



# Evaluation of soybean lines and environmental stratification using the AMMI, GGE biplot, and factor analysis methods

L.B. Sousa<sup>1</sup>, O.T. Hamawaki<sup>2</sup>, A.P.O. Nogueira<sup>3</sup>, R.O. Batista<sup>4</sup>,  
V.M. Oliveira<sup>5</sup> and R.L. Hamawaki<sup>6</sup>

<sup>1</sup>Laboratório de Melhoramento de Plantas, Instituto de Ciências Agrárias,  
Universidade Federal de Uberlândia, Uberlândia, MG, Brasil

<sup>2</sup>Programa de Melhoramento de Soja, Instituto de Ciências Agrárias,  
Universidade Federal de Uberlândia, Uberlândia, MG, Brasil

<sup>3</sup>Programa de Melhoramento de Soja, Instituto de Genética e Bioquímica,  
Universidade Federal de Uberlândia, Uberlândia, MG, Brasil

<sup>4</sup>Departamento de Genética e Bioquímica, Universidade Federal de Viçosa,  
Viçosa, MG, Brasil

<sup>5</sup>Programa de Melhoramento de Soja, Universidade Federal de Goiás,  
Palmeiras do Goiás, GO, Brasil

<sup>6</sup>Agronomy Research Center, Southern Illinois University,  
Carbondale, IL, USA

Corresponding author: L.B. Sousa  
E-mail: larissabsousa@hotmail.com

Genet. Mol. Res. 14 (4): 12660-12674 (2015)

Received March 6, 2015

Accepted June 9, 2015

Published October 19, 2015

DOI <http://dx.doi.org/10.4238/2015.October.19.10>

**ABSTRACT.** In the final phases of new soybean cultivar development, lines are cultivated in several locations across multiple seasons with the intention of identifying and selecting superior genotypes for quantitative traits. In this context, this study aimed to study the genotype-by-environment interaction for the trait grain yield (kg/ha), and to evaluate the adaptability and stability of early-cycle soybean genotypes using the additive main effects and multiplicative interaction (AMMI) analysis,

genotype main effects and genotype x environment interaction (GGE) biplot, and factor analysis methods. Additionally, the efficiency of these methods was compared. The experiments were carried out in five cities in the State of Mato Grosso: Alto Taquari, Lucas do Rio Verde, Sinop, Querência, and Rondonópolis, in the 2011/2012 and 2012/2013 seasons. Twenty-seven early-cycle soybean genotypes were evaluated, consisting of 22 lines developed by Universidade Federal de Uberlândia (UFU) soybean breeding program, and five controls: UFUS Carajás, MSOY 6101, MSOY 7211, UFUS Guarani, and Riqueza. Significant and complex genotype-by-environment interactions were observed. The AMMI model presented greater efficiency by retaining most of the variation in the first two main components (61.46%), followed by the GGE biplot model (57.90%), and factor analysis (54.12%). Environmental clustering among the methodologies was similar, and was composed of one environmental group from one location but from different seasons. Genotype G5 presented an elevated grain yield, and high adaptability and stability as determined by the AMMI, factor analysis, and GGE biplot methodologies.

**Key words:** *Glycine max.*; Genotype selections; Multivariate techniques; Grain yield

## INTRODUCTION

Soybean [*Glycine max* (L.) Merrill] is grown throughout Brazil, particularly in the State of Mato Grosso, which is the nation's largest soybean producer. This state produced approximately 26.44 million tons, in season 2015/2015, which accounts for 30% of the national soybean production (CONAB, 2015). However, because of its expansion throughout the state, soybean is cultivated in regions with different environmental conditions such as soil characteristics, temperature, photoperiod, and rainfall (Branquinho et al., 2014). As a result, the performance of soybean cultivars is influenced by interaction with these environments [genotype x environment (G x E) interaction], which makes it difficult to identify superior cultivars that are stable throughout the growing region. In particular, over the last few seasons (2011/2012, 2012/2013 and 2013/2014), the production regions of Mato Grosso have undergone periods of instability with regards to rainfall and temperature, which highlights the need to cultivate genotypes with high adaptability and stability within this region. Thus, in order to launch new cultivars, regional studies that aim to minimize environmental effects on the development of new soybean lines in different locations and/or planting seasons are necessary (Meotti et al., 2012). Cruz et al. (2012) recommends the stratification of regions for crop adaptation into more homogenous sub-regions, with the goal of minimizing the effects of G x E interactions. Several methodologies have been utilized to evaluate the performance of soybean lines and their interaction with the environment to direct the selection of the most productive, adapted, and stable lines for particular locations, regions, or planting seasons. Currently, the additive main effects and multiplicative interaction analysis (AMMI), genotype main effects and genotype x environment interaction (GGE) biplots, and factor analysis methodologies have been widely used to quantify the genotypic effects of the G x E interaction.

The AMMI method is used to evaluate the effects of G x E interactions and environmen-

tal stratification and permits the simple graphic representation of genotypes and environments in a multivariate dispersion diagram (Yokomizo et al., 2013). Several researchers have utilized this methodology for soybean cultivation (Meotti et al., 2012; Amira et al., 2013; Polizel et al., 2013). The GGE biplot, which is a modification of AMMI, is superior to the method from which it originates as it considers the effects of the genotypes along with the G x E interaction, while AMMI estimates them as additive effects (Yan et al., 2000). Factor analysis (Murakami and Cruz, 2004) is a multivariate technique that reduces the original number of abstract variables called factors. Thus, factor analysis can establish sub-environments with high correlation between the performance of lines within the same subgroup and little-to-no correlation between subgroups.

