

*Review*

## Evolution of methodology for the study of adaptability and stability in cultivated species

Leonardo Castelo Branco Carvalho<sup>1</sup>, Kaesel Jackson Damasceno-Silva<sup>2\*</sup>, Maurisrael de Moura Rocha<sup>2</sup> and Giancarlo Conde Xavier Oliveira<sup>1</sup>

<sup>1</sup>Escola Superior de Agricultura "Luiz de Queiroz" Universidade de São Paulo, Avenida Pádua Dias, 11 - Cx. Postal 9, 13418-900, Piracicaba, SP, Brazil.

<sup>2</sup>Embrapa Meio Norte, Avenida Duque de Caxias, 5650, Bairro Buenos Aires, 64,006-220, Teresina, PI, Brazil.

Received 31 October, 2015; Accepted 10 February, 2016

The GxE interaction only became widely discussed from evolutionary studies and evaluations of the causes of behavioral changes of species cultivated in environments. In the last 60 years, several methodologies for the study of adaptability and stability of genotypes in multiple environments trials were developed in order to assist the breeder's choice regarding which genotypes are more stable and which are the most suitable for the crops in the most diverse environments. The methods that use linear regression analysis were the first to be used in a general way by breeders, followed by multivariate analysis methods and mixed models. The need to identify the genetic and environmental causes that are behind the GxE interaction led to the development of new models that include the use of covariates and which can also include both multivariate methods and mixed modeling. However, further studies are needed to identify the causes of GxE interaction as well as for the more accurate measurement of its effects on phenotypic expression of varieties in competition trials carried out in genetic breeding programs.

**Key words:** Adaptability, stability, GxE interaction, genetic breeding, covariates.

### INTRODUCTION

For genetic breeding programs, there is the inherent difficulty in identifying varieties with superior performance in various environments because, even when isolating from the space factor, that is, when such genotypes are planted in similar sites (usually resulting from a subclass of places obtained via stratification), they have accentuated interaction with different crops both within the same year and with different years (Eberhart and

Russell, 1966).

The ability that certain genotypes have to grow well in a wide range of environmental conditions is, therefore, of major importance for Agronomy, especially in places where such conditions are extremely variable and, until mid-1950's, the effects of the interaction genotypes x environments were estimated only via general mean, according to the mean performance of varieties in various

\*Corresponding author. Email: [kaesel.damasceno@embrapa.br](mailto:kaesel.damasceno@embrapa.br).

locations and years (Finlay and Wilkinson, 1963).

Sprague and Federer (1951) were pioneers by showing how the variance components can be used to separate the effects of genotypes, environments, and the interaction between them, equaling the mean square obtained by analysis of variance (ANOVA) to their respective mathematical expectations. Then, Plaisted and Peterson (1959) proposed a new methodology for evaluating the influence of this interaction, which consists of applying a combined analysis of variance, that is, an analysis considering all varieties at all locations in a given year, also known as "two-factor" analysis.

The variation observed between varieties is dynamic in some cases (Finlay and Wilkinson, 1963), and the breeders find themselves faced with the choice between selecting varieties adapted to a particular range of environments (or specific sites), or obtaining varieties with broad adaptability and which, therefore, have a good performance in a range of larger environments. Varieties having site adaptability can be very useful, especially when it comes to environments with unusual conditions, or difficult cultivation, or even extreme conditions.

Several authors, such as Salmon (1951), Horner and Frey (1957), and Sandison and Bartlett (1958), discussed the theme using techniques that consider the interaction genotypes x sites, or genotypes x years (or crop) as an adaptability measure. Such techniques are of low precision when it comes to many environments or genotypes to be evaluated.

Meanwhile, and in a non-integrated way, experiments evaluating the nature of phenotypic stability gave experimental support for the understanding of the interaction genotypes x environments (Lewis, 1954; Dobzhansky and Levene, 1955; Williams, 1960). Gripping and Langridge (1963), for instance, conducted a study on the influence of heterosis on phenotypic plasticity in *Arabidopsis thaliana* and concluded that the hybrids of this species showed greater stability than homozygous individuals.

From the 1960's, several methodologies for the evaluation of adaptability and stability of genotypes in multi-environments trials have been developed, most of them still used nowadays in breeding programs for plant species cultivated worldwide. Among these, the most widely used have been based on simple linear regression (Yates and Cochran, 1938; Finlay and Wilkinson, 1963; Eberhart and Russell, 1966), multivariate analysis (Mandel, 1971; Kempton, 1984; Gauch, 1988; Zobel et al., 1988; Crossa, 1990; Yan et al., 2000), and mixed models (Piepho, 1997; Resende and Thompson, 2004).

Given this, the study proposes to discuss the technical and practical aspects of the main methodologies available in literature and used to evaluate the adaptability and stability of genotypes, as well as to trace an overview on methods that have been proposed more recently and the challenges for the evaluation of genotypes in trials of multiple environments by genetic breeding programs at the present time.

## METHODOLOGIES BASED ON ANOVA COMPONENTS

In addition to the method of Plaisted and Peterson (1959), until the beginning of the 1960s, some methodologies to evaluate the phenotypic stability of genotypes were based only on ANOVA components, among which outstands the Wricke methodology (1962), popularly known as "Ecovalence". Such parameter is estimated by the decomposition of the sum of squares of the GxE interaction (quite similar to the model of Plaisted and Peterson, which in turn proposes the decomposition of variance of the GxE interaction) in parts related to genotypes in an isolated manner, which is given by the expression:

$$\omega_i = r \sum_j (Y_{ij} - \bar{Y}_i - \bar{Y}_j + \bar{Y})^2$$

In which:  $Y_{ij}$  is the mean of genotype  $i$  in the environment  $j$ ;  $\bar{Y}_i$  and  $\bar{Y}_j$  correspond to the mean of genotype  $i$  and to the mean of environment  $j$ , respectively; and  $\bar{Y}$  is the general mean.

The parameter  $\omega_i$  measures, as a stability factor, the contribution of each genotype for the GxE interaction component, in which the genotypes that contribute less to the interaction are considered the most stable.

## METHODOLOGIES BASED ON REGRESSION METHODS

Finlay and Wilkinson (1963), based on Yates and Cochran (1938), proposed a methodology using linear regression models to compare the performance of a set of varieties evaluated in multiple sites and years in which, for each variety, a regression of their mean was obtained regarding the overall mean of all varieties in each site per year. In addition, each environment was classified as favorable or unfavorable according to the mean of all varieties in that environment.

