (Crossa et al., 1989; Flores et al., 1996; Ibanmez et al., 2001). A measure of similarity between two genotypes, *m* and *n*, in a given test environment is:

$$S_{i(m,n)} = [H_i - (m_i + n_n)/2]/(H_i - L_i)$$

where *Hi* is the highest mean yield of a genotype in a test environment *i* ; *Li* is the lowest mean yield of a genotype in a test environment *i* ; m_i is the mean yield of genotype *m* in a test environment *i* and n_i is the mean yield of genotype *n* in test environment *i*. Similarity index between two genotypes (*m* and *n*) was defined as the average of $S_{i(m,n)}$ across test environments when more than one test environment was used (Westcott, 1987). Eigenvalues of the PCOA are usually from the greatest to the least and the first eigenvalue is often called the leading eigenvalue. The PCOA is according to the sequential accumulation of the test environments according to their rank order, the environments being ranked in ascending order according to their overall means (Crossa et al. 1989). Each analysis produced a two-dimensional plot according to the first two PCOA axes. Using the eigenvectors of the main PCOA axes via the initial distance matrix can be visualized. Also, minimum spanning tree plots were drawn and those most stable genotypes with high mean yield (performance) were most distant from the center of the plot across sequential cycles (Flores et al. 1996). Ordination methods like PCOA displaying a set of data points in two dimensions make associations visible among the items in a higher dimensional space. It is a excellent tool to visualize large datasets of plant breeders with high

dimensionalities; it not only maintains the main trends in data but most of the information on details gets lost and when the intrinsic dimensions of data set are relatively high, conclusions can be misleading. PCOA can perform all calculations and plots by GENSTAT 12.1 (VSN International, 2009).

FACTOR ANALYSIS

Factor analysis (FA), as a multivariate statistical method, is used to explain variability in terms of a lower number of artificial variables or factors. It explores for linked variations in relation to artificial factors and the variables are modeled as linear combinations of potential factors as well as Error term. The FA is an ordination method related to PCA, the factors of the former being similar to the PCA of the later. It is first used as a psychometric model and is equivalent to low rank estimation of the matrix of original variables. A large number of related variables are reduced to a small number of factors (Cattell, 1965), and variation is described in terms of these factors. These general factors are common to all studied variables and in terms of factors, are unique to each variable. The axes of the initial factors may be rotated to oblique locations to conform to hypothetical ideas. The FA is related to the PCA, but the two are not completely identical; as FA uses a regression model to test hypotheses, the PCA is a descriptive method (Bartholomew et al., 2008).

The FA has been used to grasp interrelationships between different yield components of crops as well morphological properties of plants (Tadesse and Bekele, 2001; Tabrizi et al., 2011). Factors are conceptualized as real entities such as yield, but the components of PCA are abstractions that may not map easily onto real phenomena. PCA analyses total variance but FA shares variances that are analyzed. Godshalk and Timothy (1988) used a similar procedure to investigate several traits of switchgrass (*Panicum virgatum* L.) genotypes, and Saftner et al., (2008) compared the instrumental and sensory quality properties of blueberry fruit (*Vaccinium corymbosum* L.) genotypes. Peterson and Pfeiffer (1989) used FA to investigate investigation the underlying patterns and associations of multi-environment trials of wheat. They grouped the 56 locations into seven distinct regions or mega-environments. Fritsche-Neto et al., (2010) applied FA to GE interaction stratification in maize and reported that stratification of the test environment by FA was more selective in joining similarities according to a genotype's yield performance. Dettori et al., (2011) studied several quality traits of durum wheat in multi-environment trials and found that various quality traits could be regarded in low numbers of factors and one of the breeding lines indicated good quality traits as well as high mean yield in Italy.

THE ADDITIVE MAIN EFFECT AND MULTIPLICATIVE INTERACTION

The essential statistical background for the additive main effect and multiplicative interaction (AMMI) models was developed in 1952 by Williams (1952) after the invention PCA and ANOVA procedures. The AMMI model

consists of fitting an additive model (ANOVA) for producing general means, genotypes' means, and environments' means, and then fitting a multiplicative model (PCA) for the residual of an additive model or a GE interaction. It should be noted that PCA can be applied to original data or to GE interaction values; PCA is the first option and AMMI is the second. However, the usual investigation target is rather to use only one to a few PCA axes to summarize patterns in the GE interaction. The AMMI model came into widespread use in different scientific fields (Gollob 1968; Mandel 1971). The AMMI model is usually referred to as biplot analysis, even though this term was actually intended to refer to a graph or plot containing two kinds of points (Gabriel 1971). The AMMI model is an appropriate choice when both main effects and GE interaction are important.