Therefore, the goal of the present study was to evaluate the adaptability and stability of the grain yield of early cycle soybean lines in the State of Mato Grosso using the AMMI, GGE biplot, and factor analysis methods, and to compare the efficiency of these methods.

## MATERIAL AND METHODS

The experiments were conducted during two agricultural years (2011/2012 and 2012/2013), in five cities of the State of Mato Grosso. Twenty-seven early-cycle soybean genotypes were evaluated, consisting of 22 lines developed by Universidade Federal de Uberlândia (UFU) soybean breeding program, and five controls: UFUS Carajás, MSOY 6101, MSOY 7211, UFUS Guarani, and Riqueza. The geographic location, altitude, and season in which the trials were completed can be found in Table 1.

**Table 1.** Location, altitude, and season of medium cycle soybean genotype competition trials in the State of Mato Grosso.

Environment	Altitude (m)	Location	Season
Alto Taquari	851	17°49'34"S and 53°16'56"W	2011/2012
Lucas do Rio Verde	390	13°03'01"S and 55°54'40"W	2011/2012
Sinop	384	11°50'53"S and 55°38'57"W	2011/2012
Alto Taquari	851	17°49'34"S and 53°16'56"W	2012/2013
Lucas do Rio Verde	390	13°03'01"S and 55°54'40"W	2012/2013
Querência	350	12°35'49"S and 52°11'59"W	2012/2013
Rondonópolis	227	16°28'15"S and 54°38'08"W	2012/2013

Prior to implantation in each experiment, a soil sample was collected for chemical and physical analyses to obtain an accurate lime and fertilization recommendation for use in soybean culture (Novais et al., 1999). The soil was prepared conventionally, with one aration and two harrowings. Prior to sowing, the area was furrowed and fertilized in accordance with the results of soil analyses for each location.

Sowing was carried out with seeds previously treated with the fungicide Fludioxonil (Maxim® XL) and the insecticide Tiametoxam (Cruiser® 350 FS), both at 200 mL of the commercial product every 100 kg of seeds. After the seeds were treated, an inoculation was carried out with *Bradyrhizobium japonicum* for affusion in the planting furrow with Biomax® at a ratio of 42 x 10<sup>8</sup> bacteria cells/mL of product using 50 L/ha. Weed management was carried out with pre- and post-emergent herbicides, and pest and disease management was carried out according to crop appropriate technical recommendations (Embrapa, 2011).

An experimental design of randomized blocks with three replications was utilized in all experiments. The experimental units (plots) of each experiment were composed of five

lines measuring 5 m in length, spaced 0.50 m, totaling 300,000 plants/ha. When phenologic stage R8 (Fehr and Caviness 1977) was reached, the usable area of each plot (three central lines disregarding 0.50 m of the edges) was harvested, cleaned, and the grains weighed. The grains were processed and weighed in order to estimate the yield (kg/ha), which was corrected for a moisture content of 13%, as follows:

$$PF = PI \frac{100 - UI}{100 - UF} \quad (\text{Equation 01})$$

where  $PF$  is the final corrected weight of the sample ( $\text{kg ha}^{-1}$ );  $PI$  is initial weight of the sample (g);  $UI$  is the initial moisture content of the sample (%);  $UF$  is the final moisture content of the sample (13%).

The grain yield data were subjected to an individual analysis for each experiment. Later, the homogeneity of the residual variance ( $>MSE/<MSE$ ) was evaluated in order to conduct the joint variance analysis (Ramalho et al., 2012). The effects of genotypes and environments were considered fixed.

Investigation of the  $G \times E$  interaction was carried out by first decomposing it into a complex fraction between pairs of environments, as described by Cruz and Castoldi (1991). The complex fraction was obtained through the following equation:

$$C = \sqrt{(1-r)^3} \sqrt{Q_1 Q_2} \quad (\text{Equation 02})$$

where  $Q_1$  and  $Q_2$  correspond to the average squares of the genotypes in environments 1 and 2, respectively, and  $r$  corresponds to the correlation between the averages of the genotypes in the two environments.

The phenotypic and genotypic correlations between the pairs of environments was estimated. The coefficient of phenotypic correlation was estimated using the following equation:

$$r_f = \frac{\text{COV}(Y_{ij}, Y_{ij'})}{\sqrt{\hat{V}(Y_j) \hat{V}(Y_{j'})}} \quad (\text{Equation 03})$$

where  $\text{COV}(Y_{ij}, Y_{ij'})$  is the phenotypic covariance of the trait  $X$  evaluated in environment  $j$  and  $j'$ ;  $\hat{V}(Y_j)$  is the phenotypic variance of trait  $X$  in environment  $j$ ; and  $\hat{V}(Y_{j'})$  is the phenotypic variance of trait  $X$  in environment  $j'$ .

