Repeatability of some phenotypic stability parameters - a resampling approach

Matheus Henrique Silveira Mendes[1]* and Magno Antônio Patto Ramalho[1]

Abstract: *The aim of this study was to estimate the repeatability of some phenotypic stability parameters: reliability index (I), regression coefficient (b_0 e b_1), ecovalence, and the mean of the absolute scores (MAS), and to check the implications of estimating stability parameters using a reduced number of environments. A dataset comprising 25 common bean lines assessed in 36 environments was used. These environments comprise different locations, sowing seasons, and years. Different experimental groups, from three to 18 environments, were simulated. For each group, analysis of variance was performed and the stability parameters were estimated. The greater repeatability estimates were observed for the mean (b_0), followed by the I and b_1. Ecovalence and MAS had the lowest values for the repeatability estimates, which is not desirable for breeding programs aiming at selection for high stability. The phenotypic stability parameter should not be estimated from fewer than 12 environments.*

Key words: *Phaseolus vulgaris, genotype x environment interaction, multiple environment trials, plant breeding, statistics.*

INTRODUCTION

Alternatives to mitigate the genotype x environment interaction are often presented in the literature (Malosetti et al. 2013). Most of these alternatives are focused on identifying the lines/hybrids with better adaptability and stability. These methods were initially proposed by Cochran and Yates (1938), and almost thirty years later, they came to mainstream research, due to the studies of Finlay and Wilkinson (1963) and Eberhart and Russel (1966). Since then, many other methods have been made available in the literature (Jarquín et al. 2016).

The greatest challenge is to identify the adaptability and stability parameters of high heritability and repeatability because the environments used to find the estimates represent quite a small sample of the conditions to which genotypes are subjected when they are cultivated. Additionally, in the event that the lines/hybrids identified as having better adaptability and stability are used in crosses to introduce these traits in other cultivars, selection will only be successful if the estimate used to measure stability has high heritability.

Genetic control of adaptability and stability parameters are not easy to study. Thus, few studies in the literature focus on gathering information about it. Studies in this field have used strategies that demand experiments be conducted with at least four replications per environment. Estimation of such

*Corresponding author:
E-mail: mendesmhs@gmail.com

a stability parameter is found through two replications at a time. Thus, it is possible to have two estimates, one using replications 1 and 2, and the other, 3 and 4, for example. Accordingly, the repeatability estimate could be estimated. The problem is that when such a strategy is used, one part of the environmental variation becomes common to both situations, a phenomenon called permanent genetic variation (Becker and Leon 1988, Soares and Ramalho 1993). Annicchiarico (1997) used another alternative that involved wheat, maize, and oat crops. This alternative compared stability parameters based on assessing genotypes in different locations within years. The repeatability of the parameters was found through comparison between estimates in different years. However, only the variation between years is considered in this methodology.

The environments where the Value for Cultivation and Use (VCU) experiments are conducted should represent the conditions the new cultivars will be exposed to in the future (Gauch and Zobel 1988). In other words, the environment effect on the VCU must always be random. Therefore, it is necessary that the target environments represent the future conditions the recommended cultivars will find.

Publications on studies that assess adaptability and stability through a small number of environments are common in the literature. It would be important to assess the implications of this practice of using a reduced number of environments in this type of study that focuses on cultivar recommendations.

The adoption of computer simulation and resampling methods are strategies for collecting information of this type (Li et al. 2012, Kleinknecht et al. 2016). Although these approaches have been broadly used to compare methods of conducting segregating populations, as well as the number of individuals to be sampled within the population (Silva Filho 2013), there has been no report of its use as an assessment method to study adaptability and stability.

In light of the foregoing, the present study was conducted to propose a method able to assess the repeatability (r^2) of stability parameters, as well as to check the implications of estimating stability parameters using a reduced number of environments.

MATERIAL AND METHODS

Data from the Value for Cultivation and Use trials (VCU) evaluating lines of *Phaseolus vulgaris,* conducted in the state of Minas Gerais from 2013 to 2015, were used in this study. The VCU assessed 25 lines of carioca-type grains, i.e., cream-color grains with brown stripes. The experiments were conducted in 36 environments. These environments comprise 10 locations representing common bean production area in Minas Gerais in three different sowing seasons and different years. The experiments were sown in three different seasons, in some places, only.

