STATISTICAL MODELS IN AGRICULTURE: BIOMETRICAL METHODS FOR EVALUATING PHENOTYPIC STABILITY IN PLANT BREEDING

Daniel Furtado Ferreira[1], Clarice Garcia Borges Demétrio[2], Bryan Frederick John Manly[3],
Amauri de Almeida Machado[4], Roland Vencovsky[5]

ABSTRACT: This paper reviews the main concepts of several methods of phenotypic stability analysis and points out their advantages and limitations. It was concluded that the simple linear regression method of Eberhart & Russel (1966) and the bi segmented regression method of Silva & Barreto (1985) have only historical importance nowadays. Moreover, based on factors discussed in the paper, it is recommended that the regression models of Toler & Burrows (1998) and the Additive Main effects and Multiplicative Interactions (AMMI) model should be used simultaneously to study and to estimate phenotypic stability effects.

Key words: AMMI, non-linear statistical models, linear regression models.

MODELOS ESTATÍSTICOS NA AGRICULTURA: MÉTODOS BIOMÉTRICOS PARA AVALIAR A ESTABILIDADE FENOTÍPICA NO MELHORAMENTO DE PLANTAS

RESUMO: Este trabalho revisa os principais conceitos de diversos métodos de análise da estabilidade fenotípica e aponta suas vantagens e limitações. Concluiu-se que os métodos de regressão linear simples de Eberhart & Russel (1966) e o de regressão bi-segmentada de Silva & Barreto (1985) têm apenas importância histórica nos dias atuais. Além disso, considerando-se os aspectos discutidos, recomenda-se que os modelos de regressão de Toler & Burrows (1998) e o modelo com efeitos principais aditivos e interação multiplicativa (AMMI) sejam usados simultaneamente para estudar e estimar os efeitos da estabilidade fenotípica.

Palavras-chave: AMMI, modelos estatísticos não-lineares, modelos lineares de regressão.

1 INTRODUCTION

In the last phase of plant breeding programs, candidate varieties with market potential should be evaluated under a range of conditions similar to the real conditions that they will experience when in use. To be successful, a new variety should have high productivity and high performance for agronomic traits over a wide range of environmental conditions. Plant breeders usually agree about the importance of high production stability, but not necessarily about the appropriate definition of stability.

The basic cause of differences among genotypes (varieties) in relation to production stabilities is the genotype x environment (GE) interaction, so that the performance of the genotypes depends on the specific environmental conditions where they are grown. Part of the interaction can be explained by known environmental factors such as the incidence of disease or pests, annual rainfall, severity of dry-season, soil fertility, soil waterlogging, soil depth, photoperiod, etc. However, the most of GE interactions cannot be explained by these known factors.

Several methods have been developed by statisticians and applied by plant breeders to explain the GE interaction at the end of plant breeding programs. The main concepts of these methods are reviewed in this article and their advantages and limitations are pointed out.

2 BASIC CONCEPTS

In the last phase of plant breeding the new varieties are grown in several locations under different conditions of climate and soil fertility, and also in different seasons of the year (ACCIARESI & CHIDICHIMO, 1999; BECKER & LÉON, 1988). The different conditions, as defined by

[1] Professor do Departamento de Ciências Exatas – Universidade Federal de Lavras/UFLA – Cx. P. 3037 – 37200 000 – Lavras, MG – danielff@ufla.br – CNPq Researcher.
[2] Professora da ESALQ/USP – Av: Pádua Dias, 11 – Cx. P. 09 – 13418- 900 – Piracicaba, SP – clarice@carpa.ciagri.usp.br – CNPq Researcher.
[3] CAPES visiting Professor ESALQ/USP – bmanly@compuserve.com.
[4] Professor do Departamento de Matemática e Estatística da UFPEL – Pos doctorate in statistics – ESALQ/USP – amachado@ufpel.edu.br
[5] Professor da ESALQ/USP – Av: Pádua Dias, 11 – Cx. P. 09 – 13418-900 – Piracicaba, SP.

locations and seasons, are considered to be a single factor for environmental conditions. In these experiments, randomized complete blocks and incomplete block designs are usually used. For incomplete block experiments lattices designs are usually used due to the large number of varieties (genotypes) to test. For all designs, the interest of plant breeders has focused on modelling the genotype means estimated in the jth environment. Thus, one may consider the linear model:

$$Y_{ij} = \mu + g_i + e_j + ge_{ij} + \bar{\varepsilon}_{ij} \qquad (2.1)$$

where: Y_{ij} is the ith genotype mean observed in the jth environment, for $i = 1, 2, ..., p$, and $j = 1, 2, ..., q$; μ is an overall constant; g_i is the fixed effect of the ith genotype,

with ; $\displaystyle\sum_{i=1}^{p} g_i = 0$; e_j is the fixed effect of the jth

environment, with ; $\displaystyle\sum_{j=1}^{q} e_j = 0$; ge_{ij} is the effect of

interaction between the ith genotype and jth environment,

with $\displaystyle\sum_{i=1}^{p} ge_{ij} = \sum_{j=1}^{q} ge_{ij} = 0$; $\bar{\varepsilon}_{ij}$ is the mean error related to

the observed Y_{ij}, which is assumed to be normal with a mean 0 and a variance, σ^2 / n , where "n" is the number of replications associated to the yield mean Y_{ij}; p is the number of genotypes; and q is the number of environments.

　　　In Table 1 the general form of the results observed from the evaluation of p genotypes in q environments is presented. Thus, $\bar{Y}_{i\bullet}$ represents the marginal genotypes means, $\bar{Y}_{\bullet j}$ the marginal environments means, and $\bar{Y}_{\bullet\bullet}$ the overall mean.

　　　The GE interaction term ge_{ij} (in equation 2.1) represents the differential genotype yield responses under different environmental stimuli. In most situations, the relative performances of two genotypes change with the environment conditions, as a direct consequence of GE interaction. Therefore, one of the most important objectives of the analysis of phenotypic stability is to identify the genotypes whose phenotypic performance remains stable even when the environmental conditions change. These analyses only make sense if GE interactions are present (HUSSEIN et al., 2000).

Table 1 – Estimates of the p genotype means in each of the q environments.

Tabela 1 – *Estimativas das p medias genotípicas em cada um dos q ambientes.*

| Genotypes	Environments				Marginal Mean $\left(\bar{Y}_{i\bullet}\right)$
	1	*2*	$\cdots$	*q*	
1	Y_{11}	Y_{12}	$\cdots$	Y_{1q}	$\bar{Y}_{1\bullet}$
2	Y_{21}	Y_{22}	$\cdots$	Y_{2q}	$\bar{Y}_{2\bullet}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
p	Y_{p1}	Y_{p2}	$\cdots$	Y_{pq}	$\bar{Y}_{p\bullet}$
Mean $\left(\bar{Y}_{\bullet j}\right)$	$\bar{Y}_{\bullet 1}$	$\bar{Y}_{\bullet 2}$	$\cdots$	$\bar{Y}_{\bullet q}$	$\bar{Y}_{\bullet\bullet}$

　　　In Figure 1, the mean performances of two genotypes (A and B) are shown in two environments (E1 and E2) to illustrate the environmental effect, the presence and the absence of an interaction effect, and the two basic types of interaction. Figure 1 (a) and (b) show the absence of an interaction effect. In these cases the genotype lines are parallel, with genotype A showing a higher response than genotype B, and absence of environmental effect in 1 (b). Figure 1 (c) and (d) show interaction effects. In 1 (c) the interaction is of a simple type, with genotype B superior for both environments E1 and E2. In 1 (d) the interaction is complex. This is the most important case for the plant breeder, because genotype A has the lowest mean in environment E1, but the highest mean in environment E2. Most real situations show a mixture of cases 1 (a) to 1 (d).

　　　Phenotypic stability has two concepts, static and dynamic (BECKER & LEON, 1988). The static phenotypic stability exists when a genotype maintains its performance independently of variations in the environmental conditions. This type is called biological stability. A genotype has dynamic stability if its performance varies with environmental changes but in a predictable way. This kind of stability is called agronomic stability.