According to Gauch (1992) and Gauch et al., (2008), the AMMI model is an effective tool for several targets: (i) understanding GE interaction, (ii) identifying mega-environment patterns, (iii) improving the accuracy of yield estimates, (iv) imputing missing data, and (v) increasing the flexibility of experimental designs. The AMMI model increases the probability of successfully selecting genotypes with the highest yields (Gauch and Zobel 1996). When a special different statistical method is appropriate, it is often most easily diagnosed by means of a preliminary analysis by an AMMI model (Bradu and Gabriel 1978). Sometimes the clearest understanding of a dataset emerges from several statistical analyses, each revealing various features of the data. If the design result in adjusted data is judged superior to the raw data, then those adjusted data should be supplied to the AMMI model (Gauch, 2006). Finally, these advantages imply larger selection gains in plant breeding and more reliable recommendations.

The AMMI model presents a new research tool with the possibility of producing adjusted means that often have predictive accuracy equivalent to original means. It can improve accuracy as much as a double or triple the data collection effort might (Gauch and Zobel, 1997). Most strategies for improving accuracy need a particular experimental design but the AMMI model has no such requirements, so it is applicable to historical data of experimental design. Thus, the power of the AMMI model to extract additional information will often make hard-won historical data worthy of renewed interest (Gauch et al., 2008). The three most common strategies of plant breeders for analyzing yield data are ANOVA (additive model), PCA (multiplicative model), and linear regression (Finlay and Wilkinson 1963). These approaches are largely subsumed and integrated by the AMMI model and this advantage is increased with dataset size and noise level.

Analysis of the GE interactions was made from the AMMI model. The model AMMI equation is:

$$Y_{ij} = \mu + g_i + e_j + \sum_{n=1}^{N} \lambda_n \gamma_{in} \delta_{jn} + \rho_{ij} \quad (1)$$

Where Y_{ij} is the yield of the *i*th genotype in the *j*th environment; μ is the grand mean; g_i and e_j are the genotype and environment deviations from the grand mean, respectively; λ_n is the eigenvalue of the IPC analysis axis n; γ_{in} and δ_{jn} are the genotype and environment eigenvectors for axis n; n is the number of principal components retained in the model and ρ_{ij} is the error term.

Zobel (1994) suggested the two EV stability parameter of AMMI according to the blow relation:

$$EV = \sum_{n=1}^{N} \gamma_{In}^2 / n \tag{2}$$

The AMGE and SIPC parameters according to Sneller et al., (1997) are expressed as:

$$AMGE = \sum_{n=1}^{N} \sum_{g=1}^{M} \lambda_n \gamma_{in} \delta_{jn} \tag{3}$$

$$SIPC = \sum_{n=1}^{n} \lambda_n^{0.5} \gamma_{in} \tag{4}$$

where M is the number of environments. Another stability parameter of AMMI according to the blow equation was proposed by Annicchiarico (1997).

$$D = \sqrt{\sum_{n=1}^{N} (\lambda_n \gamma_{in})^2} \tag{5}$$

The AMMI's stability value (ASV) is suggested by Purchase (1997):

$$ASV = \sqrt{\frac{SSIPC\ 1}{SSIPC\ 2} (PC\ 1)^2 + (PC\ 2)^2} \tag{6}$$

where, SS, sum of squares, IPC1, interaction of principal component analysis one, IPC2, interaction of principal component analysis two. For effective interpretation of GE interactions via AMMI model a new parameter as modified AMMI's stability value (MASV) is proposed:

$$MASV = \sqrt{\sum_{n=1}^{N-1} \left(\frac{SSIPC_{\ n}}{SSIPC_{\ n+1}}\right)(PC_{\ n})^2 + (PC_{\ N})^2} \tag{7}$$

The ASTAB stability parameter (Rao and Prabhakaran, 2005) is calculated using this formula:

$$ASTAB = \sum_{n=1}^{n} \lambda_n \gamma_{ni}^2 \tag{8}$$

Also, four I_i stability indexes (Rao and Prabhakaran, 2005) for simultaneous selection of both mean yield and stability were computed based on ASTAB stability parameter and mean yield as:

$$I_i = \frac{Y_{i.}}{\mu} + \alpha \frac{(1/ASTAB_i)}{(\sum_{n=1}^{n} ASTAB)/N} \quad (9)$$

where $Y_{i.}$ is the mean yield of the *i*th genotype; μ is the general mean; α is the ratio of weights given to the stability components (α= 1 for I_1, α= 0.66 for I_2, α= 0.43 for I_3, and α= 0.25 for I_4,). Most of the mentioned the AMMI stability parameters were used successfully in analyzing multi-environment trials data by Sabaghnia et al., (2008a) in lentil (*Lens culinaris* Medik), Dehghani et al., (2010) in chickpea (*Cicer arietinum* L.) and Sabaghnia et al., (2012b) in durum wheat (*Triticum turgidum* L.). These authors reported that the AMMI model as an appropriate statistical tool for investigating multi-environment trials.