The coefficient of genotypic correlation was estimated using the following expression:

$$r_f = \frac{\hat{\phi}_{g(jj')}}{\hat{\phi}_{g(jj')} + \hat{\phi}_{ga(jj')}} \quad (\text{Equation 04})$$

where  $\hat{\phi}_{g(jj')}$  is the genetic variability of  $X$  between environment  $j$  and  $j'$ ; and  $\hat{\phi}_{ga(jj')}$  is the variability in the interaction.

The statistical significance of the coefficient of phenotypic correlation was tested using a  $t$ -test, at a 5% level of probability, and the significance of the genotypic correlations was

evaluated using bootstrap with 5000 simulations.

When a significant G x E was detected, the stratification, adaptability, and phenotypic stability analysis was conducted using the AMMI methodology (Zobel et al., 1988), GGE biplot (Yan et al., 2000), and factor analysis (Murakami and Cruz, 2004).

The AMMI analysis was initially proposed by Mandel (1971); it allows greater detailing of the sum of squares of the G x E interaction. The AMMI analysis considers additive models for the principal effects: genotypes ( $g_i$ ) and environments ( $e_j$ ), and multiplicatives for the effects of the interaction,  $(ge)_{ij}$  (Malosetti et al., 2013). Thus, the average response of a genotype  $i$  in an environment  $j$  is given by:

$$\bar{Y}_{ij} = \mu + g_i + a_j + \sum_{c=1}^q \sqrt{\lambda_c} \alpha_{ic} \gamma_{jc} + \delta_{ij} + e_{ij} \quad (\text{Equation 05})$$

in which  $\bar{Y}_{ij}$ : average observed for the response variable of genotype  $i$  in environment  $j$ ;  $\mu$ : general average;  $g_i$ : effect of the genotype  $i$ ,  $i = 1, 2, \dots, g$ ;  $a_j$ : effect of the environment  $j$ ,  $j = 1, \dots, a$ ;  $\lambda_c$ : autovalue of the  $c^{\text{th}}$  principal component in relation to the G x E interaction;  $\gamma_{jc}$ : autovalue of the  $c^{\text{th}}$  principal component in relation to environment  $j$ ;  $\delta_{ij}$ : residual or noise not explained by the principal components;  $e_{ij}$ : average experimental error, where  $e_{ij} \sim N(0; V_e/r)$ ,  $r$  is the number of replications.

The index  $c$  varies from 1 to  $q$ , where 1 = minimum ( $g-1$ ;  $a-1$ ), the value of  $q$  corresponds to the post of the matrix of G x E interaction.

The GGE biplot analysis allows the response of each genotype to specific environments to be explored. The GGE biplot multiplicative model is similar to the AMMI multiplicative model (Souza, 2010) and was carried out considering the simplified model for two principal components centered on the environment (Yan, 2001):

$$\bar{Y}_{ij} - \mu_j = \lambda_1 \gamma_{i1} \alpha_{j1} + \lambda_2 \gamma_{i2} \alpha_{j2} + \bar{\varepsilon}_{ij} \quad (\text{Equation 06})$$

where:  $\bar{Y}_{ij}$ : average of genotype  $i$  in location  $j$ ;  $\mu_j$ : average of all genotypes in a location;  $\lambda_1 \gamma_{i1} \alpha_{j1}$ : the first principal component (CP1) of the genotype effect + G x E interaction, result of the principle components analysis applied to the matrix of interactions;  $\lambda_2 \gamma_{i2} \alpha_{j2}$ : is the second principal component (CP2) of the genotype effect + G x E interaction, result of the principle component analysis applied to the matrix of interactions;  $\lambda_1$  and  $\lambda_2$ : are the autovalues associated with CP1 and CP2;  $\gamma_{i1}$  and  $\gamma_{i2}$ : are the scores of CP1 and CP2, respectively, for the genotypes;  $\alpha_{j1}$  and  $\alpha_{j2}$ : are the scores of CP1 and CP2, respectively, for the environments;  $\bar{\varepsilon}_{ij}$  is the residual of the model with NID (0,  $\sigma^2/r$ ), in which  $\sigma^2$  is the variance of the error among plots and  $r$  is the number of replicates.

Factor analysis is a multivariate technique used in the study of environmental stratification that allows elevated numbers of environments to be reduced to a small number of abstract variables called factors (Cruz et al., 2014). These factors can be independent or they can be correlated with other factors; however, they are weakly correlated with the variables of other factors (Johnson and Wichern, 1992). Thus, it is suggested that factor analysis can form subgroups of environments such that there are high correlations in the trait of interest within subgroups and little or no correlation among subgroups (Cruz et al., 2014). The analyses were carried out with the help of the computational programs Genes (Cruz, 2013), Estabilidade (Ferreira, 2002), and GGE Biplot (Yan and Tinker, 2006).

## RESULTS AND DISCUSSION

Through joint analysis of variance, significant effects for environments, genotypes, and G x E interaction were observed, indicating differential behavior of the soybean genotypes in relation to the cultivation environment (Table 2).

**Table 2.** Summary of the joint variance analysis for the trait grain yield (kg/ha) evaluated in 27 early cycle soybean genotypes in seven environments.

Source of variation	d.f.	Average Square
Blocks/Environments	14	610,072.40564
Genotypes (G)	26	1,763,945.33591**
Environments (E)	6	75,874,149.99765**
G x E	156	1,457,188.75345**
Error	364	331,673.77927
Average		3,616.52
CV (%)		15.92
h <sup>2</sup>		81.20

\*\*Significant to P = 0.01 determined by the F test. d.f. = degrees of freedom, CV(%) = coefficient of variation, h<sup>2</sup> = coefficient of genotypic determination.