　　　There are several methods to measure phenotypic stability by modelling the GE interaction. These include

　　　a) Univariate parametric methods including simple and bi-segmented linear regression, variance components with mixed models, descriptive statistics, and non-linear regression models;
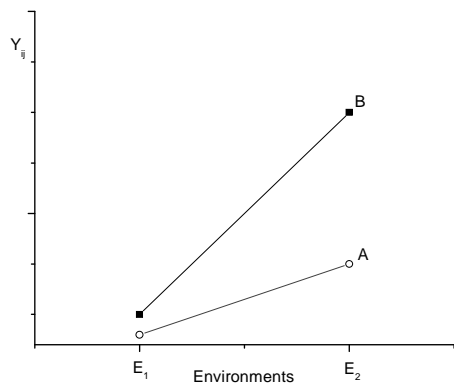
　　　b) Non-parametric methods including variance of genotype rank values; and
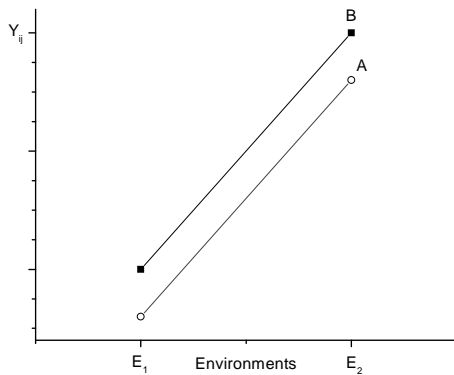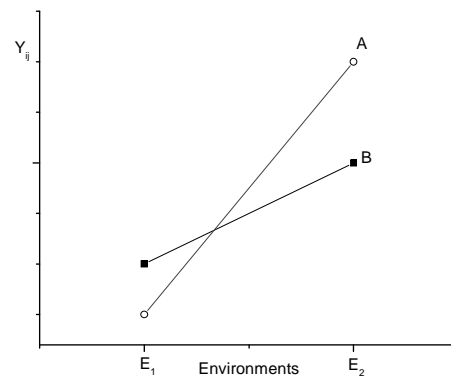
(a) Illustration of null Genotype x Environmental (GE) interaction, genotype and environmental effects.
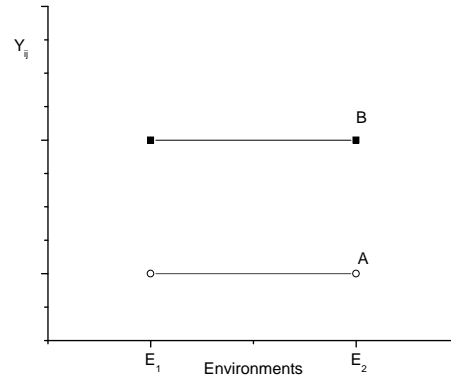
(b) Illustration of null GE interaction, environmental and genotype effects.

(c) Illustration of simple GE effect

(d) Illustration of complex GE effect

Figure 1 − Patterns of the genotype behaviors in different environments outlining the two fundamental types of interaction, considering only a simple case of two genotypes and two environments.

Figura 1 − Padrões de comportamento dos genótipos em diferentes ambientes destacando os dois tipos fundamentais de interação, considerando somente um caso simples de dois genótipos e dois ambientes.

c) Multivariate methods including AMMI (Additive Main effects and Multiplicative Interactions) analysis.

3 METHODS FOR ESTIMATING PHENOTYPIC STABILITY

3.1 Methods based on analysis of variance

All the methods described in this section are based on the model of equation (2.1). For the specific case of measuring phenotypic stability by using the static concept, Roemer (1917) proposed the use of the variance of each genotype over the environments. Thus, to estimate the static phenotype stability of the ith genotype, the following estimator could be used:

$$S_i^2 = \frac{\sum_{j=1}^{q}(Y_{ij} - \bar{Y}_{i\bullet})^2}{q-1} \tag{3.1}$$

A genotype is then considered to be stable if the sample estimate is not significantly different from zero, i.e., if the hypothesis H_0: $\sigma_i^2 = 0$ is not rejected, which means that this genotype will not have yield changes with changes in the environment. A problem with this method is that, in general, genotypes with high phenotypic stability

measured through the environmental variance show low yield. In consequence, plant breeders do not use this method to evaluate the phenotypic stability of the genotype yields, or other related random variables. However, it is useful to evaluate the phenotypic stability of traits that should maintain their levels. Among these are qualitative traits such as resistance to diseases or tolerance to environmental stresses.

The simplest method to evaluate the stability by using the dynamic concept is due to Wricke (1964). In this method, it is proposed to use the sum of square of the GE populational effects $\omega_i = \sum_{j=1}^{q} g_{ij}^2$ and a sample estimator is given by:

$$W_i = \sum_{j=1}^{q} \left(Y_{ij} - \overline{Y}_{i\bullet} - \overline{Y}_{\bullet j} + \overline{Y}_{\bullet\bullet} \right)^2 \qquad (3.2)$$

This statistics for the ith genotype represents the sum of squares of the GE estimate interaction effects. If $\omega_i = 0$, the genotype is considered stable and if it is greater than 0 the genotype is considered unstable. Wricke (1964) called this parameter ecovalency, and referred to it as genotype ability to answer to environmental changes. A high ecovalency implies a low ω_i.

The variance component of each genotype throughout the environments is another related measure of phenotypic stability, as proposed by Shukla (1972). This genotype variance estimator is given by:

$$\hat{\sigma}_i^2 = p / \left[(q - 1)(p - 1) \right] W_i - QMGE / (p - 2) \quad (3.3)$$

where $QMGE$ is the genotype x environment interaction mean square. In this case, a drawback is the possibility of obtaining a negative estimate of the genotype variance component.

The procedure to obtain geometrically the genotype ecovalency estimator W_i is shown Figure 2. The dotted line represents the genotype response obtained by adding the overall constant to the environmental effect. The solid line represents the same genotype responses, added to the ith genotype effect. The difference between the two lines represents the genotype effect and the deviations of the solid line from the true genotype values (points) represent the GE interaction effects. The sum of the squares of the interaction effects represents the ecovalency. Stable genotypes would be those whose responses are all close to the full straight line.
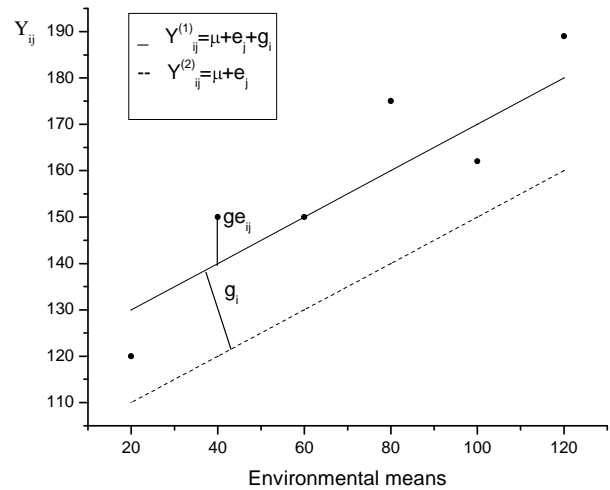
Figure 2 – Parametric representation of the responses (points) from a selected genotype and of two ways for modelling them by showing the genotypes and GE interaction effects.

Figura 2 – Representação paramétrica da resposta (pontos) de um genótipo específico e de duas formas para modelá-los, destacando os efeitos genotípicos e da interação.

3.2 Regression methods

Although there is no unanimous concept for the phenotypic stability, several methods and models have been used to explain the GE interaction. These include the use of uni-segmented and bi-segmented linear regression (CRUZ et al., 1989; EBERHART & RUSSELL, 1966; FINLAY & WILKINSON, 1963; FREEMAN & PERKINS, 1971; PERKINS & JINKS, 1968a, b; SHUKLA, 1972; SILVA, 1995a, b, 1998; SILVA & BARRETO, 1985; TOLER, 1990; TOLER & BURROWS, 1998; VERMA et al., 1978), where the basic idea consists of regressing the genotypes performances on the environmental mean yields, expressed by an environmental index, through a linear or a non-linear model in the parameters.