These authors have modeled environmental factors, simply in terms of the productivity response of genotypes. Thus, the varieties that have regression coefficients equal or close to 1(one) are considered varieties of mean stability. Among these, those that are associated with a high productivity have broad adaptation, and those associated with low productivity are weakly adapted to all environments. Varieties with coefficient significantly greater than 1.0 are considered especially adapted to favorable environments, but have low stability, and those which have coefficient lesser than 1, or tending to 0, are considered more stable and with adaptability to unfavorable environments. Therefore, the optimal variety would be the one with a good performance in all environments and with high stability (that is, regression coefficient close to 0).

Biologically, the interpretation of this factor is that such varieties are so stable that they are unable to respond to any improvement in environmental conditions. According to Eberhart and Russell (1966), the use of the regression coefficient and the deviations from the straight line as parameters of stability aimed at helping to solve this problem. Being  $i$  the environmental index for the regression of each variety, in each environment  $j$ , defined by:

$$l_j = [(\sum_{i=1}^n Y_{ij}) / v] - [(\sum_{j=1}^n \sum_{i=1}^n Y_{ij}) / vn]$$

Where  $Y_{ij}$  is the mean of the  $i$ -th variety within the  $j$ -th environment,  $v$  is the number of varieties, and  $n$  the number of environments.

Thus, the estimation of the two mentioned parameters is usually defined by the following model:

$$Y_{ij} = \mu_i + \beta_i l_j + \delta_{ij},$$

where the first parameter presented, the regression coefficient ( $\beta_i$ ), is the same proposed by Finlay and Wilkinson (1963), defined as:  $b_i = \sum_j Y_{ij} l_j / \sum_j l_j^2$ , and the second parameter ( $\delta_{ij}$ ), is estimated via the sum of squares of regression deviations, as follows:  $s_{di}^2 = [\sum_j \delta_{ij}^2 / (n - 2)] - [s_e^2 / r]$ , in which  $s_e^2 / r$  is the estimate of the error set. This procedure thus decomposes the sum of squares of the GxE interaction in two parts: the variation due to the response of each variety regarding the environmental index, and the regression deviations regarding such index. The optimal genotype then becomes the one that features high productivity, associated with a regression coefficient as close as possible to 1, and regression deviations close to 0.

Some authors claim that the approaches that only comprise regression techniques are useful only as preliminary evaluations, for they present, most of the time, large linearity deviations, making the selection of genotypes biased and applied exclusively to the set of the evaluated varieties, being seen as a very distant simplification of the reality presented in genetic breeding experiments (Witcombe and Whittington, 1971). Schlichting (1986) States that there are two important issues in methodologies that are used in the regression analysis: (1) The means and the coefficients assigned to genotypes tend to be positively correlated, that is, stable genotypes tend to have lower expression of the character in question, and (2) the assumption of linearity is not usually met.

## NONPARAMETRIC METHODOLOGIES

When data do not meet the assumptions of regression analyses, an alternative is to use nonparametric analyses. In genetic breeding, more precisely in the context of the evaluation of genotypes in MET's, there

are some proposals for adaptability and stability evaluation of genotypes based on nonparametric statistics. Studies, such as those of Hühn (1979), Nassar and Hühn (1987), Kang (1988), Lin and Binns (1988), Fox et al. (1990), and Thennarasu (1995), are among the most cited regarding this aspect.

Nassar and Hühn (1987) developed methodologies that have as a fundamental characteristic the interpretation of several measures based on the ranking of genotypes, although they are independent of each other. The most widely used, according to literature, are:  $S^{(1)}$  (mean of the absolute rank differences of a genotype over the  $n$  environments),  $S^{(2)}$  (variance among the ranks over the  $k$  environments),  $S^{(3)}$  (sum of the absolute deviations), and  $S^{(6)}$  (relative sum of squares of rank for each genotype). Such measures can be mathematically described as follows:

$$S_i^{(1)} = 2 \sum_{j=1}^{q-1} \sum_{j+1}^q |r_{ij} - r_{i,j+1}| / [q(q-1)]$$

$$S_i^{(2)} = \sum_{j=1}^q (r_{ij} - \bar{r}_i)^2 / (q-1)$$

$$S_i^{(3)} = \sum_{j=1}^q (r_{ij} - \bar{r}_i)^2 / \bar{r}_i$$

$$S_i^{(6)} = \sum_{j=1}^q |r_{ij} - \bar{r}_i| / \bar{r}_i$$

In which:  $r_{ij}$  is the rank of  $i$ th genotype in the  $j$ th environment (for  $q$  environments), and  $\bar{r}_i$  is the mean rank across all environments for the  $i$ th genotype.

Kang (1988) and Fox et al. (1990) proposed other nonparametric methods, which, in turn, calculate only one statistic and classify the most stable genotypes. Kang nonparametric stability (Rank-sum) uses both "trait single value" and stability variance (Shukla, 1972), and the genotype with the lowest ranksum is commonly the most favorable one (both the highest yielding genotype and the genotype with the lowest stability variance are ranked 1). Fox et al. is based, in many cases, in the "TOP third" statistical ranking, where a stratified ranking of the genotypes at each environment separately is done. The proportion of sites at which the genotype occurred in the top third are expressed in TOP ranking.

Thennarasu (1995) proposed four statistics ( $NP^{(1)}$ ,  $NP^{(2)}$ ,  $NP^{(3)}$ , and  $NP^{(4)}$ ) for stability measures. These measures are based on ranks of adjusted means of the genotypes in each environment, and stable genotypes as those whose position in relation to the others remained unaltered in the set of environments. Thennarasu measures are defined as:

$$NP_i^{(1)} = \frac{1}{n} \sum_{j=1}^n |r_{ij} - M'_{di}|$$

$$NP_i^{(2)} = \frac{1}{n} \left( \sum_{j=1}^n |r_{ij} - M'_{di}| / M_{di} \right)$$

$$NP_i^{(3)} = \frac{\sqrt{\sum (r_{ij} - \bar{r}'_i)^2 / n}}{\bar{r}_i}$$

$$NP_i^{(4)} = \frac{2}{n(n-1)} \left[ \sum_{j=1}^{n-1} \sum_{(j'+j+1)}^n |r_{ij} - r_{ij'}| / \bar{r}_i \right]$$

In which:  $r_{ij}$  is the rank of adjusted values ( $x_{ij}^* = x_{ij} - x_i$ ),  $\bar{r}'_i$  and  $M'_{di}$  are the mean and median ranks for adjusted values, and  $\bar{r}_i$  and  $M_{di}$  are the same parameters computed from the phenotypic unadjusted values.