The results of the AMMI model can be used to construct a biplot with a point for each genotype and for each environment, located in a graph indicating the main effects on the abscissa and the GE interaction scores on the ordinate (Gauch 1992; Gauch and Zobel, 1996). Such a graph as AMMI-1 biplot indicates, at a glance, both the main effects and the GE interaction effects for both genotypes and environments. Another useful biplot as AMMI-2 biplot indicates interaction PCA1 scores on the abscissa and interaction PCA2 scores on the ordinate (Gauch, 1992). Biplots can readily provide deep insights into a large, complex experiment (Kempton 1984; Zobel et al., 1988). Mega-environment analysis is included for the AMMI1 model through biplots (Gauch and Zobel 1997). One of the main objectives in the evaluation of multi-environment trials is to identify superior genotypes for a target area and to determine if this area can be subdivided into different mega-environments to better guide breeding strategies (Kang, 2002). The AMMI-2 biplot is an efficient means for detecting the possible mega-environments in multi-environment trials. The identification of mega-environments is involved with investigation of the annually repeatable GE interaction (Gauch and Zobel, 1996). For a particular mega-environment, genotypes are studied on the basis of mean yield and stability performance across test environments.

SHIFTED MULTIPLICATIVE MODEL

The shifted multiplicative model (SHMM) proposed by Seyedsadr and Cornelius (1992) groups genotypes into classes within which crossover interactions do not exist and within such groups, the genotype with the best mean would be the best. Multiplicative models for multi-environment trials have been used for studying GE interactions and for developing methods for grouping test environments and genotypes into groups with negligible crossover interaction (Cornelius et al., 1993; Crossa et al., 1993; Crossa and Cornelius, 1997; Abdalla et al., 1997). These models have an additive component (such as interception of linear regression, main effects of s environments and genotypes) and a multiplicative component (GE interaction effect). The SHMM model is a re-parameterization of the Tukey's (1949) model for testing non-additivity. The

singular vectors on effects for genotypes and test environments for the ordered components are primary, secondary, and so forth (Cornelius and Seyedsadr, 1997).

Cornelius et al., (1992) defined sufficient conditions for the absence of significant genotype crossover interaction in a set of environments and genotypes in the first SHMM model (SHMM1= model with one multiplicative term). In SHMM models, differences among genotypes in a special test environment are proportional to genotype differences in any other environment, but differences among environments with respect to the performance of a special genotype are proportional to environmental differences with respect to performance of any other genotype (Crossa et al., 1993). When an SHMM model is fitted to the dataset of multi-environment trials, secondary and perhaps even higher-order effects must be included if a sufficient fit is to be achieved. In clustering via a SHMM model, the measurement of distance between two test environments is taken as the residual mean square after fitting SHMM1 to the data from the two test environments subject to an additive constraint (Cornelius et al., 1993).

Cornelius et al., (1993) grouped 41 winter wheat (*Triticum aestivum* L.) genotypes into non-crossover interaction clusters via SHMM clustering method. Abdalla et al., (1997) clustered several durum wheat cultivars and related test locations via the SHMM model. Trethowan et al., (2001) used the SHMM clustering of test environments to investigate long-term associations between test locations for multi-environment trials on bread-wheat. They demonstrated the usefulness of SHMM for identifying key testing environments around the world. The SHMM clustering of genotypes is essentially by the same strategy as for clustering environments. The distance between two genotypes is defined using a constrained solution, when an SHMM1 model is fitted to the subset of data.

SITE REGRESSION BIPLOT

A usual phenomenon in most multi-environment trials is that environment is the predominant source of yield variation, and genotype and GE interaction are relatively small (Gauch and Zobel, 1996). The large magnitude of the environment effect is not relevant to genotype evaluation and only the genotype main effect and GE genotype are relevant to genotype evaluation. Therefore, it is essential to remove the environment effect from data and to focus on the other variation sources (G+GE). The GE received much attention because the G interaction is so much more straightforward to visualize and use. The GE interaction is validated by the numerous measures of stability index (Kang, 2002). Selection based on genotype effect alone may be justified if the GE interaction is known to be random and cannot be exploited (Yan et al., 2000). Exploration of the GE interaction began to make much sense following the advent of the concept of crossover interaction (Baker, 1990) or rank change (Huehn, 1996). Therefore, it seems that investigation into GE is much more meaningful when it is treated in conjunction with genotype effect. Sabaghnia et al., (2008b) in lentil (*Lens culinaris* Medik), Dehghani et al., (2009) in corn (*Zea*

mays L.) hybrids and Sabaghnia et al., (2012a) in durum wheat (*Triticum turgidum* L.), applied the GGE biplot model in evaluation of GE interaction, identification of mega-environments' structure and visualization of the "which-won-where" pattern in multi-environment trials. They detected mega-environment patterns for the mentioned crops and reported this method as an excellent tool for visual multi-environment trials' data analysis.