The average grain yield during these experiments was 3616.52 kg/ha, which was 20.15% greater than the national average of the State of Mato Grosso (3010.00 kg/ha), in season 2012/2013, demonstrating that the lines evaluated in this study have the potential to be launched as new cultivars that are recommended for this region (Conab, 2013). The coefficient of variation (CV) for grain yield was 15.92%, indicating experimental precision and control of environmental causes (Table 2). This value is considered acceptable for treatment as a quantitative trait and is consequently greatly influenced by the environment. Furthermore, the coefficient of variation in this study was lower than that obtained in soybean for grain yield (Santi et al., 2012; Yokomizo et al., 2013; Cavalcanti et al., 2014).

Heritability (h<sup>2</sup>) is one of the most important genetic parameters in breeding, as it enables the breeder to predict genetic gain, and it can be used to select the most appropriate breeding method for the trait. In advanced generations, the parameter h<sup>2</sup> is termed a coefficient of genotypic determination because it deals with pre-fixed genotypes, that is to say, those with a high degree of homozygosity and those that have already been selected. The coefficient of genotypic determination supplies the genotypic fraction present in the total phenotypic variance. In this way, it measures the reliability of the phenotypic value as an indicator of the genotypic value (Ramalho et al., 2012). The value of h<sup>2</sup> found in this study was 81.20% greater than that reported by Pinheiro et al. (2013) (maximum of 36%) for the same trait, indicating that the amount of genetic variation was lower than the amount of environmental variation.

Estimates of the phenotypic correlations between the pairs of environments were not significant, with the exception of the pairs formed by the environments Alto Taquari (2011/2012) and Alto Taquari (2012/2013) with Rondonópolis (2012/2013), which presented values close to 0.7. All of the other correlations were of low magnitude. These results are consistent with those reported by Branquinho et al. (2014) for some cities in Goiás where grain yield traits of middle cycle soybeans were distributed in different strata indicating the absence of significant correlations among these environments.

Estimates of genotypic correlations among all combinations of environments sur-

passed the phenotypic correlations that can be explained by the high genetic variability for grain yield within each environment (Table 3). Barros et al. (2012) reported similar results in studies conducted with 29 soybean genotypes in six environments in the State of Mato Grosso.

**Table 3.** Coefficient of phenotypic correlation (rf) and genotypic correlation (rg) among genotypes and decomposition of the partly complex interaction.

Environment		rf	rg	G x E (%)
Alto Taquari (2011/2012)	Lucas do Rio Verde (2011/2012)	0.32 <sup>ns</sup>	0.49	81.59
Alto Taquari (2011/2012)	Sinop (2011/2012)	0.36 <sup>ns</sup>	0.49	70.14
Alto Taquari (2011/2012)	Alto Taquari (2012/2013)	0.25 <sup>ns</sup>	0.34	83.86
Alto Taquari (2011/2012)	Lucas do Rio Verde (2012/2013)	0.18 <sup>ns</sup>	0.30	90.41
Alto Taquari (2011/2012)	Querência (2012/2013)	-0.10 <sup>ns</sup>	-0.14	98.25
Alto Taquari (2011/2012)	Rondonópolis (2012/2013)	-0.41 <sup>ns</sup>	-0.55	111.81
Lucas do Rio Verde (2011/2012)	Sinop (2011/2012)	0.25 <sup>ns</sup>	0.34	71.92
Lucas do Rio Verde (2011/2012)	Alto Taquari (2012/2013)	-0.16 <sup>ns</sup>	-0.23	102.59
Lucas do Rio Verde (2011/2012)	Lucas do Rio Verde (2012/2013)	0.29 <sup>ns</sup>	0.49	83.29
Lucas do Rio Verde (2011/2012)	Querência (2012/2013)	-0.10 <sup>ns</sup>	-0.14	93.19
Lucas do Rio Verde (2011/2012)	Rondonópolis (2012/2013)	0.07 <sup>ns</sup>	0.09	82.45
Sinop (2011/2012)	Alto Taquari (2012/2013)	0.11 <sup>ns</sup>	0.13	92.37
Sinop (2011/2012)	Lucas do Rio Verde (2012/2013)	0.27 <sup>ns</sup>	0.39	75.39
Sinop (2011/2012)	Querência (2012/2013)	0.15 <sup>ns</sup>	0.18	91.95
Sinop (2011/2012)	Rondonópolis (2012/2013)	-0.35 <sup>ns</sup>	-0.41	116.32
Alto Taquari (2012/2013)	Lucas do Rio Verde (2012/2013)	0.26	0.38	83.02
Alto Taquari (2012/2013)	Querência (2012/2013)	0.33 <sup>ns</sup>	0.40	79.95
Alto Taquari (2012/2013)	Rondonópolis (2012/2013)	-0.62 <sup>ns</sup>	-0.75	125.78
Lucas do Rio Verde (2012/2013)	Querência (2012/2013)	0.16 <sup>ns</sup>	0.24	82.92
Lucas do Rio Verde (2012/2013)	Rondonópolis (2012/2013)	0.03 <sup>ns</sup>	0.05	89.08
Rondonópolis (2012/2013)	Rondonópolis (2012/2013)	-0.04 <sup>ns</sup>	-0.05	102.18

<sup>ns</sup>:not significant by *t*-test at a P = 0.05 level of significance.