These regression models have been used most often by plant breeders since the first simple linear regression model proposed by Finlay & Wilkson (1963), and much research focused on searching for improvements in these methods.

The simple linear regression model (EBERHART & RUSSELL, 1966; FINLAY & WILKINSON, 1963) is defined by:

$$Y_{ij} = \beta_{0i} + \beta_{1i}\overline{Y}_{\bullet j} + \delta_{ij} + \overline{\varepsilon}_{ij}$$

where β_{0i}, β_{1i}, δ_{ij} and $\overline{Y}_{\bullet j}$ are the regression coefficients (intercept and slope), the deviation from the regression and the jth environmental mean, respectively. Usually $\overline{Y}_{\bullet j}$ is replaced by , $I_j = \overline{Y}_{\bullet j} - \overline{Y}_{\bullet\bullet}$ to give

$$Y_{ij} = \beta_{0i} + \beta_{1i}I_j + \delta_{ij} + \overline{\varepsilon}_{ij} \qquad (3.4)$$

This regression considers that the effects in the complete biometrical model (2.1) are independent. If the genotypes were considered simultaneously, there would be no covariance among the GE interaction and environmental effects. If, on the other hand, the genotypes were considered separately, there would be a covariance and the regression coefficient will be the standardized version of that covariance (COMSTOCK & MOLL, 1963).

The interaction effect can be modeled by adopting an equivalent model to that presented in equation (3.4) by

$$ge_{ij} = \phi_i e_j + \delta_{ij} \qquad (3.5)$$

where ϕ_i represents the regression coefficient of the environmental effects on the GE effects for the ith genotype; δ_{ij} is the regression deviate for the ith genotype and for the jth environment (Figure 3). The genotype response fits better when we choose model (3.4), or equally with model (3.5), than when the following model is chosen:

$$Y_{ij} = \mu + e_j + g_i$$

The residual mean square of each genotype is also used to classify a genotype according to it stability. The residual mean square $\left(\sigma_{di}^2\right)$ for the regression with the ith genotype is estimated by S_{di}^2 and a lack of fit test (DRAPER & SMITH, 1998), $H_0 : \sigma_{di}^2 = 0$, can be applied. Many authors use this test to define whether the ith genotype is stable or not. If that hypothesis is not rejected, then the ith genotype is considered stable (EBERHART & RUSSELL, 1966).

The linear regression models have received criticism from the scientific community (BECKER & LEON, 1998; TOLER & BURROWS, 1998), but they are nevertheless often used. The main criticism is based on the fact that the environmental index (I_j) is not independent of the response variable (Y_{ij}), because the index is estimated as the environmental mean response, although the effect of this dependence decreases as the number of genotypes increases. A second criticism is related to the use of
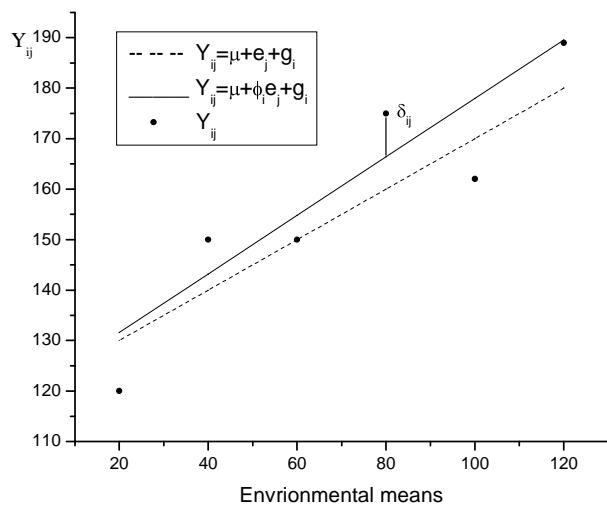

Figure 3 – Modelling the genotype yield by a simple additive linear model and by a regression model on environmental parametric effects.

Figura 3 – Modelagem da produtividade genotípica por meio de um modelo linear aditivo simples e de um modelo de regressão sobre os efeitos ambientais paramétricos.

biased estimators of the regression coefficients, because the independent variable is measured with error (STORCK & VENCOVSKY, 1994). This bias depends on p and on the ratio σ_e^2 / σ^2, where σ_e^2 is the environmental variance and σ^2 is the residual variance. Thus, the bias reduces as the environmental effects increase. Another problem that can affect the inference is the violation of the assumption of the residual environmental variances homogeneity.

The genotypes are classified according to their performance based on the regression coefficients. The genotypes response patterns depend on the test of the hypothesis $H_0 : \beta_{1i} = 1$, and the genotypes are classified as:
- responsive genotype to improvements of the environmental conditions, $(\beta_{1i} > 1)$ recommended for favorable environments;
- less responsive genotype to improvements of the environmental conditions $(\beta_{1i} < 1)$, recommended to unfavorable environments;
- broad stable genotype $(\beta_{1i} = 1)$, recommended for most environmental conditions.

With those tests, should come a lack of fit test $\left(H_0 : \sigma_{di}^2 = 0\right)$, and a multiple comparison test for the

genotype means (β_{0i}). This provides a general protocol to evaluate the response pattern over environments and the phenotypic stability and adaptability of each genotype. An adapted genotype shows a high mean performance over all environments. The favorable environments are those where farmers adopt agricultural practices with high technology, or have high soil fertility and appropriate climate conditions. Many plant breeders cannot understand the reason why stable and responsive genotypes are only recommended for favorable environments. To answer this question, it is necessary to observe that a larger β_{1i} value will potentially provide lower yields as the environmental conditions worsen. Therefore, the risk of a poor performance is high when a genotype that is classified as responsive is chosen for unfavorable environments. This can exclude such genotype from the market. In the same way a genotype with $\beta_{1i} < 1$, classified as not responsive, should not be recommended for favorable environments because this genotype does not respond very positively to improved environmental conditions.

The basic differences between the methods proposed by Eberhart & Russel (1966) and Finlay & Wilkinson (1963) are that the former suggested a logarithmic transformation of the data, while the second used the regression parameters and deviations to evaluate the stability, adaptability and the pattern of genotype response to the environment.

A criticism of the use of simple linear regression models is based on the potential non linear pattern of genotype responses to environmental variation. The first proposal to solve this deficiency was presented by Verma et al. (1978). They separated the environments into two groups (favorable and unfavorable) and fit a simple linear regression model separately to each part. The division is made based on the environmental index $I_j = \overline{X}_{\bullet j} - \overline{X}_{\bullet\bullet}$ that represents the deviation of each environmental mean from the overall mean. They considered as unfavorable environments those with negative or zero indices and as favorable environments those with positive indices. It should be noted that this classification of the environments is rather simple and may not be satisfactory when there are few genotypes.

There is also a problem with this approach because of sampling errors in the classification of environments,

when the number of environments is small. For this reason, Silva & Barreto (1985) proposed a bi-segmented linear regression model, where each segment is a straight line. Cruz et al. (1989) modified this approach. Their model is discontinuous in the junction of the two straight line segments from the unfavorable to favorable environments. They adopted this modification because of the existence of negative residual correlation among the estimators of the regression coefficients of the favorable and unfavorable environments.

This model is based on the fact that an ideal genotype (Figure 4) possesses:

• high yield performance;

• high stability with a good fitting model, i.e., $\sigma^2_{di} = 0$;

• low sensitivity to adverse conditions; and

• is capable of responding positively when environmental conditions are improved.