The biplot method (Gabriel, 1971) was expanded by Kempton (1984) and Zobel et al., (1988) highlighting the extensive usefulness of G+GE biplot (Yan et al., 2000). This method has strongly captured the imagination of plant breeders and agronomists. To explain GE interactions, a GGE biplot helps analyze multi-environment trials' data (Yan and Kang, 2002). These aspects make GGE biplot the most comprehensive tool in plant breeding. The GGE model deals with analysis of multi-environment trials' data and identifies (i) mega-environment for understanding the target environment, (ii) genotype evaluation for each mega-environment, (iii) understanding causes of GE interaction. The crossover interaction concept has led to investigations to identify homogeneous groups of environments with negligible crossover (Crossa and Cornelius, 1997). A further development of this concept is the emphasis on the "which-won-where" pattern (Gauch and Zobel, 1997). The GGE model is the most effective, useful and elegant way to reveal the "which-won-where" pattern of multi- environment trials' dataset. If there are important crossovers, the repeatability of the 'which-won-where' pattern is more important. This is a critical issue to division of the target environment into different mega-environments (Cooper et al., 1993) and presence of complex mega-environments.

The GGE model describes what is called genotype main effect in terms of GE interaction by definition of a constant value for a genotype across test environments. The genotypic PCA1 score of GGE model indicates a tendency of th genotypes to respond to environmental factors represented by the environmental PCA1 scores. The yield of genotype relative to PCA1 of GGE model is not the same in all environments; rather, it is proportional to the location of PCA1 scores. Thus, the GGE model emphasizes the fact that the genotype main effect not only has a genotypic basis but is also dependent on environmental conditions. Therefore, testing PCA1 scores not only detects genotypes with better overall performance but also suggests environmental conditions that facilitate identification of these genotypes. Yan and Rajcan (2002) reported that interactions between genotypic effects and environmental factors were the major causes of GE interaction for winter wheat yield due to PCA1 and PCA2 scores of GGE model. Also, understanding of the GE interaction is achievable if genotypic and environmental covariates are used in multi-environment trials (Yan and Kang, 2002).

STATISTICAL PACKAGES

All the reported multivariate statistical methods are difficult to apply by plant breeders without suitable and user-friendly software. MATMODEL is

software for AMMI and joint linear regression, which is available freely (Gauch, 2007). CROPSTAT of the International Rice Research Institute (IRRI, 2008) is freely available for performing ANOVA, joint linear regression, AMMI and pattern analysis. To compute SHMM model and for generating clusters of environments or genotypes, the Fortran-based program as EIGAOV is available from P.L. Cornelius, University of Kentucky, USA. Also, AGROBASE (Agronomix Software, 2009) commercial software performs ANOVA, joint linear regression and AMMI model; GGEbiplot (Yan, 2001) commercial software performs ANOVA, joint linear regression, AMMI model and GGE model; and GENSTAT 12.1 (VSN International, 2009) commercial software performs ANOVA, joint linear regression and AMMI model.

COMPARISON OF DIFFERENT STATISTICAL PROCEDURES

In recent decades the use of simple or first order multivariate procedures (PCA, PCOA and FA) in analysis of multi-environmental trials has been limited but the use of complex or second order multivariate procedures (AMMI, SHMM and GGE) in analyzing multi-environmental trials is significantly increased. It seems that most plant breeders like to determine the nature and pattern of GE interaction using more efficient statistical methods as well as possible. Second order multivariate methods have a good ability to partition a signal-rich model from a noise-rich discarded residual (Cornelius and Crossa, 1999) while cluster analyses lack that ability and are therefore quite vulnerable to noise (Smith and Gauch, 1992). After simple multivariate procedures; the AMMI model began to attract the attention of plant breeders after Zobel et al., (1988), which has become a popular tool among researchers for understanding the GE interaction. Then, the SHMM and the GGE biplot are suggested to explore the structure of the GE interaction. Common features of AMMI, SHMM, and GGE models are that they all use PCA, but they differ in processes of data transformation prior to PCA application and differ in methods of interpretation in terms of parameters and graphs. The AMMI model treats three sources of total variation (G, E and GE interaction) separately, whereas SHMM subtracts a single value (the shift parameter) from every matrix cell. The GGE model subtracts the environment main effect and then performs PCA on the remaining variation (G+GE interaction). Comparison of the AMMI1, SHMM2 and GGE2 models for the purpose of GE interaction exploration indicated that the AMMI model is the only contender among these three options, because it analyzes the GE interaction itself apart from other variation sources.