The grain yield in kg ha^{-1} was the trait under consideration. This data set was kindly provided by partner researchers from the Federal University of Lavras, the Federal University of Viçosa, Embrapa Rice and Beans, and the Agricultural Research Company of the State of Minas Gerais (Empresa de Pesquisa Agropecuária do Estado de Minas Gerais - EPAMIG).

Analyses of variance were carried out for each environment. Combined analysis of 36 environments was conducted after finding the homogeneity of error variances by considering the effect of the environment as random, and the line effect as fixed. Some parameters often used in the literature were estimated in order to assess the adaptability and stability of the genotypes. The following methods were used:

a) Wricke Methodology (Wricke and Weber 1997), which estimates ecovalence (ω_i) by considering the following model: $Y_{ij} = \mu + g_i + e_j + (ge)_{ij} + \varepsilon_{ij}$; in which Y_{ij} is the mean of line *i* in environment *j*; μ is the overall mean; g_i is the fixed effect of line *i*; e_j is the random effect of the environment *j*; $(ge)_{ij}$ is the random effect of the interaction between genotype *i* and environment *j*; and ε_{ij} is the average experimental error.

The ecovalence of line *i* (ω_i) corresponds to the sum of the squares of the interaction effects $(ge)_{ij}$ between line *i* and the environments in which it was evaluated, i.e.:

$$\omega_i = \sum_{j=1}^{K} (\bar{y}_{ij} - \bar{y}_{i.} - \bar{y}_{.j} - \bar{y}_{..})^2 = \sum_{j=1}^{K} ge_{ij}^2)$$

b) Non-segmented Linear Regression methodology of Eberhart and Russel (1969) using the following model:

$\bar{y}_{ij} = b_{0i} + b_{1i}I_j + \delta_{ij} + \varepsilon_{ij}$; in which $\bar{y}_{ij}$ is the mean of line i in environment j; b_{0i} is the intercept or the mean of line i; b_{1i} is the regression coefficient of line i; I_j is the environmental index estimated according to the expression: $I_j = \bar{y}_{.j} - \bar{y}_{..}$, in which $\bar{y}_{.j}$ is the mean of environment j and $\bar{y}_{..}$ is the overall mean; δ_{ij} is the regression deviation of line i in environment j, and ε_{ij} is the mean experimental error.

The coefficient of determination (R_i^2), which provides data adjustment to the estimated regression line, was also estimated, according to the following expression:

$R_i^2 = \dfrac{SS_{MODEL}}{SS_{TOTAL}} \times 100$, in which SS_{MODEL} is the model sum of squares and SS_{TOTAL} is the total sum of squares.

c) Method of Annicchiarico (1992), which estimates the reliability index of line i (I_i). This index was defined by the equation $I_i = \bar{p}_i - Z_{(1-\alpha)}S_i$, in which $\bar{p}_i$ is the mean of the percentage of line i in k environments, $Z_{(1-\alpha)}$ is the quantile of the normal distribution, with α being the level of significance, which was established at 5%; and $S_i = \sqrt{\sum\limits_{j=1}^{k} \dfrac{p_{ij} - \bar{p}_{i.}}{k-1}}$, in which

pij is the performance of line i in the percentage of the mean of the environment, i.e.: $p_{ij} = \dfrac{\bar{y}_{ij}}{\bar{y}_{.j}} \times 100$ **.**

d) Methodology of Gauch and Zobel (1996), known as Additive Main Effects and Multiplicative Interaction (AMMI). The model adopted was $\bar{Y}_{ij} = \mu + g_i + e_j + \sum\limits_{c=1}^{g} \sqrt{\lambda_c \alpha_{ic} \Upsilon_{jc}} + \varphi_{ij} + \bar{\varepsilon}_{ij}$; $\bar{Y}_{ij}$, μ, g_i, e_j, and $\bar{\varepsilon}_{ij}$ were already described within this model. Thus, λ_c is the eigenvalue of the cth principal component related to the GE interaction; α_{ic} is the eigenvalue of the cth principal component related to line i; Υ_{jc} is the eigenvector of the cth principal component related to environment j; and φ_{ij} is the residue not explained by the principal components (PC). The mean of the absolute scores weighed through the explanation of each PC (MAS) was estimated by considering the first two PCs in each line i, according to Pereira et al. (2009).

Experimental groups were simulated after the aforementioned estimates were obtained. The number of environments taken into account ranged from 3 to 18. The choice of environments in each simulated group was conducted through random sampling. Thus, two sets of environments were obtained, i.e., the first set was sampled in the 36 possible environments, with the size ranging from 3 to 18 environments; then the second sample was taken without replacement, with the same size as the first sample, that is, without repetition of environments present in the first group. Combined analysis of variance was applied to each simulated group, and the aforementioned stability parameters were estimated.