The ideal genotype has a regression coefficient smaller than 1 for unfavorable environments and greater than 1 for favorable environments. On the other hand, an undesirable genotype shows regression coefficient greater than 1 for unfavorable environments and smaller than 1 for a favorable environments.
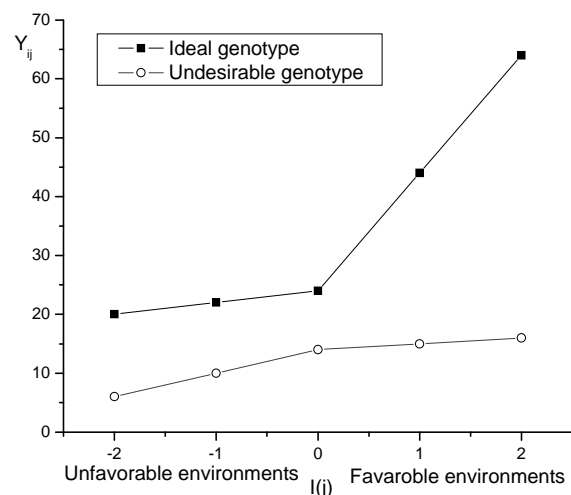

Figure 4 – Patterns for the bi-segmented linear model of Silva & Barreto (1985) showing the favorable and unfavorable environments and desirable (concave) and undesirable (convex) genotypes.

Figura 4 – Padrões para o modelo bi-segmentados de Silva and Barreto (1985), mostrando os ambientes favoráveis e desfavoráveis e os genótipos desejáveis (côncavo) e indesejáveis (convexo).

The models of Cruz et al. (1989) and of Silva & Barreto (1985) are represented simultaneously in the expression:

$$Y_{ij} = \beta_{0i} + \beta_{1i} I_j + \beta_{2i} T(I_j) + \delta_{ij} + \overline{\varepsilon}_{ij} \quad (3.6)$$

where the definitions of Y_{ij}, I_j, δ_{ij} and β_{0i} are the same as presented in the simple linear model; β_{1i} represents the linear regression coefficient related to the unfavorable environments; and $\beta_{1i} + \beta_{2i}$ represents the slope of the linear response to the favorable environments. Finally, we define $T(I_j)$ as: $T(I_j) = 0$ if $I_j \le 0$, or $T(I_j) = I_j - \overline{I}_+$ if $I_j > 0$ for the model of Cruz et al. (1989); and $T(I_j) = 0$ if $I_j \le 0$ or $T(I_j) = I_j$ if $I_j > 0$ for the model of Silva & Barreto (1985).

where $\overline{I}_+ = \sum_{j=1}^{q} I_j \Gamma(I_j > 0) \Big/ \sum_{j=1}^{q} \Gamma(I_j > 0)$ is the mean of positive environmental indexes I_j and $\Gamma(I_j > 0)$ an indicator function.

The ordinary least square parameter estimators, the tests of the hypotheses of interest $H_0 : \beta_{1i} = 0$, $H_0 : \beta_{1i} = 1$ and $H_0 : \beta_{1i} + \beta_{2i} = 1$ and the lack of fit test were described by Cruz & Regazzi (1994).

Based on the results of the hypotheses tests, the ideal genotype can be defined as being one that possesses:
- high yield performance - high β_0;
- $\beta_1 < 1$, i.e., low response in unfavorable environments;
- $\beta_1 + \beta_2 > 1$, i.e., high response in favorable environments; and
- highly stable behavior, i.e., $\sigma_d^2 = 0$.

As with the simple linear regression model, the criticism that the environmental index is not independent of the response variable (Y_{ij}) still remains. With regards of the bias on the regression coefficients, Storck & Vencovsky (1994) proposed corrections to the estimators and to the hypothesis tests. Nevertheless, if the number of environments is small, the model is inappropriate. Inferences may also be affected by the environmental errors variances heterogeneity.

A new method using bi-segmented models was proposed by Silva (1998). In the original bi-segmented model, the junction of the straight line segments associated with the favorable and unfavorable environments occurs for all of genotypes at the zero value for the environmental index. However, according to Silva (1998) the genotypes can have different junction points and this can be used to distinguish among them. The model proposed by Silva (1998) was

$$Y_{ij} = \beta_{0i} + \beta_{1i} I_j + (\beta_{2i} - \beta_{1i})(I_j - X_{0i}) \Gamma(I_j - X_{0i}) + \delta_{ij} + \overline{\varepsilon}_{ij} \quad (3.7)$$

where: $\Gamma(I_j - X_{0i}) = 0$ if $I_j \le X_{0i}$ or $\Gamma(I_j - X_{0i}) = 1$ if $I_j > X_{0i}$.

The innovation is the introduction of the parameter X_{0i} (the point of junction of the two straight line segments) in the model. This is non-linear in the parameters but can be fitted by non linear estimation methods, such as the modified Gauss-Newton method (GALLANT, 1987). However, Silva (1998) proposed a simpler trial and error estimation method based on ordinary least squares. By using a fixed value of the parameter X_{0i}, a model using ordinary regression methods is used. The value of the parameter X_{0i} was then varied and this process repeated. The final model is chosen for each genotype as the one that maximizes the coefficient of determination R^2 in the regression (Figure 5).

For this model, larger values of X_{0i} indicate lower risks under adverse environmental conditions, but the genotypes are less responsive to improvements in the environmental conditions. On the other hand, the smaller

values of X_{0i} indicate higher risks but more responsive genotypes to improvements in environmental conditions.

With Silva's (1998) regression model, the criticism of a lack of independence between the dependent and independent variables still remains. To solve this problem, Toler (1990) and Toler & Burrows (1998) proposed simple uni- and bi-segmented models, where the environmental index is a parameter to be estimated. These models use indicator variables for the favorable and unfavorable environments and are unusual because the independent variable is not observed. These models are:

$$Y_{ij} = \beta_{0i} + \beta_{1i} \mu_j + \delta_{ij} + \overline{\varepsilon}_{ij} \quad (3.8)$$

$$Y_{ij} = \beta_{0i} + \left[Z_j \beta_{1i} + (1 - Z_j) \beta_{2i} \right] \mu_j + \delta_{ij} + \overline{\varepsilon}_{ij} \quad (3.9)$$
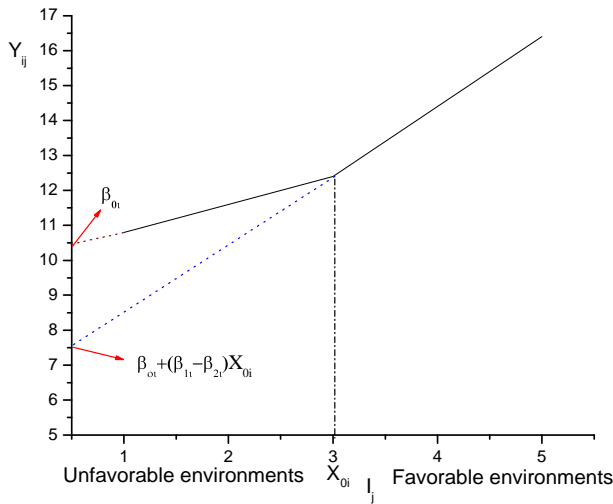
Figure 5 – Representation of Silva's (1998) bi-segmented model, showing the junction parameter for the two straight line segments (X_{0i}) corresponding to the unfavorable and favorable environments.

Figura 5 – Representação do modelo bi-segmentado de Silva (1998), mostrando o parâmetro de junção dos dois segmentos de linha (X₀ᵢ) correspondente aos ambientes desfavoráveis e favoráveis.

where Z_j is an indicator variable that is 1 if $\mu_j \leq 1$, and zero otherwise; and μ_j is the effect of the jth environment.

Restrictions on the parameters are necessary to allow the model to be fitted to data. These restrictions are:

$$\sum_{i=1}^{p} \beta_{1i} = p \quad \text{and} \quad \sum_{j=1}^{q} \mu_j = 0$$

for model (3.8) and

$$\sum_{i=1}^{p} \beta_{1i} = \sum_{i=1}^{p} \beta_{2i} = p \quad \text{and} \quad \sum_{j=1}^{q} \mu_j = 0$$

for model (3.9) .

Essentially, these models differ from the previous one because the environmental index μ_j is now a parameter of the model. Hence, least square methods for non-linear models must be used to estimate the parameters.

Toler & Burrows (1998) suggested a protocol for determining the genotype response patterns over environments. Initially, the plant breeder should fit the bi-segmented model (3.9) and apply Student *t* test for the

hypothesis $H_0 : \beta_{1i} = \beta_{2i}$. If this hypothesis is not rejected, the simple non-linear regression model (3.8) should be applied. Then, the hypothesis that the common regression coefficient of the ith genotype is equal to 1, $H_0 : \beta_i = 1$ should be tested. The genotype is then classified according to its response pattern. If the regression coefficient were larger than 1, the genotype is responsive, if it is equal to 1, the genotype has a wide range of adaptation, and if it is smaller than 1, the genotype is non-responsive.