The mega-environment identification of the GGE2 biplot is comparable to an AMMI1 biplot while no GGE biplot has yet been developed that corresponds to the AMMI2. Therefore, the AMMI model is superior to the GGE model for mega-environment identification through a biplot in the complex interaction. The AMMI model is unique in analyzing effects separately, without confounding the genotype effect with GE effects, which as been a basic requirement in plant

breeding. In contrast the SHMM model is completely unaware of this simple and important distinction. Finally, it seems that the AMMI model is better than other methods (SHMM and GGE) for analyzing multi-environment trials datasets. Agronomists attempt to improve environmental conditions, whereas plant breeders try to improve genotypes. Therefore, it seems that it is better to consider these effects (G+E+GE interaction) separately, and so the AMMI model is preferred.

CONCLUSIONS

Among different clustering methods, there are four methods that benefit most from the special F-test for determination of a cutoff point, a good procedure for classification of genotypes in multi-environment trials. The main reason highlighting the AMMI model as the most appropriate one for breeding programs is that the ANOVA section of the AMMI model can separate effects of genotype and environment from the GE interaction, and the PCA section of the AMMI model can separate the signal-rich portion of the GE interaction. The AMMI model offers better opportunities than GGE and SHMM models for graphic analysis of the GE interaction and mega-environment identifications but options of GGE biplot software are more acceptable for most researchers. Therefore, it seems that developing similar statistical packages for an AMMI model could encourage plant breeders as well as other researchers to use this powerful statistical procedure in their investigations.

REFERENCES

Abdalla, O.S., Crossa, J. & Cornelius, P.L. (1997): Results and biological interpretation of shifted multiplicative model clustering of durum wheat cultivars and test sites. Crop Sci. 37, 88–97.

Abou-El-Fittouh, H.A., Rawlings, J.O. & Miller, P.A. (1969): Classification of environments to control genotype by environment interactions with an application to cotton. Crop Sci. 9, 135–140.

Agronomix Software Inc. (2009): Agrobase Generation II. Version 18. Agronomix Software Inc. Winnipeg, MB. Canada.

Allard, R.W. & Bradshaw, A.D. (1964): Implication of genotype × environmental interaction in applied plant breeding. Crop Sci. 5, 503–506.

Annicchiarico, P. (1997): Joint regression vs AMMI analysis of genotypeenvironment interactions for cereals in Italy. Euphytica 94, 53–62.

Bartholomew, D. J., Steele, F., Galbraith, J., & Moustaki, I. (2008): Analysis of Multivariate Social Science Data (2 ed.). New York: Chapman & Hall/CRC.

Bradu, D. & Gabriel, K.R. (1978): The biplot as a diagnostic tool for models of two-way tables. Technometrics 20, 47–68.

Calinski, T. & Corsten, L.C.A. (1985): Clustering means in ANOVA by simultaneously testing, Biometrics 41, 39–40.

Cattell, R.B. (1965): A Biometrics Invited Paper. Factor Analysis: An Introduction to Essentials II. The role of factor analysis in research. Biometrics 21, 405–435.

Cooper, M., DeLacy, I.H. & Eisemann R.L. (1993): Recent advances in the study of genotype × environment interactions and their application to plant breeding, p.116–131. In: Imrie B.C. & Hacker J.B. (Eds.), Focused plant improvement: towards responsible and sustainable agriculture, Proc. 10th Australia plant

breeding conference. Vol. I. Organizing Committee, Australian Convention and Travel Service, Canberra, Australia.

Cornelius, P.L. & Crossa, J. (1999): Prediction assessment of shrinkage estimators of multiplicative models for multi-environment cultivar trials. Crop Sci. 39, 998–1009.

Cornelius, P.L. & Seyedsadr, M. (1997): Estimation of general linear–bilinear models for two-way tables. Journal of Statistical Computation and Simulation 58, 287–322.

Cornelius, P.L., van Sanford, D.A. & Seyedsadr, M. (1993): Clustering cultivars into groups without rank-change interactions. Crop Scie. 33, 1193–1200.

Crossa J., Westcott, B. & Gonzalez, C. (1989): The yield stability of CIMMYT'S maize germplasm. Euphytica 40, 245–251.

Crossa, J. & Cornelius, P.L. (1997): Sites regression and shifted multiplicative model clustering of cultivar trial sites under heterogeneity of error variances. Crop Sci. 37, 406–415.

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

Crossa, J., Cornelius, P.L., Seyedsadr, M.S. & Byrne, P. (1993): A shifted multiplicative model cluster analysis for grouping environments without genotypic rank range. Theor. Appl. Genet. 85, 577–586.

Dehghani, H., Sabaghnia, N. & Moghaddam, M. (2009): Interpretation of genotype-by-environment interaction for late maize hybrids' grain yield using a biplot method. Turk. J. Agric. Fores. 33, 139–148.

Dehghani, H., Sabaghpour, S.H. & Ebadi, A. (2010): Study of genotype × environment interaction for chickpea yield in Iran. Agron. J. 102, 1–8.

DeLacy, I.H., Basford, K.E., Cooper, M., Bull, J.K. & McLaren, C.G. (1996): Analysis of multi-environment data - An historical perspective. In M. Cooper & G.L. Hammer, eds. Plant adaptation and crop improvement, p. 39–124. Wallingford, UK, CABI.