The above procedure allowed the r^2 estimates of the adaptability and stability parameters to be found, without repetition of any environments, neither years, seasons, or locations were considered in the model, only the lines were repeated. Analysis of variance was applied to each parameter using the following model: $q_{io} = \mu + g_i + \varepsilon_{io}$, in which q_{io} is the estimated value of the parameter for line i (i = 1, 2, ..., 25) in the simulation o (o = 1 and 2); μ is the overall mean; g_i is the effect of line i; and ε_{io} is the error associated with observation q_{io}.

Based on the expected mean squares of the aforementioned analysis of variance, the (r^2) of each parameter could be estimated through the following equation: $\hat{r} = \dfrac{\hat{\Phi}_g}{\hat{\Phi}_g + \dfrac{\sigma^2}{r}}$, in which $\hat{\Phi}_g$ is the deviation squared effect of the line obtained from the equation $\hat{\Phi}_g = (QM_g - QM_e)/r$, in which QM_g is the mean square associated with the variation among genotypes; QM_e is the mean square associated with the variation between simulations; and r is the number of simulations ($r = 2$). The (r^2) was also estimated for the mean of the lines in each analysis performed. In each different size of experimental group, 5000 estimates for (r^2) could be found.

The coincident behavior of the best lines in the simulations, in regards to analysis, could be seen by considering all environments. The simulations were repeated 10,000 times in each number of environments in the experimental groups. The algorithm to perform these analyses was written in the R software (R Core Team 2016).

RESULTS AND DISCUSSION

The summary of combined analysis of variance in the 36 environments (results not presented) shows that the

environmental sources of variation, the genotype x environment (GE) variations, were significant (P ≤ 0.01). The environmental source of variation was responsible for most of the total sum of squares. It should be noted that the assessment environments comprised two years, the sowing seasons in each year, and the locations. The growing conditions for common bean are often subject to many changes in crop years, locations, and seasons within the same year, as it is often reported in the literature concerning experiments conducted in Minas Gerais (Lima et al. 2013, Lima et al. 2014, Ferreira et al. 2015).

Five of the 25 lines assessed were common check lines, i.e., cultivars already recommended for planting. The other lines were obtained from the breeding programs of UFLA, UFV, and Embrapa Arroz e Feijão, i.e., they had to be divergent, as it was confirmed through analysis of variance. The occurrence of a significant GE interaction in experiments using common bean lines, as per the current study, is often related to other experiments conducted in the region. This reinforces the fact that the responses of the lines often do not coincide in the different environments the lines are assessed in.

The results found through the analyses performed are quite favorable for the aims of the current study, i.e., expressive differences between environments and lines and, above all, differences in the GE interaction. The simulations based on this data set likely reflect common situations involving the common bean crop in Minas Gerais and, certainly, throughout Brazil.

There is no consensus among researchers concerning the stability concept (Becker and Leon 1988). The adaptability concept in the present study is associated with the ability to advantageously use the environmental stimulus to achieve high grain yield. Accordingly, it can be inferred that the lines assessed herein had different adaptations due to the wide variation in grain yield (Table 1).

Lin et al. (1986) say that there are basically three concepts of stability: Type 1 – biological stability: the genotype shows little variation among environments; Type 2 – agronomical stability: stability is assessed as to whether or not the response to environmental variation is aligned with the mean performance of all lines assessed in the experiments; and Type 3 – which is more associated with behavior predictability. The lines are more stable if the mean square regression deviation, which assesses stability, is lower, or if the coefficient of determination of the regression equation tends to the unit. Lin and Binns (1988) have proposed a fourth type, which associates stability with unpredictable environmental factors alone, i.e., the line will be more stable if the mean square of the interaction between the lines and the years within locations is small. This type will not be dealt with in the present study.

The methodology of Eberhart and Russel (1966) allows assessment of the three types of stability. When the b_1 of the linear regression is not different from zero (0), the line would show biological stability. None of the lines exhibited such behavior. Notice that the lowest estimate of b_1 was 0.86, for line 15. Most of the 25 lines assessed herein can be classified as having agronomical stability, because b_1 did not differ from one (1); however, some studies have shown a b_1 estimate higher than 1, i.e., these lines may be responsive to most environmental conditions, which is desirable. Notice that all R² estimates were of high magnitude. This shows that all of them have type 3 stability, according to Lin and Binns (1988). It can also be inferred that all lines had good predictability.