The methodology proposed by Lin and Binn (1988) somewhat differs from the cited ones, since it is not explicitly based on values obtained from the ranking of genotypes, considering that this methodology is based on absolute values obtained for the rated character and not on the numbering of the ranking of genotypes. Lin and Binns proposed obtaining the stability statistics  $P_i$ , which is given by the expression:

$$P_i = \sum_{j=1}^n (X_{ij} - M_j)^2 / 2n$$

In which:  $P_i$  = superiority index of  $i$ -th cultivar;  $X_{ij}$  = productivity of  $i$ -th cultivar planted in the  $j$ -th site;  $M_j$  = maximum response obtained among all the cultivars in the  $j$ -th site;  $n$  = number of sites. This expression can be developed into:

$$P_i = \left[ n(\bar{X}_i - \bar{M})^2 + \sum_{j=1}^n (X_{ij} - \bar{X}_i - M_j + \bar{M})^2 \right] / 2n$$

In which:  $\bar{X}_i$  is the mean of the character obtained in  $n$  environments; and  $\bar{M}$  the mean of the maximum responses of genotypes in all environments.

Such methodologies, despite being considered as alternatives, especially for their lack of robustness, may be considered preferable when applications of regression methods are not possible, or when there is a need for analyses of easier implementation and of most objective interpretation. This fact is clearly observed, for instance, in the methodologies of Kang (1988), Lin and Binns (1988), and Fox et al. (1990), which feature a single value as a stability measure.

## METHODOLOGIES BASED ON MULTIVARIATE METHODS

With the advent of more sophisticated computational

resources and computers with greater processing capacity, other methodologies, such as those based on multivariate analysis, became more accessible and are preferably used to the extent that most statistical packages were made available. One of the bases for conducting adaptability and stability analyses via multivariate models is the principal component analysis (PCA), which has as its essence the application of the SVD (*Singular Value Decomposition*) method, which, in turn, performs the linear decomposition of variables contained in a data array in an iterative manner, in order to summarize the information contained in a smaller number of explanatory vectors.

Among the most used methodologies in adaptability genetic studies, the following models outstands: AMMI (*Additive Main Effect and Multiplicative Interaction*) and GGE Biplot (*Genotype plus Genotype by Environment*) (Kempton, 1984; Zobel et al., 1988; Crossa, 1990; Gauch, 1992; Yan et al., 2000). In the AMMI models, developed by Mandel (1971) and popularized by Zobel et al. (1988) and Gauch (1992), the magnitude of the GxE interaction is estimated according to the response of each variable (here considered as environments) in a rather original approach, by the combination in a single model between ANOVA and the principal component analysis (PCA). The idea is to consider the effect of the GxE interaction as multiplicative component (more realistic in biological terms), and other effects (genotypes and environments) as purely additive effect components (Duarte and Vencovsky, 1999). Thus, the AMMI statistical model can be expressed as:

$$Y_{ij} = \mu + g_i + a_j + \sum_{k=1}^n \lambda_k \gamma_{ik} \alpha_{jk} + e_{ij}$$

Where:  $Y_{ij}$  is the mean response of the genotype  $i$  in the environment  $j$ ;  $\mu$  is the general mean of the trials;  $g_i$  is the fixed effect of genotype  $i$ ;  $a_j$  is the fixed effect of the environment  $j$ ;  $\lambda_k$  is the  $k$ -th singular value (scalar) of the original interaction array;  $\gamma_{ik}$  is the element corresponding to the  $i$ -th genotype in the  $k$ -th singular vector of the column of the interaction array;  $\alpha_{jk}$  is the element corresponding to the  $j$ -th environment in the  $k$ -th singular vector in the line of the array; and  $e_{ij}$  is the residual effect.

The AMMI methodology uses SVD multivariate technique to reduce the information contained in a data array  $n \times m$  (genotypes and environments, respectively) in vectors that accumulate, in a systematic manner (in order of importance), the greater part of variation contained in the data and, consequently, in GxE interaction. Since its disclosure, this methodology has been widely used for studies on adaptability in several important cultivated species such as wheat (Kempton, 1984; Crossa et al., 1999; Paderewski et al., 2011), Corn (Hirotsu, 1983; Ndhlela et al., 2014), soybeans (Gauch, 1988; Zobel et al., 1988; Yokomizo et al., 2013), sugar cane (Silveira et al., 2013), and rice (Samonte et al., 2005).

Another method based on the same principle that has

been increasingly used in recent years is the GGE Biplot, proposed by Yan et al. (2000), which, in general terms, is similar to the AMMI model, with the key difference that, in the multiplicative component for the decomposition via SVD, only the effect on the environment is excluded, consequently considering the effects of genotypes and of the interaction together. Therefore, models that consider the effect of the interaction as multiplicative, besides capitalizing the GxE interaction more efficiently (Zobel et al., 1988), have advantages, such as quantification of each genotype and environment to the sum of squares of the interaction, and provide an easy interpretation of results by *Biplot* graphs (Gabriel, 1971; Kempton, 1984).

Both the AMMI methodology and GGE Biplot have the additional advantages of generating information on the genotypes with broad adaptability (combination of phenotypic mean and stability information on the same graph), and aiding in the delineation of agronomic areas via identification of mega-environments (defined as the group of environments with similar GxE interaction standard, and consequently with little change in the ranking of the genotypes evaluated), which may indicate the most representative environments of each site and genotypes with specific adaptation to each region.

## METHODOLOGIES BASED ON MIXED MODELS

The usual hypothesis test of the variance analysis assumes independence of the main effects of the model; when such assumption is met, the effects can be tested using the mean square of the residue. Thus, within the context of the trials in multiple environments, any differences found between the effects of genotypes should, theoretically, be the same for any environment tested. However, if there is an interaction between the components of the model (in the specific case, between the effects of genotypes and environments), the hypothesis test is reformulated, leading then to decision making regarding the nature of such effects, that is, the *a priori* definition about which effects must be considered as fixed and which as random.

According to Freeman (1973), if the same set of genotypes is tested in several environments, the hypothesis test for the significance of the effects of genotypes must be performed in relation to the mean square of the interaction rather than the residue, as previously mentioned. When, for instance, such environments are considered as a random sample of all possible environments, it is assumed the use of a mixed model, or even of a completely random one, which means that, for instance, deviations from the normal range should be fully taken into consideration in order to assume the validity of the inferences made from the model.

By assuming the effects of genotypes as random, the BLUP's (*Best Linear Unbiased Predictors*) can be

obtained, which is not possible by the methods of adaptability and stability aforementioned. The BLUP's of the effects of genotypes and of GxE interaction eliminate their noises through the deliberation of such effects by a regressor factor, which is usually referred to as "repeatability" (which, in practical terms, is usually the character's heritability), leading, therefore, to the *Shrinkage* estimates of such effects and to the prediction of genetic values (Searle et al., 1992; Piepho, 1997; Resende, 2007).