Dettori, M., Crossa, J., Ammar, K., Peña R.G. & Varela, M. 2011. Three-mode principal Component analysis of genotype-by-environment-by-trait data in durum wheat. J. Crop Improv. 25, 619–649.

Eberhart, S.A. & Russell, W.A. (1966): Stability parameters for comparing varieties. Crop Sci. 6, 36–40.

Edwards, A.W.F., & Cavalli-Sforza, L.L. (1964): Reconstruction of evolutionary trees. pp. 67–76 in Phenetic and Phylogenetic Classification, ed. V. H. Heywood and J. McNeill. Systematics Association pub. no. 6, London.

Eisemann, R.L., Cooper M., & Woodruff D.R. (1990): Beyond the analytical methodology, better interpretation and exploitation of GE interaction in plant breeding, p. 108–117. In: Kang, M.S. (Ed.), Genotype-by-environment interaction and plant breeding, Louisiana State University Agric. Center, Baton Rouge, LA.

Finlay, K.W. & Wilkinson, G.N. (1963): The analysis of adaptation in a plantbreeding programme. Aust. J. Agric. Res. 14, 742–754.

Flores, F., Moreno, M.T., Martinez, A. & Cubero, J.I. (1996): Genotype × environment interaction in faba bean : Comparison of AMMI and Principal Coordinate models. Field Crops Res. 47, 117–127.

Fox, P.N. & Rosielle, A.A. (1982): Reducing the influence of environmental main-effects on pattern analysis of plant breeding environments. Euphytica 31, 645-656.

Freeman, G.H. & Dowker, B.D. (1973): The analysis of variation between and within genotypes and environments. Heredity 30, 97–109.

Fritsche-Neto, R., Miranda, G.V., de Lima, R.O., de Souza, H.N. (2010): Factor analysis and SREG GGE biplot for the genotype × environment interaction stratification in maize. Ciência Rural 40, 1043–1048.

Gabriel, K.R. (1971): The biplot graphic display of matrices with application to principal component analysis. Biometrika 58, 453–467.

Gauch, H.G. & Zobel, R.W. (1996): AMMI analysis of yield trials. In M.S. Kang & Gauch, H.G. eds. Genotype-by-environment interaction, p. 85-122. Boca Raton, FL, CRC Press.

Gauch, H.G. & Zobel, R.W. (1997): Identifying mega-environments and targeting genotypes. Crop Sci. 37, 311–326.

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

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

Gauch, H.G. (2006): Statistical analysis of yield trials by AMMI and GGE. Crop Sci. 46, 1488–1500.

Gauch, H.G. (2007): MATMODEL version 3.0: Open source software for AMMI and related analyses. Available at http://www.css.cornell.edu/staff/gauch (verified 12 August 2012). Crop and Soil Sciences, Cornell Univ., Ithaca, NY.

Gauch, H.G., Piepho, H.P. & Annicchiarico, P. (2008): Statistical analysis of yield trials by AMMI and GGE: Further considerations. Crop Sci. 48, 866–889.

Godshalk, E.B., & Timothy D.H. (1988): Factor and principal component analyses as alternatives to index selection. Theor. Appl. Genet. 76, 352–360.

Gollob, H.F. (1968): A statistical model which combines features of factor analytic and analysis of variance techniques. Psychometrika 33, 73–115.

Gower, J.C. (1966): Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika 53, 325–338.

Gower, J.C. (1967): Multivariate analysis and multidimensional geometry. Statistician 17: 13–28.

Gower, J.C. (1971): Statistical methods of comparing different multivariate analyses of the same data. In F. R Hodson, D. G. Kcadall, and P. Tau eds. Mathematics in the Archaeological and Historical Sciences. Chicago: Aldine pp. 138–149.

Guitard, A.A. (1960): The use of diallel correlations for determining the relative locational performance of varieties of barley. Can. J. Plant Sci. 40, 645–651.

Habgood, R.M. (1977): Estimation of genetic diversity of self-fertilising cereal cultivars based on genotype-environment interactions. Euphytica 26, 485–489.

Hanson, W.D. (1970): Genotypic stability. Theor.Appl.Genet. 40:226–231.

Hanson, W.D. (1994) Distance statistics and interpretation of southern states regional soybean tests. Crop Sci. 34, 1498–1504

Hernandez, C.M., Crossa, J. & Castillo, A. (1993): The area under the function: an index for selecting desirable genotypes. Theor. Appl. Genet. 87, 409–415.

Huehn, M. (1979): Beitrage zur erfassung der phanotypischen stabilitat. EDV Med. Biol. 10, 112–117.

Huehn, M. (1996): Non-parametric analysis of genotype × environment interactions by ranks. p. 213–228. In M.S. Kang and H.G. Gauch (ed.) Genotype by environment interaction. CRC Press, Boca Raton, FL.