The greatest challenge faced by breeders lies in recommending a new cultivar with the lowest possible risk in their decision, because the process of recommending a new cultivar is laborious and expensive. Any mistake in to the decision made impairs the breeders' work. Thus, one of the methods most adopted by breeders is the cultivar recommendation risk estimate. The ideal scenario is to have a line with a reliability index (I) above 100%, because at this value, the line is expected to have a performance greater than the average of all lines in different environments. Table 1 shows that none of the lines had an I estimate above 100%; however, it can be noted that some of them had a higher adoption risk; for example, number 21, whose I estimate was 79.9%, showed a higher probability of being below the average.

One of the most expressive conclusions based on the AMMI analysis is the explanation for each PC. The restriction lies in a large number of lines and environments, as occurred in the present study, because the contribution of the first PC, which is that which actually matters, is often of low magnitude. Thus, the first PC explained 17.43%; and the second, 13.33% of the variation. It is difficult to use a larger number of PCs to explain the variation, especially through graphics. Thus, it was decided to put the sum of the scores of each of the lines by considering the two first PCs only, as was done by Pereira et al. (2009). Notice that the MAS value ranged from 0.63 (line 13) to 6.23 (line 17), i.e., line 19 was more

Table 1. Estimates of mean grain yield (kg ha^{-1}), the reliability index (I) of the method of Annicchiarico, the regression coefficient (b_1), the R² method of Eberhart & Russel, the ecovalence (ω_i) of Wricke, and the weighted average of the absolute scores of the first and second principal components (MAS) of the 25 common bean lines assessed in 36 environments

Line	Mean	I	b_1	R²	ω_i (%)	MAS
1	2645 a[1]	94.08	1.13*	0.91	3.67	4.18
2	2758 a	96.58	1.09	0.82	6.66	5.00
3	2697 a	95.32	1.08	0.84	5.74	5.01
4	2629 a	93.03	1.12*	0.90	3.64	3.55
5	2627 a	91.03	1.14*	0.87	5.35	3.96
6	2743 a	98.42	1.13*	0.90	4.16	2.28
7	2647 a	93.00	1.05	0.81	6.39	6.02
8	2746 a	99.44	1.13*	0.93	2.72	1.11
9	2428 b	85.50	1.05	0.89	3.66	0.96
10	2488 b	91.55	0.95	0.90	2.49	1.76
11	2294 c	80.38	0.99	0.90	2.76	0.77
12	2504 b	92.53	0.89*	0.88	2.96	3.61
13	2578 b	88.80	1.01	0.85	4.49	0.63
14	2388 c	84.70	0.91	0.86	3.69	3.57
15	2364 c	87.20	0.86**	0.85	3.75	0.89
16	2448 b	89.26	0.86**	0.84	3.92	4.18
17	2652 a	94.62	0.96	0.82	5.03	6.23
18	2398 c	85.88	0.92	0.84	4.07	3.98
19	2504 b	93.03	1.00	0.93	1.98	0.80
20	2384 c	86.71	0.92	0.90	2.61	1.08
21	2248 c	79.93	0.98	0.89	3.01	2.34
22	2496 b	91.63	0.83*	0.89	2.79	1.27
23	2513 b	85.06	1.13*	0.87	5.00	4.19
24	2332 c	81.93	0.90*	0.80	5.46	4.93
25	2624 b	95.73	0.97	0.86	4.02	1.38

[1] Values followed by the same letter are not different according to the Scott & Knott test at 5% probability; * Significant according to the t test (b_1 =1) at 5% probability.

biologically stable and line 17 was less stable. Observe that these MAS values are quite consistent with the ecovalence, i.e., the measure of contribution of line *i* to the GE interaction (Table 1).

The current study was conducted to answer some questions that would be hard to answer except through computer simulation. One of these questions concerns the implications of the adaptability and stability parameter estimates when the number of environments is quite small.

One strategy of breeders is selection to develop greater stability in lines/hybrids. It seems this strategy has not yet been adopted, mainly due to the difficulty in obtaining a stability parameter among the many options available in the literature that can lead to success in selection. In order to assess the feasibility of selection for a trait, the most important information is the heritability/repeatability of the line. If its r^2 is of low magnitude, success in selection will clearly be limited (Bernardo 2010, Hallauer et al. 2010).