Overall, regarding studies on plant genetics, the studies using mixed models were very scarce until the beginning of the last decade; however, their use in the evaluation of the most diverse cultures is increasing, in view of the advantages that this approach offers regarding difficulties (loss of experimental plots, heterogeneity of environmental variances, etc.) routinely found in agronomic experiments, especially in studies requiring many trials, such as the GxE interaction (Bastos et al., 2007; Carbonel et al., 2007; Mathews et al., 2007; Verardi et al., 2009; Borges et al., 2010; Mendes et al., 2012; Silva et al., 2012; Farias Neto et al., 2013; Gouvêa et al., 2013; Rodrigues et al., 2013; Gomez et al., 2014; Torricelli et al., 2014).

The basic model for the application of the methodology of mixed models was initially presented by Henderson (1973), and Resende (2007) defines it in a matricial way for the analysis of trials in the multi-environments, such as:  $Y = Xb + Zg + Tga + e$ . Where, according to this model, the relationship of the arrays in terms of means and variances is given by:

$$E \begin{bmatrix} y \\ g \\ ga \\ e \end{bmatrix} = \begin{bmatrix} Xb \\ 0 \\ 0 \\ 0 \end{bmatrix}; \text{Var} \begin{bmatrix} g \\ ga \\ e \end{bmatrix} = \begin{bmatrix} I\sigma_g^2 & 0 & 0 \\ 0 & I\sigma_{ga}^2 & 0 \\ 0 & 0 & I\sigma_e^2 \end{bmatrix}$$

Where:  $y$  is the phenotypic data vector,  $b$  is the vector of effects of the combination repetition-site (previously determined as fixed),  $g$  is the vector of genotypic effects (random),  $ga$  is the vector of the dual interaction GxE (consequently random), and  $e$  is the vector of residues (naturally random).  $X$ ,  $Z$ , and  $T$  are the incidence arrays of these effects, respectively. Thus, the prediction of genetic values via BLUP regarding a particular character, considering the effects of genotypes and GxE interaction as random, can be described by:

$$w_{ij} = y_{.j} + (\sigma_{gi}^2 + N\sigma_g^2 / \sigma_{gi}^2 + N\sigma_g^2 + \sigma^2)y_{.i} - y_{..} + (\sigma_{gi}^2 / \sigma_{gi}^2 + \sigma^2)y_{ij} - y_{.i} - y_{.j} + y_{..}$$

A simple statistic, based on the mixed modeling, is the one proposed by Resende (2007), which is a Harmonic Mean of the Relative Performance of Genotypic Values (HMRPGV):

$$\frac{n}{\sum_{j=1}^n 1/Vg_{ij}}$$

Where:  $n$  is the number of environments evaluated and  $Vg_{ij}$  is the genotypic value corresponding to genotype  $i$  in the environment  $j$ .

For the estimation of variance components, the *REML* (Maximum Residual Likelihood) method has been used, developed by Patterson and Thompson (1971), in which the values are estimated by the maximization of the likelihood function of residues instead of the observed data. Therefore, in trials that use the mixed models approach, especially in the case of unbalanced experiments, REML/BLUP analysis has been the most indicated (Resende and Thompson, 2004; Schaeffer, 2004).

## METHODOLOGIES USING COVARIATES

Eberhart and Russell (1966) highlight that although the genetic variability is notorious in terms of adaptability, it is difficult to explore it to its fullest, both because of the difficulty in evaluating (or even conceptualizing) adaptability, and the evident problem in quantifying the complexity of factors that influence natural environments.

The same authors state that the use of the general mean of the varieties in each environment shows that planting seasons (which cause differences mainly caused by unforeseen factors such as rainfall) are much more influential in the response of the varieties than the differences inherent in environments such as soil type. In addition, the breeders are inclined to disregard the importance of the results obtained in unfavorable environments, therefore leading to a successive loss of varieties which can justify huge adaptability.

Therefore, the use of an environmental index linked to means of varieties in each environment, such as those used by Finlay and Wilkinson (1963), Eberhart and Russell (1966), and Perkins and Jinks (1968), as the only factor of environmental information is not optimal, and the mathematical relationship between other environmental factors, such as rainfall, temperature, soil type, and the variable response, might be able to generate indexes less biased and more independent of the effect of variety within the analysis. Hardwick and Wood (1972) go further and claim that the fact that the deviations from the regression are not independent of the environmental mean also invalidates the use of the second parameter proposed by Eberhart and Russell (1966).

Freeman (1973) States that, in the face of great difficulty in efficiently capitalizing the GxE interactions to find which environments can maximize genotypes of interest, the use of other variables can be useful to find the factors that are behind the real difference between the genotypes. Freeman and Perkins (1971) reiterate that the use of a regression index would need to be based on measures independent of the environment, whether of physical or biological quality. Fripp and Caten (1971), therefore, through this type of approach, compared the

use of physical and biological variables and found that when the number of genotypes evaluated is large, this approach provides a value similar to that which uses the environmental mean.

However, Perkins (1972) found differences in genotype groups by the use of multiple regression based on climatic factors. Shukla (1972) and Wood (1976) used similar approaches, in which a correlation between a linear combination of genotypes and a linear combination of environmental factors were performed. According to Wood (1976), such an approach, when compared to others, provided a more logical explanation for the genotypic variation in different environments.

Overall, information about environmental measures are hardly available, but considering that the performance of a genotype can considerably vary from one environment to another, it is extremely important that the environmental cause of such change of behavior is measured in order to determine whether such differences can be due to factors inherent to climate or soil, or even due to management strategies. However, it has been noted that even when the experimental sites representing each region are fully selected, the management factors, soil characteristics, and climatic factors are usually not taken into consideration (Schlichting and Levin, 1986).

Some studies, such as those of Beckett (1982), aimed to quantify the environmental factors responsible for the interaction. This author performed a linear regression of each environmental variable in relation to productivity, aiming to identify the predominant factor and possibly the most influential one on the component of the interaction. Nevertheless, according to Weisberg (2005), when there are several factors in equal magnitude influencing the interaction, or when these factors present a certain degree of correlation between them, the simple linear regression analysis can be inappropriate.

From the 1980's, the use of environmental variables and the prediction of their influence on the productivity of some species have been widely applied in the studies on the GxE interaction, and currently several authors have been inserting environmental information, whether as characterization factors and environmental stratification as covariates in the analysis models of GxE interaction (Haun, 1982; Denis, 1988; Van Eeuwijk et al., 1996; Vargas et al., 1998; Crossa et al., 1999; Van Eeuwijk et al., 2005; Voltas et al., 2005; Thomason and Phillips, 2006; Vargas et al., 2006; Boer et al., 2007; Ramburan et al., 2011; Heslot et al., 2014).