Ibanmez, M.A., Direnzo, M.A., Samame, S.S., Bonamico, N.C. & Poverene, M.M. (2001): Genotype–environment interaction of lovegrass forage yield in the semi-arid region of Argentina. J. Agric.Sci. 137, 329–336.

IRRI. (2008): CROPSTAT for Windows, version 5. Available at http://www.irri.org/science/software/irristat.asp (verified 27 August 2012).

Biometrics and Bioinformatics Unit, Int. Rice Research Inst., Los Baños, Philippines.
Jolliffe. I.T. (2002): Principal Component Analysis. Springer-Verlag, 2nd edition, 502 PP.
Kang, M.S. (2002): Quantitative genetics, genomics, and plant breeding. Wallingford, UK, CABI.
Kempton, R.A. (1984): The use of biplots in interpreting variety by environment interactions. J. Agric. Sci., 103, 123–135.
Lin, C.S. & Butler, G. (1990): Cluster analyses for analyzing two-way classification data. Agron. J. 82, 344–348.
Lin, C.S. & Thompson, B. (1975): An empirical method of grouping genotypes based on a linear function of the genotype environment interaction. Heredity 34, 255–263.
Lin, C.S. (1982): Grouping genotypes by a cluster method directly related to genotype–environment interaction mean square. Theor. Appl. Genet. 62, 277–280.
Lin, C.S., Binns, M.R., & Lefkovitch, L.P. (1986): Stability analysis : Where do we stand? Crop Sci. 26, 894–900.
Lin, C.S., Butler, G., Hall, I. & Nault, C. (1992):Program for investigating genotype-environment interactions. Agron. J. 84, 121–124.
Mandel, J. (1971): A new analysis of variance model for non-additive data. Technometrics 13. 1–18.
Muir, W., Nyquist, W. E. & Xu, S. (1992): Alternative partitioning of the genotype - by - environment interaction. Theor. Appl. Genet. 84, 193–200.
Mungomery, V.E., Shorter, R. & Byth, D.E. (1974): Genotype x environment interactions and environmental adaptation. I. Pattern analysis - Application to soybean populations. Aust. J. Agric. Res. 25, 59–72.
Perkins, J.M. & Jinks, J.L. (1971): Analysis of genotype x Environment interaction in triple test cross data. Heredity 26, 203–209.
Perkins, J.M. (1972): The principal component analysis of genotype environmental interactions and physical measures of the environment. Heredity 29, 51–70.
Peterson, C.J., & Pfeiffer, W.H. (1989): International winter wheat evaluation: relationships among test sites based on cultivar performance. Crop Sci. 29, 276–282.
Polignano G.B., Uggenti P. & Perrino P. (1989): Pattern analysis and genotypic × environmental interactions in faba bean (*Vicia faba* L.) populations. Euphytica 40, 31–41.
Purchase, J.L. (1997): Parametric analysis to describe G × E interaction and yield stability in winter wheat. Ph.D. thesis. Dep. of Agronomy, Faculty of Agriculture, Univ. of the Orange Free State, Bloemfontein, South Africa.
Rao, A.R. Polignano Prabhakaran, V.T. (2005): Use of AMMI in simultaneous selection of genotypes for yield and stability. Ind Soc Agril Statist 59, 76–82.
Sabaghnia, N, Mohammadi, M. & Karimizadeh, R. (2012a): Targeting improved genotypes onto durum wheat growing locations using site regression model. Natura Montenegrina 11, In Press.
Sabaghnia, N, Mohammadi, M. & Karimizadeh, R. (2012b): The evaluation of genotype × environment interactions of durum wheat's yield using of the AMMI model. Agric. Fores. 55, 5–21.
Sabaghnia, N., Dehghani, H. & Sabaghpour, S.H. (2008a): Graphic analysis of genotype by environment interaction for lentil yield in Iran. Agron. J. 100, 760–764.

Sabaghnia, N., Sabaghpour, S.H. & Dehghani, H. (2008b): The use of an AMMI model and its parameters to analyze yield stability in multi-environment trials. J. Agric. Sci. 146, 571–581.

Saftner, R, Polashock, J. Ehlenfeldt, M. & Vinyard, B. (2008): Instrumental and sensory quality characteristics of blueberry fruit from twelve cultivars. Postharvest Biol. Tech. 49, 19–26.

Seyedsadr, M. & Cornelius, P.L. (1992): Shifted multiplicative models for nonadditive two-way tables. Commu. Statis. Simu. Comp. 21, 807–822.

Shukla, G.K. (1972): Some statistical aspects of partitioning genotype-environmental components of variability. Heredity 29, 237–245.

Smith, M.F. & Gauch, H.G. (1992): Effects of noise on AMMI and hierarchical classification analyses. S. Afr. Stat. J. 26, 121–142.