Estimates of decomposition of the partly complex G x E interaction greater than 50% denote a predominance of complex interactions, while estimates lower than 50% denote predominance of simple interactions. The complex interactions make identification of superior genotypes difficult, and for this reason justify the analyses of phenotypic adaptability and stratification to be performed (Cruz et al., 2014) (Table 3). All environment pairs presented complex interactions (Table 3). In other words, there was inconsistency in the superiority of the genotype with the environmental variation, which hampers the indication of cultivars and lines (Cruz and Castoldi, 1991; Vencovsky and Barriga, 1992; Pelúzio et al., 2008). Glasenapp et al. (2015) also observed predominance when there were complex interactions between grain yield and other traits that related to yield.

Five estimates of the decomposition of the G x E interaction between the environment exceeded 100%, which can occur in situations when the correlations among the environments are negative (Cruz et al., 2012) (Table 3). This result is consistent with the findings of Pelúzio et al. (2008) in studies evaluating 20 soybean cultivars four planting seasons in the State of Tocantins.

Faced with a significant G x E interaction and seeking to mitigate its effects upon cultivar recommendation, an alternative is to perform a study of adaptability and phenotypic stability and of environmental stratification. Thereby, concerning the decomposition of the G x E interaction by the AMMI methodology, one can observe in Table 4 that all of the principal components were significant. Meotti et al. (2012) observed a similar result when performing studies identifying stable cultivars that are adapted to different planting seasons.

For the principal components, we could observe that in order to get a greater per-

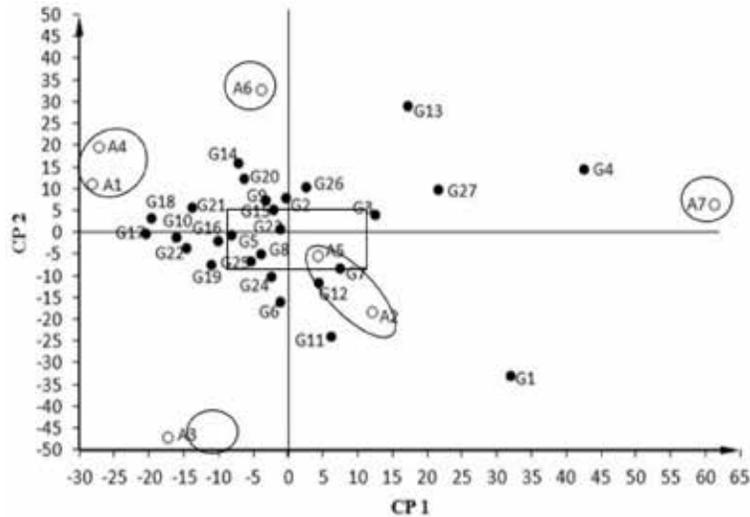
centage of explained variance of the sum of squares of the interaction, that is 78.93%, three principal components were necessary to exceed the limit of 70%, as suggested by Ramalho et al. (2012) for the AMMI analysis. A similar result was observed by Maia et al. (2014) who evaluated 44 soybean genotypes in different environments, and reported a value of 70% when including the third principal component. The first two principal components capture most of the variation, and there is a reduction in the subsequent components. Gauch and Zobel (1996) demonstrated that the excessive inclusion of multiplicative terms (principal components) could seriously reduce the accuracy of the analysis. Therefore, in order to build a graphic biplot, they adopted the first two principal components, which is similar to the procedure carried out in other studies of adaptability and phenotypic stability in soybean (Correa, 2007; Asfaw et al., 2009). While there is no consensus on the minimum proportion of the sum of squares of the G x E interaction that should be accumulated by the first principal component to construct a biplot, graphic dispersion is desirable as it permits inferences to be made about the adaptability, stability, and yield of genotypes.

The interpretation of stability using the AMMI model is performed from the distance of representative points of the genotypes and environments to the zero score of the two principal components, thus Figure 1 shows that genotype G23 is the most stable, followed by genotypes G2, G3, G5, G8, G15, and MSOY 7211.