Van Eeuwijk et al. (1996), in a seminal study, summarizes some methods based on factor analysis for the insertion of information about environmental covariates for the explanation of the GxE interaction, and, according to the author, such models are just an extension of the most general case:

$$Y_{ij} = \mu + \alpha_i + \beta_j + \rho_i z_j + e_{ij}$$

Where:  $\rho_i$  is a coefficient that reflects the sensitivity of the

genotype  $i$ , and  $z_j$  is the measure of the environmental variable  $z$  in the environment  $j$ . According to what was expressed, this strategy can be useful for the inclusion of a single environmental covariate such as "rainfall".

However, the idea of factorial regression can be generalized for the inclusion of other covariates, as follows:

$$Y_{ij} = \mu + \alpha_i + \beta_j + \rho_{i1}z_{j1} + \rho_{i2}z_{j2} \dots + \rho_{im}z_{jm} + E_{ij}$$

where  $\rho_{im}z_{jm}$  corresponds to the effect of a variable  $m$  in genotype  $i$  within the environment  $j$ . The successive addition of many environmental variables can reduce the accuracy of the prediction, considering that these variables may be modeling only the non-addictive part of the GxE interaction, that is, to the extent that more variables are added, the same can be inflated with the residue.

Thus, one can resort to the use of a reduction index of covariates of the model by the expression:  $\zeta_j = \sum_{h=1}^H \lambda_h z_{jh}$ , then becoming that which incorporates the synthetic covariate  $\lambda_h$ , with an initially unknown value, which is the more likely linear combination (via least-squares criterion) that can be generated from the available variables, therefore obtained via data set. The model becomes more thrifty (with reduced degrees of freedom) and can be written as:

$$Y_{ij} = \mu + \alpha_i + \beta_j + \rho_i(\sum_{h=1}^H \lambda_h z_{jh}) + E_{ij}$$

Where:  $H$  is the number of environmental covariates (Van Eeuwijk et al., 1996; Vargas et al., 1998; Crossa et al., 1999).

Some studies have been using explanatory covariates in the most variable way possible. Voltas et al. (2005) used the factorial regression and GGE Biplot methodologies for cultivation zoning and subsequent selection of superior genotypes, coupled with detection of main environmental factors that influenced the GxE interaction in 21 wheat genotypes evaluated in 8 environments. On the other hand, Yan and Tinker (2006), in a job evaluating 145 barley genotypes in 25 environments, use the combination of the two approaches aforementioned, by the integration of both into a single mathematical model. However, these authors only used genotype covariates (21 productivity components characters) for the explanation of the interaction regarding the productivity character.

Vargas et al. (2006) used more completed factorial regression models described by Van Eeuwijk et al. (1996) to decompose the GxE interaction effect on corn, with the aid of both genotype variables (*QTL*'s predicted via molecular markers) and environmental variables, estimating what the authors called *QTL* x environment interaction. Ramburan et al. (2011), studying varieties of sugar cane, used 14 environmental covariates (mean temperature per day, daily regimen of rain, mean daily

evaporation, soil moisture, among others) combined with the principal component analysis (PCA) to characterize the relative influence of each of the variables on the different environments. The authors then modeled the GxE interaction via the AMMI model, and verified through correlation analysis the relationship between the main components of such analysis with the most significant environmental variables.

According to Resende (2007), using regression methods, as well as their combination in multivariate models, are disadvantageous when there are experimental evidences of experimental unbalance factors or heterogeneity of variances between sites. Considering that such strategies take the effect of genotypes as fixed, their use becomes incoherent when you want to estimate variance components and other genetic parameters based on these experiments. Thus, only when a prediction of the genotypic values (as opposed to the use of phenotypic means) is made, actual values regarding the cultivation and use of a variety can be obtained.

## COMBINED APPROACHES

An advantageous approach can be the combination of multiplicative and mixed models. Piepho (1998), Resende and Thompson (2004), and Resende (2007), describe in detail the methods named as Analysis of Factors under Mixed Models (*FAMM*) and Principal Component Analysis under Mixed Models (*PCAM*). In the latter, rather than the data array with purely phenotypic values, values previously predicted are used considering random effects (both genotypes and environments, or both). Thus, for a PCA analysis under mixed models, you can adopt the equations relating to:

$$Y = Xb + Z(Q \otimes I_g)(Q^{-1} \otimes I_g)a + \bar{e}$$

Where:  $Q = V_m$  is the matrix of eigenvectors associated with  $m$  covariates.

More research is necessary before the complete use of environmental variables in the evaluation of adaptability and stability of genotypes (both in cultivated species and in natural populations), considering the question of how to properly analyze the environmental information is still not well established (Schlichting, 1986; Ramburan et al., 2011).

## RESULTS AND DISCUSSION

Although the development of methodologies specifically applied to the adaptability and stability studies on cultivated species only occurred in the last 50 years, considerable advances that have occurred in the fields of statistics (mainly regarding multivariate methods), computer science, as well as in the concept of GxE

interaction, enable a leap for such methodologies, which is quite important considering the current context, in which trials of evaluation of genotypes in a number of environments are carried out.

The methodologies based on linear regression only feature important limitations. Among them outstands the use of the mean of all the cultivars in each environment, such as environmental index, which may not occur in this case; the independence between variables, especially when the number of genotypes evaluated is small, which is a restriction of the use of regression, especially considering the current scenario where there is a need for an increasingly amount of tests for the adaptability and stability evaluation (Table 1).

The Ecovalence methodology has the advantage of easy interpretation, considering that it is based on the interpretation of a single numeric value ( $W_i$ ), however, it does not provide information about environments. In addition, the use of a single value to determine adaptability and stability can be difficult to apply to different objectives, such as the simultaneously broad recommendation and local recommendation of genotypes. In such situations, it may be preferable the use of methodologies, such as those based on multivariate analysis, that have the ability to reduce the data in order to provide a simple interpretation of the results. However, the methodologies based on linear regression can still be applied in a situation with smaller number of MET's.

Among the advantages of methodologies based on multivariate methods are the possibility of application of a biologically more realistic concept of GxE interactions, the ease of interpretation of results – provided by the use of Biplots charts – and the information level generated by the analysis (Table 1). To obtain information about each genotype evaluated and each environment is very useful to the breeders, since it enables a better separation of the concepts of adaptability (wide and local) and stability. Therefore, the use of methodologies such as AMMI and GGE Biplot is encouraged for most cases.

Nonparametric methods, such as those of Lin and Binns (1988), Hühn (1979), Nassar and Hühn (1987), Kang (1988), Fox (1990), and Thennarasu (1995), may be an alternative when the prerequisites of other methodologies are not understood (for instance, when the data do not clearly follow any probability distribution). However, such methodologies are less robust than the others because they are based on the ranking of genotypes only. In addition, the ambiguity caused sometimes by obtaining more than one ranking, such as the methods of Thennarasu ( $N_1$ ,  $N_2$ ,  $N_3$ , and  $N_4$ ) and Nassar and Hühn ( $S^1$ ,  $S^2$ ,  $S^3$ , and  $S^6$ ), can hinder the decision of plant breeders when such decision is based only on one of these methodologies. It is worth mentioning that the methods of Lin and Binns avoid this last factor, for it provides a single measure ( $P_i$ ) for the interpretation of adaptability and stability (Table 1).