On the other hand, if the hypothesis of equality of regression coefficients of the two straight lines segments on the bi-segmented regression model is rejected then the hypotheses $H_0 : \beta_{1i} = 1$ and $H_0 : \beta_{2i} = 1$ should be tested. In this case, the genotypes can be classified in a similar way as discussed for the uni- and bi-segmented linear models. Hence, the ideal is a genotype that shows a high yield β_{0i} , $\beta_{1i} < 1$ and $\beta_{2i} > 1$, while an undesirable genotype shows $\beta_{1i} > 1$ and $\beta_{2i} < 1$.

3.3 The additive main effect and multiplicative interactions (AMMI) method

There are multivariate methods for the study of phenotypic stability, including AMMI as discussed by Crossa (1990), Gauch Junior (1985), Gauch Junior & Zobel (1988), Yau (1995) and Zobel et al. (1988). Many studies have applied both multivariate and univariate techniques and these methods have been useful for identifying stable and adapted genotypes (ACCIARESI & CHIDICHIMO, 1999; DIAS & KRZANOWSKI, 2003; FLORES et al., 1996, 1998; HOHLS, 1995; MEDINA et al., 1999; TAI, 1999; VARGAS et al., 1999; YAU, 1995).

The aim of the AMMI analysis is to model the interaction effects through a principal component model (JOHNSON & WICHERN, 1998). The AMMI model was developed by Gabriel (1971) and Gollob (1968), and has been applied and extended by many other authors. This model is defined by:

$$Y_{ij} = \mu + g_i + e_j + \sum_{k=1}^{m} \lambda_k r_{ik} s_{jk} + \delta_{ij} + \overline{\varepsilon}_{ij} \qquad (3.10)$$

The interaction effect is $ge_{ij} = \sum_{k=1}^{m} \lambda_k r_{ik} s_{jk} + \delta_{ij}$,

where λ_k is the eigenvalue associated with the kth principal component; r_{ik} is the ith element of the eigenvector for λ_k associated with genotypes, s_{jk} is jth

element of the eigenvector for λ_k associated with environments, and m is the number of retained components axes.

Let X be a matrix of the GE interactions parametric effects, defined by:

$$X = \begin{bmatrix} ge_{11} & \cdots & ge_{1q} \\ \vdots & \ddots & \vdots \\ ge_{p1} & \cdots & ge_{pq} \end{bmatrix}$$

The X matrix $(p \times q)$ with rank $r = min\ (p\text{-}1,\ q\text{-}1)$, is submitted to the singular value decomposition in the following way:

$$X = R\Lambda S^t \qquad (3.11)$$

where $R\ (p\ x\ r)$ and $S\ (q\ x\ r)$, R and S are column orthonormal, $R^t R = S^t S = I_r$ and $\Lambda = diag\left\{\lambda_k^{1/2}\right\}$, $k = 1, 2, ..., r$; λ_k is the kth non-null eigenvalue of XX^t or of X^tX, R and S are matrix of eigenvector of the related r eigenvalues.

In general, $m \le r$ non-null eigenvalues are kept, since they explain most of the total variation, given by:

$$tr\left(X^t X\right) = tr\left(XX^t\right) = \sum_{i=1}^{p}\sum_{j=1}^{q}\delta_{ij}^2 = \frac{SSGE}{n}$$

If the proportion of explanation is large for small m, the technique is considered efficient. Therefore, the interaction effects can be predicted by:

$$\hat{\delta}_{ij} = \sum_{k=1}^{m}\hat{\lambda}_k\,\hat{r}_{ik}\,\hat{s}_{jk}$$

For a formal evaluation to the lack of fit for the AMMI model, an analysis of variance can be accomplished as presented in Table 2 for m kept principal components.

An important characteristic of this method is the possibility of obtaining plots of the principal components kept in the analysis of the AMMI model, and plots of the scores of PC's axes against the mean yield. Genotype and environment scores can be plotted on the same graph and used to visually identify stability and the similarity between genotypes and environments. Therefore, it is useful to perform genotypes ecological zoning, through cluster analysis methods to identify genotype and environment relationships when two or more principal

Table 2 – Summarized analysis of variance of the AMMI model, considering the retention of m principal components, and decomposition of the degrees of freedom of the GE interaction using Gollob's (1968) method.

Tabela 2 – Resumo da análise de variância do modelo AMMI, considerando a retenção de m componentes principais e a decomposição dos graus de liberdade da interação GA usando o método de Gollob (1968).

Source of Variation	Degrees of Freedom		Sum of Square
Genotypes (G)	$p - 1$		
Environments (E)	$q - 1$		
GxE	$(p - 1)(q - 1)$		$SSGE$
PC1	$v_1 = p + q - 1\ \text{-}2\times 1$	Gollob	$n\lambda_1$
PC2	$v_2 = p + q - 1 - 2\times 2$	(1968)	$n\lambda_2$
$\vdots$	$\vdots$		$\vdots$
PCm	$v_s = p + q - 1 - 2\times m$		$n\lambda_s$
Deviation	$(\,p - 1\,)(\,q - 1\,) - \sum_{i=1}^{m} v_i$		$SSGA - \sum_{i=1}^{m} n\lambda_i$
Mean error	v		SSE

components are retained. Figure 6 shows a biplot of PC1 vs mean genotypes yields. A stable genotype is located as close as possible of the zero level of the PC1 axis. An ideal genotype has a high mean yield performance and is stable. An undesirable genotype has low stability as well low mean yield. If in the same PCi vs PCk graph the environmental and genotypic scores are plotted we can perform ecological zoning, by choosing the genotype and environment groups that are located close and in the same region of the biplot.

The AMMI model has some disadvantages. If the number of components retained in the model is large ($m \geq 3$), it is difficult to describe the behavior of the GE interaction effects due to the impossibility of obtaining graphs in more than three dimensions. It is possible to plot all pairs of components but, in this case, each component accounts only for a small portion of the total GE variation. Also, the environmental response pattern can not be estimated directly from the AMMI model. Hence, the AMMI model is mainly useful to identify the most stable and responsive genotypes, visually, in two or at most three dimensions. As the environmental genotype response pattern is almost always required by the plant breeders, this must be identified through one of the methods discussed previously.
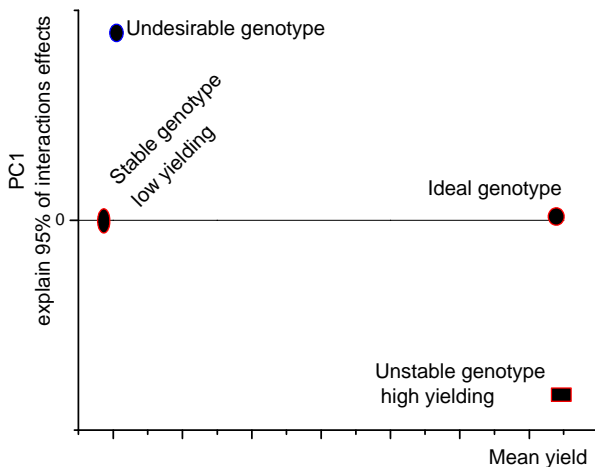

Figure 6 – Illustration of a biplot graph with PC1 vs mean yield, presenting the main types of genotypes and patterns of stability and adaptation.

Figura 6 – Ilustração do gráfico biplot com CP1 vs a produtividade média, apresentado os tipos principais de genótipos e padrões de estabilidade e adaptabilidade.