Sneller, C.H., Kilgore-Norquest, L. & Dombek, D. (1997): Repeatability of yield stability statistics in soybean. Crop Sci. 37, 383–390.

Tabrizi, N.F.M., Rashidi, V.. Yarnia, M. & Farajzadeh, E. (2011): Relationships between morphological characteristics and bean yield under drought stress and non-stress conditions. J. Food Agric. Envir. 9, 264–267.

Tadesse, W. & Bekele E. (2001): Factor analysis of components of yield in grasspea (*Lathyrus sativus* L.). Lathyrus Lathyrism Newsl. 2, 91–93.

Trethowan, R.M., Crossa, J., van Ginkel, M. & Rajaram, S. (2001): Relationships among bread wheat international yield testing locations in dry areas. Crop Sci. 41, 1461–1469.

Tukey, J.W. (1949): One degree of freedom for non-additivity. Biometrics 5, 232–242.

van Oosterom, E.J., Kleijn, D., Ceccarelli, S. & Nachit, M.M. (1993): Genotype-by environment interaction in barley in the Mediterranean region. Crop Sci. 33, 669–674.

VSN International. (2009): GENSTAT statistical package for windows version 12.1. Reference Manual. Clarendon Press, Oxford, UK.

Westcott, B. (1987): A method of assessing the yield stability of crop genotypes. J. Agric. Sci. 108, 267–274.

Williams, E.J. (1952): The interpretation of interactions in factorial experiments. Biometrika 39, 65–81.

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

Yan, W. & Rajcan, I. (2002): Biplot evaluation of test sites and trait relations of soybean in Ontario. Crop Sci. 42, 11–20.

Yan, W. (2001): GGEbiplot—a Windows application for graphical analysis of multi-environment trial data and other types of two-way data. Agron. J. 93, 1111–1118.

Yan, W., & Kang, M.S. (2002): GGE biplot analysis: A graphical tool for breeders, geneticists, and agronomists. CRC Press, Boca Raton, FL.

Yan, W., Hunt, L.A., Sheng, Q. & Szlavnics, Z. (2000): Cultivar evaluation and mega-environment investigation based on the GGE biplot. Crop Sci. 40, 597–605.

Zobel, R.W. (1994): Stress resistance and root systems. p. 80–99. In Proc. Of the Workshop on Adaptation of Plants to Soil Stress. 1–4 Aug. 1993. INTSORMIL Publ. 94–2. Inst. of Agriculture and Natural Resources, Univ. of Nebraska, Lincoln.

Zobel, R.W., Wright, M.J. & Gauch, H.G. 1988. Statistical analysis of a yield trial. Agron. J. 80, 388–393.

Naser SABAGHNIA

VIŠEFAKTORIJALNA STATISTIČKA ANALIZA INTERAKCIJE GENOTIP × ŽIVOTNA SREDINA KOD OGLEDA SA VIŠE FAKTORA ŽIVOTNE SREDINE U PROGRAMIMA OPLEMENJIVANJA

SAŽETAK

U završnim fazama u programima oplemenjivanja biljaka, veliki broj novih poboljšanih genotipova je testiran u širokom opsegu faktora spoljašnje sredine i osnovne statističke metode koje se koriste za modeliranje ovog sistema, i pri tom mogu biti prilično komplikovane. Obično prisustvo dejstva interakcije genotipa i sredine (GE) komplikuje izbor najpovoljnijih genotipova za određene uslove spoljašnje sredine. Postoji nekoliko dostupnih statističkih metoda za analizu rezultata ogleda u koje je uključeno više faktora spoljašnje sredine uključujući niz jednofaktorijalnih i višefaktorijalnih procedura. Jednofaktorijalne metode imaju neadekvatan kapacitet da u potpunosti objasne strukturu GE interakcije, jer oni pokušavaju da definišu GE interakciju na osnovu jednog ili dva parametra, dok je multiplikativna GE interakcija mnogo kompleksnija, te se ne može ograničiti na samo nekoliko parametara. Nasuprot tome, višefaktorijane statističke metode istražuju više aspekata GE interakcije i pokušavaju da uzmu u obzir više informacija. Najčešće korišćene višefaktorijalne statističke metode su: analiza grupe (CA), analiza glavnih djelova (PCA), analiza glavnih koordinata (PCOA), faktorska analiza (FA), dodatni glavni efekat i multiplikativna interakcija (AMMI), izmjenjen multiplikativni model (SHMM), sajt biplot regresija (GGE). Ovaj rad daje pregled ovih višefaktorijalnih statističkih metoda za analizu podataka u ogledu sa više faktora spoljašnje sredine. Nekoliko AMMI parametara stabilnosti je razmatrano i poređena su tri ova značajna modela (AMMI, GGE i SHMM).

Ključne riječi: adaptacija, biplot, analiza stabilnosti, prinos