A problem lies in how to arrive at these estimates of stability parameters. In the past, this procedure was based on considering cases of line/hybrid assessment in different environments, with four or more replications. In this case, the experiments would be analyzed by considering two replications at a time, which would be chosen at random. That means, making parameter estimates for each combined analysis of two replications, and then proceeding with analysis of variance of the parameter involving the results of both situations. Based on the expected mean squares of the analyses of variance, it was necessary to estimate the r^2 of the parameter. This procedure has already been adopted in some cases using plant species in Brazil (Soares and Ramalho 1993).

Another strategy was promoted by Eagles and Frey (1977), Virk et al. (1985), Leon and Becker (1988), and Jalaluddin and Harrison (1993). They used the classification correlation of the stability estimators by sub-dividing the environments

where the genotypes were assessed in two pre-set sub-groups. The strategy adopted in the present study was similar; however, the resampling process was repeated 10,000 times for the sub-group of each size of sample environments. The stability parameters estimated in each pair of sub-groups were subjected to analysis of variance and the r^2 of these parameters was estimated.

Table 2 shows that r^2 varied among the parameters assessed. Regardless of the sub-group size, r^2 was always greater for the mean. A similar result was reported by Jalaluddin and Harrison (1993), who used another comparison methodology applied to wheat. The second estimate with a higher r^2 magnitude was the reliability index of Annicchiarico (Annicchiarico 1992). The r^2 estimates of b_1 in the linear regression of E & R (Eberhart and Russel 1966) are in an intermediate position. Becker (1987), who is cited by Annicchiarico (1997), showed that the r^2 of b_1 exhibited low heritability when he worked with wheat and barley. Lower r^2 values were found through the Wricke method (Wricke and Weber 1997) and through the MAS (Gauch and Zobel 1996). These two last estimates, regardless of the sample size, always exhibited low magnitude

Table 2. Repeatability estimates (r^2) of the Mean, the Reliability index (I) of the Annicchiarico method, the regression coefficient (b_1) of the method of Eberhart & Russel, the ecovariance (ω_i) of Wricke, and the weighted average of the absolute scores of the first and second principal components (MAS) found through simulations applied to groups of experiments with different numbers of environments in assessment of 25 common bean lines

Sub-Group Size	Mean			I			b_1			ω_i			MAS		
3	0.49[1]	0.48[2]	0.5[3]	0.29	0.27	0.3	0.08	0.06	0.09	-0.08	-0.09	-0.06	-0.01	-0.03	0.01
4	0.56	0.547	0.56	0.35	0.33	0.36	0.13	0.11	0.15	-0.05	-0.07	-0.03	0.03	0.01	0.04
5	0.63	0.63	0.64	0.41	0.4	0.42	0.2	0.19	0.22	-0.03	-0.04	-0.01	0.07	0.05	0.08
6	0.67	0.66	0.67	0.44	0.43	0.45	0.24	0.22	0.26	-0.01	-0.03	0.01	0.1	0.08	0.11
7	0.71	0.703	0.71	0.5	0.49	0.51	0.28	0.26	0.29	0.03	0.01	0.04	0.12	0.11	0.14
8	0.73	0.73	0.73	0.52	0.51	0.53	0.33	0.31	0.34	0.05	0.03	0.06	0.16	0.14	0.17
9	0.75	0.75	0.76	0.55	0.54	0.56	0.38	0.36	0.39	0.06	0.05	0.08	0.17	0.16	0.19
10	0.77	0.77	0.78	0.57	0.56	0.58	0.41	0.4	0.42	0.1	0.08	0.11	0.21	0.2	0.22
11	0.79	0.79	0.79	0.6	0.6	0.61	0.44	0.42	0.45	0.1	0.09	0.12	0.22	0.2	0.23
12	0.81	0.8	0.81	0.62	0.62	0.63	0.47	0.46	0.48	0.13	0.11	0.14	0.23	0.22	0.25
13	0.82	0.82	0.82	0.643	0.64	0.65	0.5	0.5	0.51	0.15	0.14	0.16	0.25	0.24	0.26
14	0.83	0.83	0.83	0.66	0.65	0.66	0.53	0.52	0.54	0.17	0.16	0.18	0.27	0.26	0.29
15	0.84	0.84	0.85	0.68	0.67	0.68	0.56	0.55	0.57	0.19	0.18	0.2	0.3	0.28	0.31
16	0.85	0.85	0.85	0.69	0.69	0.7	0.58	0.57	0.58	0.21	0.2	0.22	0.31	0.3	0.33
17	0.86	0.86	0.86	0.71	0.7	0.71	0.59	0.58	0.6	0.22	0.21	0.23	0.32	0.31	0.33
18	0.87	0.86	0.87	0.71	0.71	0.72	0.62	0.61	0.62	0.24	0.23	0.25	0.35	0.33	0.36