3.4 Other methods

Other methods that have been used can be considered complementary to the regression methods. For instance, in the method of Lin & Binns (1988, 1991) each genotype is compared with the genotype with maximum performance in each environment, and an index is obtained. This index is then partitioned into estimates of the genetic and GE interaction effects. Hence, it is possible to identify the genotype that contributes the most to the GE interaction. Other methods propose the estimation of a reliability index for each genotype, which is used for evaluating the phenotypic stability (ANNICCHIARICO, 1992; ANNICCHIARICO et al., 1995). The genotype response yield in the jth environment is expressed as the percentage of the jth environmental mean values and, for each genotype, the mean and the standard deviation over environments are estimated by using the transformed data. The reliability index is obtained by calculating the $100\,\alpha\,\%$ quantile from the normal distribution with this mean and standard deviation. The authors suggested a value of $\alpha = 0.25$.

4 EXAMPLE

Common bean yield (g/plot) data set involving nine genotypes and eight environments will be used to illustrate some of the described analyses. In each environment, a randomized complete block experiment was carried out with four replicates. The analysis of variance (Table 3) shows significant (*P<0.01*) genotype, environment and the interaction effects. The significant interaction effect justifies the phenotypic stability analysis.

Figure 7 shows the studentized residuals plot against environments indicating no heterogeneity of variances. The Shapiro Wilks' test was performed on residuals and the normality hypothesis was not rejected at the 5% level of significance.

Table 4 shows estimates of the genotype means and ecovalencies. Genotype 3 showed the largest ecovalency estimate, followed by genotypes 9 and 5. Genotypes 3 and 6 showed the lowest mean yield. Genotype 2, which has the smallest ecovalency estimate and the second largest mean, is considered stable and adapted. Genotypes 2, 4 and 6 also have low ecovalency estimates.

Table 3 – Summarized multi environment analysis of variance among nine common bean genotypes in eight locations.

Tabela 3 – Resumo da análise de variância conjunta entre oito genótipos de feijão em oito localidades.

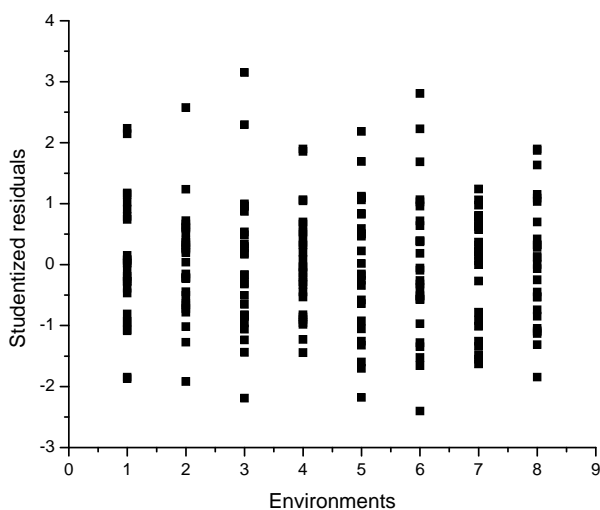| SV	DF	MS	F	P>F
Environments(E)	7	4343484.51	104.80	0.000
Blocks/Env (B/E)	24	41485.19	1.00	0.460
Genotypes (G)	8	185589.32	4.48	0,000
G x E	56	69908.50	1.69	0.008
Residual	128	41446.25		
Total	287	-		
CV	15.87%			
Overall mean	1283.12			

Figure 7 – Studentized residuals against environments plot showing no evidence of variance heterogeneity.

Figura 7 – Gráfico dos resíduos estudentizados em função de ambientes, mostrando que não evidências de heterogeneidade de variâncias.

Table 4 – Genotypes mean and ecovalency estimates for the 9 common bean varieties.

Tabela 4 – Médias genotípicas e estimativas de ecovalência para 9 variedades de feijão.

Genotypes	$\overline{Y}_{i\bullet}$	W_i
1	1317.50	344934.6
2	1394.79	193741.2
3	1192.50	1303388.3
4	1306.67	225704.1
5	1244.79	544904.2
6	1158.33	239935.5
7	1215.00	378683.7
8	1414.79	299659.7
9	1303.75	423669.8

Table 5 shows the result of fitting the uni-segmented linear regression model of Eberhart & Russell (1966). Genotypes 1, 2, 3, 4, 6, 8, and 9 have slope estimates that are not significantly different from 1 (*P>0.05*), and are, therefore, all considered to be broad stable genotypes. Genotype 5 has a slope greater than 1, and is considered less responsive. The lack of fit test is significant (*P<0.01*) only for genotypes 3.

The bi-segmented regression model of Cruz et al. (1989) (Table 6) shows that all genotypes have slopes for unfavorable environments that are not significantly different from unity. In favorable environments, the slopes of genotype 5 was significantly less than 1 and of genotypes 3 and 8 were significantly larger than one. Other genotypes showed slopes that are not significantly different from one. No genotype did show a desirable environmental response pattern. Genotype 3 showed a significant lack of fit.

Table 7 shows the test for equality among slopes in the favorable and unfavorable environments for the models of Toler & Burrows (1998). No genotype did show significant result (P < 0.05) in this test, and the uni-segmented model was used to model the genotype behavior.

Table 8 shows the estimates of the uni-segmented linear regression model of Toler & Burrows (1998). Only genotype 9 showed a slope significantly greater than one. Genotypes 5 and 7 had slopes significantly smaller than one. This indicates evidence that the Toler & Burrows (1998) model is more powerful than that of Eberhart & Russell (1966).

Table 5 – Parameters estimates and hypothesis test according to the Eberhart & Russell's (1966) uni-segmented linear regression model.

Tabela 5 – *Estimativas dos parâmetros e testes de hipóteses para o modelo de regressão linear uni-segmentado de Eberhart and Russell.*

Genotypes	$\hat{\beta}_{1i}$	t for $H_0 : \beta_{1i} = 1$	$Pr>/t/$	P-values for the lack of fit test
1	1.12	1.10	0.274	0.327
2	0.98	-0.14	0.885	0.604
3	1.10	0.89	0.375	0.000
4	0.96	-0.39	0.696	0.523
5	0.75	-2.27	0.025	0.254
6	0.97	-0.30	0.762	0.474
7	0.78	-1.98	0.050	0.531
8	1.13	1.14	0.254	0.449
9	1.22	1.95	0.053	0.394

Table 6 – Parameter estimates and hypothesis test to the Cruz et al. (1989) bi-segmented linear regression model.

Tabela 6 – *Estimativas dos parâmetros e testes de hipóteses para o modelo de regressão linear bi-segmentado de Cruz et al. 's (1989).*

Genotypes	$\hat{\beta}_{1i}$	t for $H_0 : \beta_{1i} = 1$	$Pr>/t/$	$\hat{\beta}_{1i} + \hat{\beta}_{2i}$	t for $H_0 : \beta_{1i} + \beta_{2i} = 1$	$Pr>/t/$
1	1.158	1.156	0.250	1.052	0.272	0.786
2	0.975	-0.183	0.855	1.001	0.007	0.995
3	0.942	-0.424	0.672	1.403	2.119	0.036
4	0.892	-0.789	0.432	1.081	0.428	0.669
5	0.906	-0.686	0.494	0.443	-2.930	0.004
6	1.115	0.847	0.399	0.677	-1.701	0.091
7	0.784	-1.581	0.116	0.774	-1.189	0.236
8	0.992	-0.059	0.953	1.389	2.046	0.043
9	1.234	1.719	0.088	1.180	0.948	0.345

Table 7 – Toler and Burrows' (1998) test for choosing uni- or bi-segmented regression models.

Tabela 7 – *Teste de Toler e Burrows (1998) para a escolha do modelo de regressão uni ou bi-segmentado.*

Genotypes	$\hat{\beta}_{2i} - \hat{\beta}_{1i}$	SE	t for $H_0 : \beta_{2i} - \beta_{1i} = 0$	$Pr>/t/$
1	-0.334	0.348	-0.960	0.339
2	0.337	0.348	0.969	0.335
3	0.563	0.349	1.616	0.109
4	0.114	0.347	0.329	0.742
5	-0.586	0.350	-1.674	0.097
6	-0.526	0.348	-1.509	0.134
7	0.112	0.348	0.321	0.749
8	0.429	0.348	1.233	0.220
9	-0.110	0.348	-0.316	0.752

Table 9 shows the eigenvalues estimates and the proportion of the total variation accounted for by each of the axes of the AMMI model. The first two components accounted for 77.8% of the total variation, and PC1 and PC2 can be retained without information loss.