Considering that the imbalance of data and the consequent loss of information are more common, or at least more likely to occur, in the current context, in which the number of environments and genotypes tested is increasing, methodologies that need to meet assumptions, such as the normality of data and absence of residual correlation, may not be the most suitable ones. For these cases, we recommend the use of more robust methodologies that include the use of mixed models, since these methodologies consider phenotypic values for obtaining predictors of real genotypic values. In the current context, in which the processing power of computers is huge, there are no more practical limitations to the application of this approach.

The inclusion of environmental variables in the adaptability and stability evaluation of genotypes is advantageous for dividing physical environments (sites) in generalized environmental factors. This type of approach can be advantageous for the ability to deal with unexpected environmental factors that often decisively influence the performance of genotypes. When including environmental effects separately, the problem of temporal variation of environments (variation between crops and years in the same site) is more elegantly approached. However, some care must be taken into consideration: Do the listed variables really affect the species in any meaningful way? How many environmental covariates must be used in order to obtain the most parsimonious model?. However, the main limitation for the use of such methodologies seems to be the difficulty in obtaining environmental data *in loco* during the exact period of performing experiments, whether by lack of interest of the plant breeder, whether by technological limitations.

The direct comparison between the several methodologies available is not an easy task, and is often inconsistent, mainly due to the fact that many are based on statistical principles quite distinct. A smart approach can be to base the choice of methodology according to the profile and characteristics of the data set to be analyzed. It is not very surprising that it is becoming costly to obtain a more detailed knowledge about the GxE interaction, considering that, in the context of evaluation of multi-environments trials, most of the efforts is usually focused on measuring the performance of the genotypes, while little or no attention is more accurately given to the evaluation of the environments. There is a need for better researches dedicated both to the study on the nature of the GxE interaction, and to the development of statistical genetic models able to comprise greater number of information related to genotypes and environments evaluated.

### Conflict of Interests

The authors have not declared any conflict of interest.

**Table 1.** Comparison between the main characteristics of the methodologies used the most for the evaluation of adaptability and stability in multi-environments trials.

Methodologies	Based Model	No. of parameters/measures	Advantages	Disadvantages/Limitations
Wricke (1962)	ANOVA	1 ( $\omega$ )	Simple interpretation; Simple calculations.	Data need to be balanced and meet the assumptions of a regression analysis.
Finlay and Wilkinson (1963)	Regression	1 ( $\beta_{ii}$ )	Simple interpretation; Simple calculations.	Assumes/requires that regression deviations are insignificant; Data need to be balanced and meet the assumptions of a linear regression analysis; Use of environmental index dependent on the mean of the evaluated genotypes.
Eberhart and Russell (1966)	Regression	2 ( $\beta_{ii}$ ; $\delta_{ij}$ )	Simple interpretation; Simple calculations.	Data need to be balanced and meet the assumptions of a linear regression analysis; Use of environmental index dependent on the mean of the evaluated genotypes.
Nassar and Hühn (1987)	Non-parametric	4 ( $S_i^{(1)}$ ; $S_i^{(2)}$ ; $S_i^{(3)}$ ; $S_i^{(6)}$ )	Simple calculations; Requires no assumptions.	Less robust model; absence of inferences about environments; 4 independent measures can hinder the conclusion.
Kang (1988)	Non-parametric	1 (Rank-sum)	Simple interpretation; Simple calculations; Requires no assumptions.	Less robust model; absence of inferences about environments.
Lin and Binns (1988)	Non-parametric	1 ( $P_i$ )	Simple interpretation; Simple calculations; Requires no assumptions.	Less robust model; absence of inferences about environments.
Fox et al. (1990)	Non-parametric	1 (TOP)	Simple interpretation; Simple calculations; Requires no assumptions.	Less robust model; absence of inferences about environments.
Thenarasu (1995)	Non-parametric	4 ( $NP^{(1)}$ , $NP^{(2)}$ , $NP^{(3)}$ , $NP^{(4)}$ )	Simple calculations; Requires no assumptions.	Less robust model; absence of inferences about environments; 4 independent measures can hinder the conclusion.
AMMI (Zobel et al., 1988; Gauch, 1992)	Additive/Multiplicative	---	Biologically realistic; Simple and Graphical interpretation; Inferences about environments and genotypes.	Limited when applied to unbalanced data or with significant residual correlations; Graphical interpretation is disadvantageous when the number of genotypes and environments is very large.
GGE Biplot (Yan et al., 2000)	Multiplicative	---	Biologically realistic; Simple and Graphical interpretation; Inferences about environments and genotypes.	Limited when applied to unbalanced data or with significant residual correlations; Graphical interpretation is disadvantageous when the number of genotypes and environments is very large.
HMRPGV (Mixed Models)	Mixed	1 (HMRPGV)	Simple interpretation; Applicable to unbalanced data; Tolerates residual correlation; Array of kinship can be inserted; Obtains BLUP predictors; Applicable to any size of data file.	Difficulties inherent to the degree of complexity of the approach that considers mixed models.
Covariate based methods	Regression; Multiplicative; Mixed	---	Works with information about types of environments and not with specific environments; Models better the influence of unforeseen factors; Offers the same advantages of Multiplicative and mixed Models, when allied to them.	Data collection for the inclusion of environmental and/or genotypic covariates in the model is still difficult; There are no precise criteria about which covariates should be chosen yet.