[1] 5000 estimates, on average; [2] Lower limit of the confidence interval in the average of r^2 estimates at 5% probability; [3] Upper limit of the confidence interval in the average of r^2 estimates at 5% probability.

r^2 estimates - lower than 0.35. This outcome shows that these two parameters are not useful to breeders working in selection programs focused on enhancing stability. Annicchiarico (1997) observed that the repeatability estimate depended on the line/hybrid group assessed and on the function of the group of locations when he used some methods to measure stability, including the AMMI.

The number of environments involved in the process also affected r^2. Regardless of the parameter used, the r^2 estimate showed an upward trend depending on the number of environments involved (Table 2). In general terms, improvement was more expressive up to 11 and 12 environments; from these numbers on, the r^2 estimate showed a downward trend. It was possible to estimate the confidence interval of the mean r^2 estimate in each situation since there were 5000 r^2 estimates in each sub-group of differing numbers of environments. Regardless of the parameter, the larger the number of environments involved, the lower the magnitude of the confidence interval of r^2. Thus, it becomes clear that the increased number of environments makes not only the higher r^2 estimate possible, but also higher precision estimates.

Table 3. Estimates of coincidences in the Mean, the reliability index (I) of the Annicchiarico method, the regression coefficient (b_1) of the method of Eberhart & Russel, the ecovalence (ω_i) of Wricke, and applied to the weighted mean of the absolute scores of the first and second principal components (MAS) found through simulation of groups of experiments using different numbers of environments in which 25 common bean lines were assessed

Sub-Group Size	Coincidence									
	Mean		I		b_1		ω_i		MAS	
3	61.1		55.1		48.1		42.7		39.3	
	60.6	61.5	54.6	55.5	47.5	48.6	42.2	43.1	38.9	39.8
4	64.4		57.0		53.0		46.9		41.8	
	64.0	64.9	56.6	57.4	52.4	53.5	46.5	47.3	41.3	42.2
5	67.0		59.0		55.8		49.3		43.8	
	66.6	67.4	58.6	59.4	55.3	56.4	48.9	49.7	43.4	44.3
6	69.3		60.4		59.1		51.8		45.7	
	68.9	69.7	60.0	60.8	58.6	59.6	51.3	52.2	45.2	46.1
7	71.5		61.7		61.5		53.8		47.1	
	71.2	71.9	61.3	62.1	61.0	62.0	53.4	54.2	46.7	47.5
8	73.0		62.9		64.0		55.6		48.8	
	72.6	73.3	62.4	63.3	63.6	64.5	55.2	56.0	48.4	49.3
9	74.6		63.9		66.2		56.8		50.7	
	74.3	74.9	63.5	64.3	65.7	66.6	56.4	57.2	50.2	51.1
10	75.9		65.3		68.4		58.5		51.7	
	75.6	76.2	64.9	65.7	68.0	68.8	58.1	58.8	51.3	52.1
11	76.8		65.9		70.3		60.0		52.9	
	76.5	77.1	65.6	66.3	69.9	70.7	59.6	60.4	52.5	53.3
12	77.8		67.4		71.7		61.6		54.9	
	77.5	78.0	67.0	67.7	71.4	72.1	61.2	61.9	54.5	55.3
13	78.8		68.7		73.6		62.8		55.6	
	78.5	79.0	68.4	69.0	73.2	73.9	62.4	63.1	55.2	55.9
14	79.4		70.1		75.5		64.4		57.5	
	79.2	79.7	69.8	70.4	75.2	75.9	64.1	64.8	57.1	57.9
15	80.1		70.8		76.6		65.9		58.7	
	79.8	80.3	70.5	71.1	76.3	77.0	65.6	66.2	58.3	59.0
16	80.7		72.1		78.3		67.8		60.1	
	80.4	80.9	71.8	72.4	78.0	78.6	67.5	68.1	59.8	60.4
17	81.3		73.2		79.5		69.5		61.6	
	81.1	81.5	72.9	73.5	79.2	79.8	69.2	69.7	61.2	61.9
18	81.8		74.1		81.0		70.8		63.4	
	81.6	82.0	73.8	74.3	80.7	81.2	70.5	71.1	63.0	63.7

[1]mean of the 10,000 simulations; [2]Lower limit of the confidence interval of the mean of coincidence estimates at 5% probability; [3]Upper limit of the confidence interval of the mean of coincidence estimates at 5% probability.