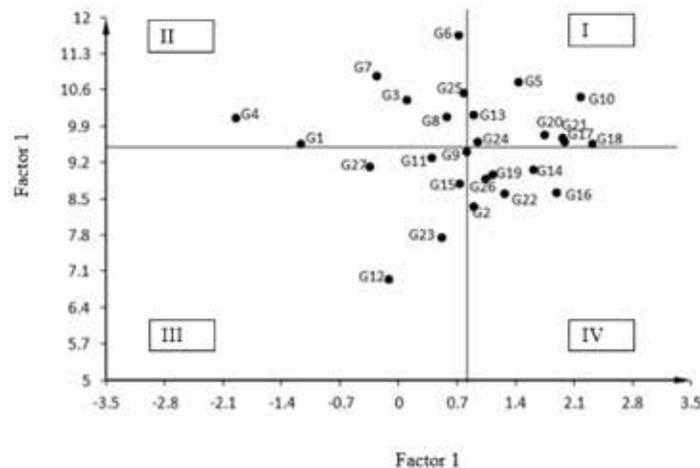
On the other hand, genotypes G1, G4, and G13 were the most unstable, and were therefore the genotypes that most contributed to the G x E interaction. In general, neither the most stable nor the most unstable genotypes are the most productive, that is, those that contribute the most to interaction. The most desirable genotypes are those that are stable and present high yield. In this case G23 was the most stable genotype, which presented a grain yield of 3326.14 kg/ha, 8.03% lower than the general average of the genotypes in all of the environments; however, it is still considered to be satisfactory as this yield is above the national average for the State of Mato Grosso (3010.00 kg/ha) (CONAB, 2013). On the other hand, genotype G5 was classified as having high stability and presented an elevated average grain yield, which exceeded the overall average (Table 3). The G1 and G4 genotypes were the most unstable and generated average yields higher than those of G23, at 3490.05 and 3683.62 kg/ha, respectively. Genotype G7 showed the greatest yield and the greatest averages for most of the environments studied, and a higher average yield of 5916.67 kg/ha was obtained in Sinop in season 2011/2012 (Table 3), which is consistent with the results reported by Amira et al. (2013), who observed that the most productive genotypes were among the most unstable. Unstable genotypes respond better to environmental oscillations. If the environment is favorable, their behavior will be favorable; however, if the environment is unfavorable, the yield produced by this genotype could be compromised. The environment that least contributed to the G x E interaction and was therefore the most stable was Lucas do Rio Verde in 2012/2013, followed by Lucas do Rio Verde safrain 2011/2012, and the most unstable were Sinop 2011/2012 and Rondonópolis 2012/2013 (Figure 1).

Through AMMI analysis, it is possible to identify the genotypes carrying specific adaptations. When considering the high proximity of the reference points for genotypes and environments, and observing the signs of the scores for genotypes and environments, sees that when the signs are the same they interact positively, and when the signs are contrary they interact negatively (Duarte and Vencovsky, 1999). Genotypes G7, G11, and G12 were the most stable and most adapted to Lucas do Rio Verde regardless of the season, with genotype G7 presenting specific adaptations to this environment (Figure 1). Genotypes G5, UFUS Carajás, G8, G19,

G22, MSOY 7211, and UFUS Riqueza were most adapted to Sinop (season 2011/2012) meaning that high yields are expected when these lines are cultivated in this environment (Figure 2).



**Figura 1.** Plot of the scores of the first two principal components in regards to environmental stratification, according to the AMMI model for the trait grain yield, for 27 early cycle soybean genotypes: G1, G2, G3, G4, G5, G6 (UFUS Carajás), G7, G8, G9, G10, G11, G12, G13, G14, G15, G16, G17, G18, G19, G20, G21, G22, G23, G24 (MSOY 6101), G25 (MSOY 7211), G26 (UFUS Guarani), and G27 (UFUS Riqueza), in seven environments; A1 (Alto Taquari in 2011/2012), A2 (Lucas do Rio Verde in 2011/2012), A3 (Sinop in 2011/2012), A4 (Alto Taquari in 2012/2013), A5 (Lucas do Rio Verde in 2012/2013), A6 (Querência in 2012/2013), A7 (Rondonópolis in 2012/2013). CP1: Principal component 1 and CP2: principal component 2.



**Figura 2.** Dispersion of 27 genotypes, G1, G2, G3, G4, G5, G6 (UFUS Carajás), G7, G8, G9, G10, G11, G12, G13, G14, G15, G16, G17, G18, G19, G20, G21, G22, G23, G24 (MSOY 6101), G25 (MSOY 7211), G26 (UFUS Guarani), and G27 (UFUS Riqueza) in relation to the representative axes for environmental stratification. Factor 1 represents region R1, composed of locations: Alto Taquari in 2011/2012 and Alto Taquari in 2012/2013, and factor 2 represents region R2, composed of locations: Lucas do Rio Verde in 2011/2012 and Lucas do Rio Verde in 2012/2013.

As for environmental stratification, five groups were formed. Two were formed by two environments: one with Alto Taquari in 2011/2012 and Alto Taquari in 2012/2013, and the other with Lucas do Rio Verde in 2011/2012 and Lucas do Rio Verde in 2012/2013 (Figure 2), albeit the correlations between the environments were not significant (Table 4). The formation of few groups and isolated environments can be explained by the complex nature of the interaction and by the fact that there was no correlation between the environments (Table 4).

**Table 4.** Proportion of the rate of interaction for each principal axis of the AMMI analysis, for 27 genotypes.

Principal component	Explanation (%)	Accumulated explanation (%)
CP1**	40.07	40.08
CP2**	21.38	61.46
CP3**	17.46	78.93
CP4**	10.62	89.55
CP5**	5.73	95.28
CP6**	4.72	100.00

\*\*Significant by the F test at a P = 0.01 level of significance. Principal component = principal axis.