## REFERENCES

- Bastos IT, Barbosa MHP, Resende MDV, Peternelli LA, Silveira LCI, Donda LR, Figueiredo ICR (2007). Avaliação da interação genótipo x ambiente em cana-de-açúcar via modelos mistos. *Pesquis. Agropecuária Trop.* 37(4):195-203.
- Beckett JL (1982). Variety x environment interactions in sugar beet variety trials. *J. Agric. Sci.* 98(02):425-435.
- Boer MP, Wright D, Feng L, Podlich DW, Luo L, Cooper M, van Eeuwijk FA (2007). A mixed-model quantitative trait loci (QTL) analysis for multiple-environment trial data using environmental covariables for QTL-by-environment interactions, with an example in maize. *Genetics* 177(3):1801-1813.
- Borges V, Soares AA, Reis MS, Resende MD, Cornélio VMO, Leite NA, Vieira AR (2010). Desempenho genotípico de linhagens de arroz de terras altas utilizando metodologia de modelos mistos. *Bragantia* 69(4):833-841.
- Carbonel SAM, Chiorato AF, Resende MD, Dias LDS, Beraldo ALA, Perina EF (2007). Estabilidade de cultivares e linhagens de feijoeiro em diferentes ambientes no estado de São Paulo. *Bragantia* 66(2):193-201.
- Crossa J (1990). Statistical analyses of multilocation trials. *Adv. Agron.* 44:55-85.
- Crossa J, Vargas M, Van Eeuwijk FA, Jiang C, Edmeades GO, Hoisington D (1999). Interpreting genotype x environment interaction in tropical maize using linked molecular markers and environmental covariables. *Theor. Appl. Genet.* 99(3-4):611-625.
- Denis JB (1988). Two way analysis using covarites. *Statistics* 19(1):123-132.
- Dobzhansky T, Levene H (1955). Genetics of natural populations. XXIV. Developmental homeostasis in natural populations of *Drosophila pseudoobscura*. *Genetics* 40(6):797.
- Duarte JB, Vencovsky R (1999). Interação genótipos x ambientes: uma introdução à análise AMMI. Ribeirão Preto: Sociedade Brasileira de Genética. P 60.
- Eberhart ST, Russell WA (1966). Stability parameters for comparing varieties. *Crop Sci.* 6(1):36-40.
- Farias Neto JTD, Moura EF, Resende MDVD, Celestino Filho P, Augusto SG (2013). Genetic parameters and simultaneous selection for root yield, adaptability and stability of cassava genotypes. *Pesquis. Agropecuária Bras.* 48(12):1562-1568.
- Finlay KW, Wilkinson GN (1963). The analysis of adaptation in a plant-breeding programme. *Crop Pasture Sci.* 14(6):742-754.
- Fox PN, Skovmand B, Thompson BK, Braun HJ, Cormier R (1990). Yield and adaptation of hexaploid spring triticale. *Euphytica* 47:57-64.
- Freeman GH (1973). Statistical methods for the analysis of genotype-environment interactions. *Heredity* 31(3):339-354.
- Freeman GH, Perkins JM (1971). Environmental and genotype-environmental components of variability. VIII. *Heredity* 27:15-23.
- Fripp YJ, Caten CE (1971). Genotype-environmental interactions in *Schizophyllum commune*. I. Analysis and character. *Heredity* 27(3):393-407.
- Gabriel KR (1971). The biplot graphic display of matrices with application to principal component analysis. *Biometrika* 58(3):453-467.
- Gauch HG (1988). Model selection and validation for yield trials with interaction. *Biometrics* 44(3):705-715.
- Gauch HG (1992). Statistical analysis of regional yield trials: AMMI analysis of factorial designs. Elsevier Science Publishers.
- Gomez GM, Unêda-Trevisoli SH, Pinheiro JB, Di Mauro AO (2014). Adaptive and agronomic performances of soybean genotypes derived from different genealogies through the use of several analytical strategies. *Afr. J. Agric. Res.* 9(28):2146-2157.
- Gouvêa LRL, Silva GAP, Verardi CK, Oliveira ALB, Gonçalves ECP, Scaloppi-Junior EJ, Gonçalves P (2013). Rubber tree early selection for yield stability in time and among locations. *Euphytica* 191(3):365-373.
- Gripping B, Langridge J (1963). Phenotypic stability of growth in the self-fertilized species. *Arabidopsis thaliana*. In *Statistical Genetics and Plant Breeding: A Symposium and Workshop*. 982:368. National Academies.
- Hardwick RC, Wood JT (1972). Regression methods for studying genotype-environment interactions. *Heredity* 28(2):209-222.
- Haun JR (1982). Early prediction of corn yields from daily weather data and single predetermined seasonal constants. *Agric. Meteorol.* 27(3):191-207.
- Henderson CR (1973). Sire evaluation and genetic trends. *J. Anim. Sci. Symp.* pp. 10-41.
- Heslot N, Akdemir D, Sorrells ME, Jannink JL (2014). Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions. *Theor. Appl. Genet.* 127(2):463-480.
- Hirotsu C (1983). Defining the pattern of association in two-way contingency tables. *Biometrika* 70(3):579-589.
- Horner TW, Frey KJ (1957). Methods for determining natural areas for oat varietal recommendations. *Agron. J.* 49(6):313-315.
- Hühn VM (1979) Beitrage zur erfassung der phanotypischen stabilitat. *EDV Med. Biol.* 10:112-117.
- Kang MS (1988) A rank-sum method for selecting high-yielding, stable corn genotypes. *Cereal Res. Comm.* 16:113-115.
- Kempton RA (1984). The use of biplots in interpreting variety by environment interactions. *J. Agric. Sci.* 103(01):123-135.
- Lewis D (1954). Gene-environment interaction: A relationship between dominance, heterosis, phenotypic stability and variability. *Heredity* 8:333-356.
- Lin CS, Binns MR (1988). A superiority measure of cultivar performance for cultivar x location data. *Can. J. Plant Sci.* 68:193-198
- Mandel J (1971). A new analysis of variance model for non-additive data. *Technometrics* 13(1):1-18.
- Mathews KL, Chapman SC, Trethowan R, Pfeiffer W, Van Ginkel M, Crossa J, Cooper M (2007). Global adaptation patterns of Australian and CIMMYT spring bread wheat. *Theor. Appl. Genet.* 115(6):819-835.
- Mendes FF, Guimarães LJM, Souza JC, Guimarães PEO, Pacheco CAP, Machado JRDA, Parentoni SN (2012). Adaptability and stability of maize varieties using mixed model methodology. *Crop Breed. Appl. Biotechnol.* 12(2):111-117.
- Nassar R, Hühn M (1987). Studies on estimation of phenotypic stability: Tests of significance for nonparametric measures of phenotypic stability. *Biometrics* 43:45-53.
- Ndhlela T, Herselman L, Magorokosho C, Setimela P, Mutimaamba C, Labuschagne M (2014). Genotypex Environment Interaction of Maize Grain Yield Using AMMI Biplots. *Crop Sci.* 54(5):1992-1999.
- Paderewski J, Gauch HG, Mądry W, Drzazga T, Rodrigues PC (2011). Yield response of winter wheat to agro-ecological conditions using additive main effects and multiplicative interaction and cluster analysis. *Crop Sci.* 51(3):969-980.
- Perkins JM (1972). The principal component analysis of genotype-environmental interactions and physical measures of the environment. *Heredity* 29(1):51-70.
- Perkins JM, Jinks JL (1968). Environmental and genotype-environmental components of variability. 3. Multiple lines and crosses. *Heredity* 23(3):339-356.
- Piepho HP (1997). Analyzing genotype-environment data by mixed models with multiplicative terms. *Biometrics* pp. 761-766.
- Piepho HP (1998). Empirical best linear unbiased prediction in cultivar trials using factor-analytic variance-covariance structures. *Theor. Appl. Genet.* 97(1-2):195-201.
- Plaisted RL, Peterson LC (1959). A technique for evaluating the ability of selections to yield consistently in different locations or seasons. *Am. Potato J.* 36(11):381-385.
- Ramburan S, Zhou M, Labuschagne M (2011). Interpretation of genotypex environment interactions of sugarcane: Identifying significant environmental factors. *Field Crops Res.* 124(3):392-399.
- Resende MDV (2007). Matemática e estatística na análise de experimentos e no melhoramento genético. Colombo: Embrapa Florestas. P 362.
- Resende MDV, Thompson R (2004). Factor analytic multiplicative mixed models in the analysis of multiple experiments. *Rev. Mat. Estat.* 22(2):31-52.
- Rodrigues WP, Vieira HD, Barbosa DH, Souza Filho GR, Candido LS (2013). Adaptability and genotypic stability of *Coffea arabica* genotypes based on REML/BLUP analysis in Rio de Janeiro State, Brazil. *Genet. Mol. Res.* 12(3):2391-2399.