Table 4. Percentage estimates of the variance explained through the first (PC1) and second (PC2) principal components of the AMMI method found by means of simulations applied to groups of experiments using different numbers of environments to assess 25 common bean lines

Sub-group size	Explained variance (%)					
	PC1			PC2		
3	67.25[1]			32.75		
	66.99[2]		67.50[3]	32.50		33.01
4	53.10			30.55		
	52.89		53.32	30.38		30.71
5	45.64			27.70		
	45.44		45.83	27.58		27.83
6	40.23			25.49		
	40.07		40.40	25.39		25.60
7	36.42			23.98		
	36.28		36.57	23.88		24.07
8	33.72			22.49		
	33.58		33.85	22.41		22.57
9	31.57			21.47		
	31.45		31.69	21.40		21.55
10	29.85			20.55		
	29.74		29.96	20.48		20.62
11	28.40			19.82		
	28.31		28.50	19.75		19.89
12	27.04			19.13		
	26.95		27.13	19.07		19.20
13	26.02			18.65		
	25.94		26.10	18.59		18.71
14	25.08			18.12		
	25		25.15	18.06		18.17
15	24.27			17.73		
	24.20		24.34	17.68		17.78
16	23.55			17.34		
	23.49		23.61	17.29		17.39
17	22.84			16.97		
	22.79		22.90	16.93		17.02
18	22.26			16.69		
	22.21		22.31	16.64		16.73

[1] Mean of the 10,000 simulations; [2] Lower limit of the confidence interval of the mean of the proportion of the variance explained through the principal component at 5% probability; [3] Upper limit of the confidence interval of the proportion of the variance explained through the principal component at 5% probability.

An important aspect in choosing the number of environments is that the number must be able to classify the lines involved, with the highest coincidence possible, considering the classification obtained when all the 36 environments were used in the analysis. Coincidence in the parameter estimates of the six best lines in the different number of environments involved, in the context of all environments, followed (as was expected) the aforementioned outcomes of r^2, i.e., the coincidence was higher for the mean estimates and lower for MAS and ω_i (Table 3).

When the AMMI method (Gauch and Zobel 1996) is used, one of the main questions is in regard to how much the first and second PCs explained the interaction, because the interpretation of the outcomes, especially of those expressed through graphics, is only consistent if the first two PCs are able to explain more than 60% of the interaction. Table 4 shows that the first two PCs explained most of the interaction when a small number of environments was used, seven or less. Such an outcome is expected, because the number of PCs is in accordance with the degrees of freedom of the environments or genotypes, whichever is smaller. As 25 lines were assessed, the number of environments established the number of PCs. Thus, in using three environments, two degrees of freedom for environment, the first and second PCs are able to explain 100% of the interaction, just as occurred in the present study. In contrast, in using 18 environments, 17 PCAs would be generated, and the first two of them would explain only 39% of the interaction (Table 4). All the variation explained by the other 15 PCAs was understood as noise. Actually, as already mentioned, when lines/hybrids are assessed in a large number of environments (which is desirable, and which sometimes happens), this method has the limitation that the two first PCs explain a small fraction of the variation due to the interaction (Annicchiarico 1997). The simulations performed here corroborate this conclusion. Note that the confidence interval of the percentage that was explained by the first two PCs was often too small, regardless of the number of environments sampled. This lends support to the previous statements.

REFERENCES

Annicchiarico P (1992) Cultivar adaptation and recommendation from alfalfa trials in Northern Italy. **Journal of Genetics and Breeding 46**: 269-278.

Annicchiarico P (1997) Joint regression vs AMMI analysis of genotype-environment interactions for cereals in Italy. **Euphytica 94**: 53-62.

Becker HC (1987) Zur Heritabilität statistischer Maßzahlen für die ertragssicherheit. **Vor. Pflanzenzüchtg 12**: 134-144.

Becker HC and Leon J (1988) Stability analysis in plant breeding. **Plant Breeding 101**: 1-23.