Only the first two PC axes (Table 10) were significant ($P<0.01$), indicating that the model with PC1 and PC2 only

is appropriate for the data.

In the plot PC2 against PC1 (Figure 8), genotypes 3 and 5 are far from the origin and are considered unstable. On the other hand, genotypes 2 and 4 are close to the origin and are considered stable. It is easy to see associations between environments and genotypes. For instance, genotypes 2 and 6 are related to environment 3,

Table 8 – Toler & Burrows' (1998) uni-segmented regression model.

Tabela 8 – Modelo de regressão uni-segmentado de Toler e Burrows (1998).

Genotypes	$\hat{\beta}_{1i}$	SE	t for $H_0: \beta_{1i} = 1$	$Pr>|t|$
1	1.120	0.105	1.152	0.251
2	0.982	0.104	-0.173	0.863
3	1.109	0.105	1.040	0.301
4	0.956	0.104	-0.425	0.671
5	0.743	0.105	-2.447	0.016
6	0.962	0.104	-0.359	0.720
7	0.777	0.105	-2.130	0.035
8	1.131	0.105	1.257	0.211
9	1.219	0.105	2.093	0.038

Table 9 – Principal component (PC) model and variation account by each axis

Tabela 9 – Modelo dos componentes principais (CP) e variação total explicada de cada eixo.

Axes	Eigenvalues	% Variation account	% Cumulative variation
1	674214.9	51.7	51.7
2	340719.2	26.1	77.8
3	160256.4	12.3	90.1
4	60049.6	4.6	94.7
5	54953.4	4.2	98.9
6	10446.3	0.8	99.7
7	4318.9	0.3	100.0

Table 10 – Summarized the AMMI model analysis of variance for the common bean data.

Tabela 10 – Resumo da análise de variância do modelo AMMI para os dados do feijoeiro.

SV	DF	MS	F	P>F
Genoytpes (G)	8	185589.32	4.48	0.0001
Environments (E)	7	4343484.49	104.80	0.0000
G x E	56	69908.50	1.69	0.0082
PC1	14	144474.63	3.49	0.0001
PC2	12	85179.79	2.06	0,0244
PC3	10	48076.91	1.16	0.3239
Lack of fit	20	19465.24	0.47	0.9734
Residual	128	41446.25		
Total	287	-		
CV	20.7822%			
Overall mean	39.4515			

and genotype 1 is related to environment 2. Genotype 3 is not associated to any particular environment.
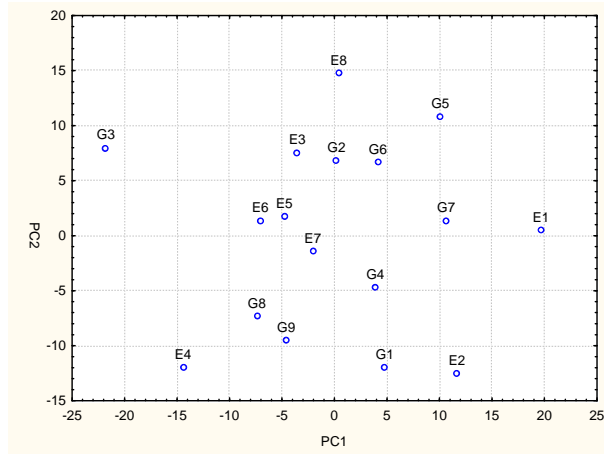

Figure 8 – Biplot of PC2 against PC1 displaying the genotypes and environments in the same graph.

Figura 8 – Biplot do CP2 contra CP1 mostrando os genótipos e ambientes no mesmo gráfico.

Figure 9 shows the plot of PC1 against genotype yields. Genotypes 8 and 2 were the most productive and stable; genotypes 1, 4, and 9 were stable with intermediate productivity and genotype 3 was the least stable with low productivity.
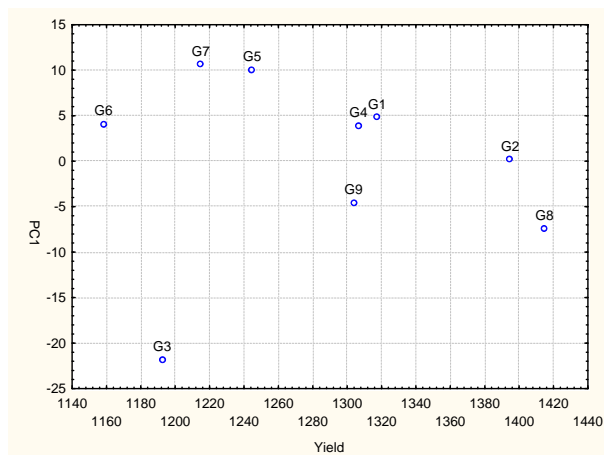

Figure 9 – Biplot of PC1 against mean yield for the 9 common bean genotypes.

Figura 9 – Biplot do CP1 contra a produtividade média dos 9 genótipos de feijão.

5 GENERAL CONSIDERATIONS

Many challenges still exist with the modelling of phenotypic stability. The AMMI model is criticized because the GE response patterns are not provided. This limitation, however, can be overcome. Toler & Burrows (1998) pointed out that there are a relationship between parameters of the AMMI model and those of the simple and bi segmented regression models. However, the authors did not clarify how to derive the regression parameters from those of AMMI model. The solution to that problem would allow the AMMI model to describe the environmental response patterns of the genotypes.

The use of mixed models for modelling and estimating GE interaction effects and for testing stability was considered by Yang (2002). This approach can be used if the genotypes, environments and GE interactions effects are considered as random effects.

The use of molecular markers to obtain information on the GE interaction in terms of QTL's (Quantitative Trait Loci) is an approach that has potential to be successful. Many biometric techniques could be developed to decompose the QTL effects of the genotypes and GE interaction, offering the potential for the improvement of the methods employed to evaluate phenotypic stability and to understand the causes of the GE interaction. Another field to be considered is the possibility of evaluating phenotypic stability for variables with non-normal distributions. No references about studies in this area have been found. Finally, the use of bootstrap and permutation tests is possible when the basic assumptions of the models are violated. Preliminary results that will be published in another article showed that bootstrap tests basically give the same results as the usual hypothesis testing for parameters of the linear uni and bi-segmented regression models, indicating the robustness of the *t* test with respect to the lack of independence between the response variable and the environmental index.

It can be conclude that the simple linear regression method of Eberhart & Russel (1966) and the bi-segmented regression method of Silva & Barreto (1985) have only historical importance nowadays. Instead, by considering the factors discussed in this paper, it is recommended that the regression models of Toler & Burrows (1998) and the AMMI model should be used simultaneously to estimate phenotypic stability effects.

A genotype with low phenotypic stability is predestined to be eliminated from the market. The use of appropriate biometrics techniques is necessary for

identifying the most adapted, responsive and stable genotypes in the final phases of the plant breeding program, where the high costs and the time spent in assays are powerful justifications to search for improved methods.

6 REFERENCES

ACCIARESI, H. A.; CHIDICHIMO, H. O. Genotype environment interaction in *Avena sativa* L.: employing AMMI and factorial correspondence models. **Pesquisa Agropecuária Brasileira**, Brasília, v. 34, n. 10, p. 1823-1830, out. 1999.

ANNICCHIARICO, P. Cultivar adaptation and recommendation from alfalfa trials in Northern Italy. **Journal Genetics and Breeding**, Rome, v. 46, n. 1, p. 269-278, Mar. 1992.

ANNICCHIARICO, P.; BERTOLINI, M.; MAZZINELLI, G. Analysis of genotype environment interactions for maize hybrids in Italy. **Journal Genetics and Breeding**, Rome, v. 49, p. 61-68, 1995.

BECKER, H. C.; LEON, J. Stability analysis in plant breeding. **Plant Breeding**, Berlin, v. 101, p. 1-23, 1988.

COMSTOCK, R. E.; MOLL, R. H. Genotype environment interaction. In: HANSON, W. D.; ROBINSON, H. F. (Eds.). **Statistical genetics and plant breeding**. Washington: National Academic Science Natural Research Council, 1963. p. 164-196, 982 p.