- Salmon SC (1951). Analysis of variance and long-time variety tests of wheat. *Agron. J.* 43(11):562-570.
- Samonte SOP, Wilson LT, McClung AM, Medley JC (2005). Targeting cultivars onto rice growing environments using AMMI and SREG GGE biplot analyses. *Crop Sci.* 45(6):2414-2424.
- Sandison A, Bartlett BO (1958). Comparison of varieties for yield. *J. Nat. Inst. Agric. Bot.* 8:351-357.
- Schaeffer LR (2004). Application of random regression models in animal breeding. *Livest. Prod. Sci.* 86(1):35-45.
- Schlichting CD (1986). The evolution of phenotypic plasticity in plants. *Ann. Rev. Ecol. Syst.* pp. 667-693.
- Schlichting CD, Levin DA (1986). Effects of inbreeding on phenotypic plasticity in cultivated *Phlox*. *Theor. Appl. Genet.* 72(1):114-119.
- Searle SR, Casella G, McCulloch CE (1992). Variance components. Hoboken: John Wiley & Sons. P 528.
- Shukla GK (1972). Some statistical aspects of partitioning genotype environmental components of variability. *Heredity* 29(2):237-245.
- Silva GO, Carvalho ADF, Vieira JV, Fritsche-Neto R (2012). Adaptabilidade e estabilidade de populações de cenoura. *Hortic. Bras.* 30(1):80-83.
- Silveira LCID, Kist V, Paula TOMD, Barbosa MHP, Peternelli LA, Daros E (2013). AMMI analysis to evaluate the adaptability and phenotypic stability of sugarcane genotypes. *Scientia Agricola* 70(1):27-32.
- Sprague GF, Federer WT (1951). A comparison of variance components in corn yield trials: II. Error, year x variety, location x variety, and variety components. *Agron. J.* 43(11):535-541.
- Thennarasu K (1995). On certain non-parametric procedures for studying genotype-environment interactions and yield stability. Ph.D. thesis. P.J. School, IARI, New Delhi, India.
- Thomason WE, Phillips SB (2006). Methods to evaluate wheat cultivar testing environments and improve cultivar selection protocols. *Field Crops Res.* 99(2):87-95.
- Torricelli R, Ciancaleoni S, Negri V (2014). Performance and stability of homogeneous and heterogeneous broccoli (*Brassica oleracea* L. var. *italica* Plenck) varieties in organic and low-input conditions. *Euphytica* 199(3):385-395.
- Van Eeuwijk FA, Denis JB, Kang MS (1996). Incorporating additional information on genotypes and environments in models for two-way genotype by environment tables. In: Kang MS, Gauch HG (eds). *Genotype-by-environment interaction*. Boca Raton: CRC Press. pp. 15-49.
- Van Eeuwijk FA, Malosetti M, Yin X, Struik PC, Stam P (2005). Statistical models for genotype by environment data: From conventional ANOVA models to eco-physiological QTL models. *Crop Pasture Sci.* 56(9):883-894.
- Vargas M, Crossa J, Sayre K, Reynolds M, Ramírez ME, Talbot M (1998). Interpreting Genotype x Environment Interaction in Wheat by Partial Least Squares Regression. *Crop Sci.* 38(3):679-689.
- Vargas M, Van Eeuwijk FA, Crossa J, Ribaut JM (2006). Mapping QTLs and QTLx environment interaction for CIMMYT maize drought stress program using factorial regression and partial least squares methods. *Theor. Appl. Genet.* 112(6):1009-1023.
- Verardi CK, Resende MD, Costa RD, Gonçalves PDS (2009). Adaptabilidade e estabilidade da produção de borracha e seleção em progênies de seringueira. *Pesquis. Agropecuária Bras.* 44(10):1277-1282.
- Voltas J, Lopez-Corcoles H, Borrás G (2005). Use of biplot analysis and factorial regression for the investigation of superior genotypes in multi-environment trials. *Eur. J. Agron.* 22(3):309-324.
- Weisberg S (2005). *Applied Linear Regression*. Wiley Series in Probability and Mathematical Statistics. Hoboken: John Wiley & Sons. P 310.
- Williams W (1960). Relative variability of inbred lines and F1 hybrids in *Lycopersicum esculentum*. *Genetics* 45(11):1457.
- Witcombe JR, Whittington WJ (1971). Study of the genotype by environment interaction shown by germinating seeds of *Brassica napus*. *Heredity* 26:37-41.
- Wood JT (1976). The use of environmental variables in the interpretation of genotype-environment interaction. *Heredity* 37(1):1.
- Wricke G (1962). On a method of understanding the biological diversity in field research. *Z. Pfl.-Zücht* 47:92-146.
- Yan W, Hunt LA, Sheng Q, Szlavnics Z (2000). Cultivar evaluation and mega-environment investigation based on the GGE biplot. *Crop Sci.* 40(3):597-605.
- Yan W, Tinker NA. (2006). Biplot analysis of multi-environment trial data: Principles and applications. *Can. J. Plant Sci.* 86(3):623-645.
- Yates F, Cochran WG (1938). The analysis of groups of experiments. *J. Agric. Sci.* 28(04):556-580.
- Yokomizo GK, Duarte JB, Vello NA, Unfried JR (2013). Análise AMMI da produtividade de grãos em linhagens de soja selecionadas para resistência à ferrugem asiática. *Pesquis. Agropecuária Bras.* 48(10):1372-1380.
- Zobel RW, Wright MJ, Gauch HG (1988). Statistical analysis of a yield trial. *Agron. J.* 80(3):388-393.