Environmental stratification and phenotypic adaptability can be studied using factor analysis. In Table 3, four final factors are shown to be necessary to reach a value above 80% of the total variation of the data, which is consistent with that indicated by Murakami and Cruz (2004). To apply the information obtained by factor analysis, it is crucial that each variable considered can be adequately represented by common factors, those that will provide values that permit inference about the environment strata and about the adaptability of the genotypes studied. Thus, commonality was observed, which represents the proportion of the variance of the variable in each environment owing to the common factors, or this can be considered a measure of the efficiency of the representation of the variable by a common part (Cruz et al., 2014). The smallest value of commonality was 0.71 (Table 4) indicating that the description of the performance of each genotype considering grain yield in each environment was adequate. According to the factorial loads of factors 1 and 2, the environments were not clustered. For factor 1, only two environments presented factorial loads above the ideal: Alto Taquari in 2012/2013 (0.88) and Rondonópolis in 2012/2013 (-0.85); however, they were not clustered because they presented opposite signs. For factor 2, only one environment presented a factorial load above 0.7, which was Lucas do Rio Verde in 2012/2013 (0.90). Similar results were observed by Mendonça et al. (2007) when verifying the impossibility of clustering environments in the absence of factorial loads greater than 0.8 in the same factor in the study of environmental stratification, adaptability and stability of soybean genotypes. However, because of commonalities above 0.7, we opted to cluster the environments with closer factorial loads and the same sign. This way, for factor 1, Alto Taquari in 2011/2012 and Alto Taquari in 2012/2013 were allocated to group 1 (Region 1), and for factor 2, Lucas do Rio Verde in 2011/2012 and Lucas do Rio Verde in 2012/2013 comprised group 2 (Region 2).

The environments Sinopin 2011/2012, Lucas do Rio Verde in 2012/2013, and Querênciain 2012/2013 were considered favorable as they presented positive indices, and Alto Taquari in 2011/2012, Lucas do Rio Verde in 2011/2012, Alto Taquari in 2012/2013, and Rondonópolis in 2012/2013 were considered unfavorable because they presented negative indices (Table 5). Thus, approximately 58% of the environments proved to be unfavorable for the potential production of these soybean genotypes. The favorable environments showed the greatest average grain yields at 5064.69, 4648.35, and 3914.96 kg/ha, respectively (Table 6).

**Table 5.** Description of the environmental index for seven environments, commonality and initial and final factorial loads resulting from the decomposition of yield in each environment, and common and specific factors.

Environment	Index	Commonality	Initial loads		Final loads	
			Factor1	Factor 2	Factor 1	Factor 2
A1	-486.92	0.71	0.67	0.32	0.47	0.24
A2	-1133.80	0.71	0.25	0.80	-0.24	0.56
A3	1448.17	0.85	0.64	0.29	0.13	0.07
A4	-281.71	0.90	0.69	-0.55	0.88	0.20
A5	1031.82	0.86	0.45	0.33	0.11	0.90
A6	298.44	0.84	0.27	-0.40	0.11	0.10
A7	-876.01	0.86	-0.73	0.36	-0.85	0.18

Environment: A1 (Alto Taquari in 2011/2012), A2 (Lucas do Rio Verde in 2011/2012), A3 (Sinop in 2011/2012), A4 (Alto Taquari in 2012/2013), A5 (Lucas do Rio Verde in 2012/2013), A6 (Querência in 2012/2013), A7 (Rondonópolis in 2012/2013).

**Table 6.** Average grain yield (kg/ha) of 27 early cycle soybean genotypes cultivated in seven environments in the State of Mato Grosso.

Genotypes	Environments							Average
	A1	A2	A3	A4	A5	A6	A7	
G1	2793.00	3490.00	5547.33	2048.33	4379.00	2522.67	3650.00	3490.05
G2	3394.00	1879.00	4666.00	3006.67	4114.00	4038.67	2781.67	3411.43
G3	1994.00	2178.00	5466.67	3698.00	5395.67	4912.00	3797.00	3920.19
G4	2778.00	2626.00	3841.33	1849.33	4552.00	4486.67	5652.00	3683.62
G5	3746.33	2970.00	5611.67	4124.33	5126.00	4262.67	2718.00	4079.86
UFUS Carajás	4002.00	3266.00	5549.67	3002.67	5798.00	3042.67	2790.33	3921.62
G7	3042.00	3330.33	5916.67	3037.67	5256.00	4693.33	3406.67	4097.52
G8	3358.00	2795.00	5674.00	3015.33	5082.00	4544.00	2511.33	3854.24
G9	3258.67	2785.67	5032.33	3432.67	4397.33	4837.33	2592.33	3762.33
G10	3766.00	2321.00	5766.67	4392.67	5338.00	4022.67	2233.33	3977.19
G11	2701.00	2082.00	5517.00	2787.67	4924.00	1736.00	2971.00	3245.52
G12	3150.33	2116.00	5650.00	2342.67	3266.00	3361.33	3062.00	3278.33
G13	2916.67	2071.00	2600.00	3555.00	4672.00	3594.67	3765.33	3310.67
G14	2095.33	1467.00	4596.00	3902.00	4956.00	4890.67	2001.33	3415.48
G15	3105.67	2271.00	4926.67	3026.67	4346.00	4446.67	2500.00	3517.52
G16	3614.33	2517.00	4299.67	3705.00	3888.67	2242.67	1950.00	3173.90
G17	4188.67	2395.00	5826.67	3747.00	4850.00	4496.00	1800.00	3900.48
G18	3701.00	2753.00	5164.67	4432.67	4470.00	3946.67	1676.00	3734.86
G19	3198.00	2346.00	5800.00	3270.33	4618.00	4074.67	2003.33	3615.76
G20	2826.00	2625.00	3968.00	4107.67	4574.00	3760.00	2048.33	3415.57
G21	3438.33	2582.00	5053.33	4252.67	4626.00	4168.00	2029.33	3735.67
G22	3407.00	2267.00	5906.00	3301.67	4400.00	4560.00	1866.67	3672.62
G23	2275.33	2323.00	5116.67	3241.00	3698.00	4253.33	2375.67	3326.14
MSOY 6101	3615.67	2089.67	5533.33	2994.33	5012.00	3120.00	2775.00	3591.43
MSOY 7211	3080.00	3300.00	5491.00	3472.67	5094.67	4200.00	2275.33	3844.81
UFUS Guarani	2679.00	1777.00	5104.33	4042.67	4382.00	4450.67	3341.67	3682.48
UFUS Riqueza	2374.67	2410.67	3121.00	2250.33	4290.00	3040.00	3420.00	2986.67
Average	3129.59	2482.72	5064.69	3334.80	4648.35	3914.96	2740.51	3616.52