CROSSA, J. Statistical analysis of multilocations trials. **Advances in Agronomy**, New York, v. 44, p. 55 85, 1990.

CRUZ, C. D.; REGAZZI, A. J. **Modelos biométricos aplicados ao melhoramento genético**. Viçosa: UFV, 1994. 390 p.

CRUZ, C. D.; TORRES, R. A. de; VENCOVSKY, R. An alternative approach to the stability analysis proposed by Silva and Barreto. **Revista Brasileira de Genética**, Ribeirão Preto, v. 12, n. 2, p. 567 580, 1989.

DIAS, C. T. dos S.; KRZANOWSKI, W. J. Model selection and cross validation in additive main effect and multiplicative interaction models. **Crop Science**, Madison, v. 43, n. 3, p. 865-873, 2003.

DRAPER, N. R.; SMITH, H. **Applied Regression Analysis**. 3. ed. New York: J. Wiley and Sons, 1998. 706 p.

EBERHART, S. A.; RUSSELL, W. A. Stability parameters for comparing varieties. **Crop Science**, Madison, v. 6, n. 1, p. 36 40, 1966.

FINLAY, K. W.; WILKINSON, G. N. The analysis of adaptation in a plant breeding program. **Australian Journal of Agriculture Research**, East Melbourne, v. 14, n. 6, p. 742 754, 1963.

FLORES, F.; MORENO, M. T.; CUBERO, J. I. A comparison of univariate and multivariate methods to analyze G x E interaction. **Field Crops Research**, Amsterdam, v. 56, n. 3, p. 271-286, Apr. 1998.

FLORES, F.; MORENO, M. T.; MARTINEZ, A.; CUBERO, J. I. Genotype environment interaction in faba bean: comparison of AMMI and principal coordinate models. **Field Crops Research**, Amsterdam, v. 47, n. 2/3, p. 117-127, Aug. 1996.

FREEMAN, G. H.; PERKINS, J. M. Environmental and genotype-environmental components of variability: VIII. relations between genotypes grown in different environments and measures of these environments. **Heredity**, Edinburgh, v. 27, p. 15-23, 1971.

GABRIEL, K. R. The biplot-graphical display of matrices with applications to principal components analysis. **Biometrika**, [S.l.], v. 58, p. 453-467, 1971.

GALLANT, A. R. *Nonlinear statistical models*. New York: J. Wiley, 1987. 610 p.

GAUCH JUNIOR, H. G. **Integrating additive and multiplicative models for analysis of yield trials with assessment of predictive success**. [S.l.]: Cornell University Department of Agronomy, 1985. Mimeo.

GAUCH JUNIOR, H. G.; ZOBEL, R. W. Predictive and postdictive success of statistical analysis of yield trials. **Theoretical and Applied Genetics**, Berlin, v. 76, n. 1, p. 1 10, 1988.

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

HOHLS, T. Analysis of genotype environment interactions. **South African Journal of Science**, Pretoria, v. 91, n. 3, p. 121-124, Mar. 1995.

HUSSEIN, M. A.; BJORNSTAD, A.; AASTVEIT, A. H. SASG x ESTAB: a SAS program for computing genotype x environment stability statistics. **Agronomy Journal**, Madison, v. 92, n. 3, p. 454 459, May/Jun. 2000.

JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**. 4. ed. New Jersey: Prentice-Hall, 1998. 816 p.

LIN, C. S.; BINNS, M. R. A superiority measure of cultivar performance for cultivar x location data. **Canadian Journal of Plant Science**, Ottawa, n. 68, p. 193-198, Jan. 1988.

LIN, C. S.; BINNS, M. R. Assessment of a method for cultivar selection based on regional trial data. **Theoretical and Applied Genetics**, Berlin, v. 82, n. 3, p. 379-388, 1991.

MEDINA, J. L.; MOORE, P. P.; SHANKS, C. H.; GILL, F. F.; CHANDLER, C. K. Genotype x environment interaction for resistance to spider mites in Fragaria. **Journal of the American Society for Horticultural Science**, Alexandria, v. 124, n. 4, p. 353-357, Jul. 1999.

PERKINS, J. M.; JINKS, J. L. Environmental and genotype-environmental components of variability: III. multiple lines and crosses. **Heredity**, Edinburgh, v. 23, p. 339-356, 1968a.

PERKINS, J. M.; JINKS, J. L. Environmental and genotype-environmental components of variability: IV. non-linear interactions for multiple inbred lines. **Heredity**, Edinburgh, v. 23, p. 525-535, 1968b.

ROEMER, T. Sind die ertragsreichen sorten ertragssicherer? **Mitt. DLG**, [S.l.], v. 32, p. 87-89, 1917.

SHUKLA, G. K. Some statistical aspects of partitioning genotype-environmental components of variability. **Heredity**, Edinburgh, v. 29, p. 237-245, 1972.

SILVA, J. G. C. da. Análise da adaptabilidade através de regressão linear segmentada: I. fundamentos. **Pesquisa Agropecuária Brasileira**, Brasília, v. 30, n. 4, p. 435 448, 1995a.

SILVA, J. G. C. da. Análise da adaptabilidade através de regressão linear segmentada: I. aplicação. **Pesquisa Agropecuária Brasileira**, Brasília, v. 30, n. 4, p. 449 462, 1995b.

SILVA, J. G. C. da. Análise da adaptabilidade por regressão segmentada com estimação da junção dos segmentos. **Pesquisa**

Agropecuária Brasileira, Brasília, v. 33, n. 7, p. 1013 1029, 1998.

SILVA, J. G. C.; BARRETO, J. N. Aplicação da regressão linear segmentada em estudos da interação genótipo x ambiente. In: SIMPÓSIO DE EXPERIMENTAÇÃO AGRÍCOLA, 1., 1985, Piracicaba. **Anais**... Piracicaba: ESALQ, 1985. p. 49-50.

STORCK, L.; VENCOVSKY, R. Stability analysis based on a bi-segmented discontinuous model with measurement errors in the variables. **Revista Brasileira de Genética = Brazilian Journal of Genetics**, Ribeirão Preto, v. 17, n. 1, p. 75-81, 1994.

TAI, G. C. C. Canonical variate analysis of genotype x environment interactions. **Canadian Journal of Plant Science**, Ottawa, v. 79, n. 3, p. 427-431, Jul. 1999.

TOLER, J. E. **Patterns of genotype performance over environmental arrays**. Clemson: Clemson University, 1990. 154 p.

TOLER, J. E.; BURROWS, P. M. Genotype performance over environmental arrays: a non-linear grouping protocol. **Journal of Applied Statistics**, [S.l.], v. 25, n. 1, p. 131-143, 1998.

VARGAS, M.; CROSSA, J.; EEUWIJK, F. A. van; RAMIRES, M. E.; SYRE, K. Using partial least square regression, factorial regression, and AMMI models for interpreting genotype x environment interaction. **Crop Science**, Madison, v. 39, n. 4, p. 9555-9567, Jul./Aug. 1999.

VERMA, M. M.; CHAHAL, G. S.; MURTY, B. R. Limitations of conventional regression analysis: a proposed modification. **Theoretical and Applied Genetics**, Berlin, v. 53, p. 89 91, 1978.

WRICKE, G. Zur berechnung der ökovalenz bei sommerweizen und hafer. **Zeitschrift für Pflanzenzüchtung = Journal of Plant Breeding**, Berlin, v. 52, n. 2, p. 127-138, 1964.

YANG, R. C. Likelihood-based analysis of genotype-environment interactions. **Crop Science**, Madison, v. 42, p. 1434-1440, 2002.

YAU, S. K. Regression and AMMI analysis of genotype x environmental interactions: an empirical comparison. **Agronomy Journal**, Madison, v. 87, n. 1, p. 121 126, Jan./Feb. 1995.

ZOBEL, R. W.; WRIGHT, M. J.; GAUCH JUNIOR, H. G. Statistical analysis of a yield trial. **Agronomy Journal**, Madison, v. 80, p. 388-393, 1988.