Bernardo R (2010) **Breeding for quantitative traits in plants**. 2nd edn, Stemma Press, Woodbury, 400p.

Cochran WG and Yates F (1938) The analysis of groups of experiments. **Journal of Agricultural Science 28**: 556-580.

Eagles HA and Frey KJ (1977) Repeatability of the stability variance parameters in oats. **Crop Science 17**: 253-256.

Eberhart SA and Russel WA (1966) Stability parameters for comparing varieties. **Crop Science 6**: 36-40.

Eberhart SA and Russel WA (1969) Yield stability for a 10-line diallel of single cross and double cross maize hybrids. **Crop Science 42**: 357-361.

Ferreira RADC, Ramalho MAP, Toledo FHRB and Abreu AFB (2015) Implications of the number of years assessment on recommendation of common bean cultivars. **Plant Breeding 134**: 599-604.

Finlay KW and Wilkinson GN (1963) The Analysis of Adaptation in a Plant-Breeding Programme. **Australian Journal of Agricultural Research 14**: 742-754.

Gauch HG and Zobel RW (1988) Predictive and postdictive success of statistical analyses of yield trials. **Theoretical and Applied Genetics 76**: 1-10.

Gauch HG and Zobel RW (1996) AMMI analysis of yield trials. In Kang MS and Gauch Jr HG **Genotype by environment interaction**. CRC Press, New York, p. 85-122.

Hallauer AR, Carena MJ and Miranda Filho JB (2010) **Quantitative genetics in maize breeding**. 3rd edn, Springer, New York, 664p.

Jalaluddin M and Harrison SA (1993) Repeatability of stability estimators for grain yield in wheat. **Crop Science 33**: 720-725.

Jarquín D, Pérez-Elizalde S, Burgueño J and Crossa J (2016) A hierarchical bayesian estimation model for multienvironment plant breeding trials in successive years. **Crop Science 56**: 1-17.

Kleinknecht KJ, Möhring FL, Meyer U and Piepho HP (2016) A simulation-based approach for evaluating the efficiency of multienvironment trial designs. **Crop Science 14**: 1-14.

Leon J and Becker HC (1988) Repeatability of some statistical measures of phenotypic stability - Correlations between single year results and multi-year results. **Plant Breeding 100**: 137-142.

Li X, Zhu C, Wang J and Yu J (2012) Computer simulation in plant breeding. **Advances in Agronomy 16**: 219-164.

Lima LK, Ramalho MAP, Abreu AFB, Toledo FHRB and Ferreira RADC (2014) Implications of predictable and unpredictable environmental factors in common bean VCU trials in Minas Gerais. **Crop Breeding and Applied Biotechnology 14**: 146-153.

Lima LK, Ramalho MAP, Ferreira RADC and Abreu AFB (2013) Repeatability of adaptability and stability parameters of common bean in unpredictable environments. **Pesquisa Agropecuária Brasileira 48**: 1254-1259.

Lin CS and Binns MR (1988) A superiority measure of cultivar performance for cultivar × location data. **Canadian Journal of Plant Science 68**: 193-198.

Lin CS, Binns MR and Lefkovitch LP (1986) Stability analysis: where do we stand? **Crop Science 26**: 894-900.

Malosetti M, Ribaut JM and van Eeuwijk FA (2013) The statistical analysis of multi-environment data: Modeling genotype-by-environment interaction and its genetic basis. **Frontiers in Physiology 4**: 1-17.

Pereira HS, Melo LC, Del Peloso MJ, Faria LC, Costa JGC, Diaz JLC, Rava CA and Wendland A (2009) Comparação de métodos de análise de adaptabilidade e estabilidade fenotípica em feijoeiro-comum. **Pesquisa Agropecuária Brasileira 44**: 374-383.

R Core Team (2016) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Silva Filho JL (2013) Optimizing the number of progenies and replications in plant breeding experiments. **Crop Breeding and Applied Biotechnology 13**: 151-157.

Soares AA and Ramalho MAP (1993) Repetibilidade de grãos e dos parâmetros de estabilidade na cultura do arroz. **Ciência e Prática 17**: 64-70.

Virk DS, Chahal SS and Pooni HS (1985) Repeatability of stability estimators for downy mildew incidence in pearl millet. **Theoretical and Applied Genetics 70**: 102-106.

Wricke G and Weber WE (1997) **Quantitative genetics and selection in plant breeding**. Walter De Gruyter and Co., Berlin, 420p.