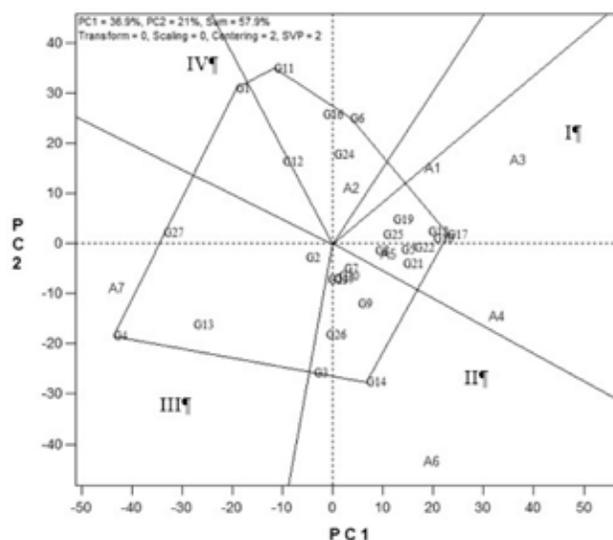
Environments: A1 (Alto Taquari in 2011/2012), A2 (Lucas do Rio Verde in 2011/2012), A3 (Sinop in 2011/2012), A4 (Alto Taquari in 2012/2013), A5 (Lucas do Rio Verde in 2012/2013), A6 (Querência in 2012/2013), A7 (Rondonópolis in 2012/2013).

Analysis of phenotypic adaptability was performed through the dispersion of scores in graphs where the axes were represented by common factors. In other words, each factor represented a region, or a stratum, in which there is similarity in the genotypic behavior (Cruz et al., 2014). In quadrants II and IV, the genotypes with adaptability specific to the region determined

by the factor are situated in two regions (R1 and R2) (Figure 2). Thus, genotypes G1, G3, G4, G6, G7, G8, and MSOY 6101 are specifically adapted to Lucas do Rio Verde, while genotypes G2, G14, G16, G19, G22, and MSOY 7211 perform better in Alto Taquari regardless of the season. Quadrant I represents broad adaptability, which in this case, was represented by genotypes G5, G10, G13, G17, G18, G20, G21, and MSOY 6101. Poorly performing genotypes can be found in quadrant III, specifically G9, G11, G12, G15, G23, and UFUS Riqueza, which are not recommended for cultivation in this region. Furthermore, the genotypes that presented the lowest average grain yields were allocated to region 1 (Table 6).

In the GGE biplot analysis, the first two principal components accounted for 57.9% of the variation and although this was lower than the ideal limit (70%), it is satisfactory, since the interaction among the environments was complex and there was great variability in the climatic conditions. This result is lower than that observed by Amira et al. (2013) (86.6%) and similar to that found by Asfaw et al. (2009) (61.50%). When using the first two principal components, four clusters of environments (mega-environments) were formed using the GGE biplot methodology (Figure 3), corroborating the results obtained by Amira et al. (2013).

The best environments were Sinop in 2011/2012 and Alto Taquari in 2012/2013, which presented a greater value for CP1, and a value closer to zero for CP2, demonstrating a greater power of discrimination between the genotypes (Figure 3) and greater yield in regards to the other environments (Table 6). It is believed that this occurred because of the good growing conditions with regards to water availability while the experiments were conducted in these locations, therefore facilitating greater photosynthetic capacity in the plants and consequently greater grain yield.



**Figura 3.** Plot of the scores of the principal components, regarding environmental stratification, according to the GGE biplot model for the trait grain yield, for 27 genotypes: G1, G2, G3, G4, G5, G6 (UFUS Carajás), G7, G8, G9, G10, G11, G12, G13, G14, G15, G16, G17, G18, G19, G20, G21, G22, G23, G24 (MSOY 6101), G25 (MSOY 7211), G26 (UFUS Guarani), and G27 (UFUS Riqueza), in seven environments: A1 (Alto Taquari in 2011/2012), A2 (Lucas do Rio Verde in 2011/2012), A3 (Sinop in 2011/2012), A4 (Alto Taquari in 2012/2013), A5 (Lucas do Rio Verde in 2012/2013), A6 (Querência in 2012/2013), and A7 (Rondonópolis in 2012/2013). PC1: principal component 1, PC2: principal component 2.

Genotypes G5, G10, G17, G18, and G22 were the most stable and productive and were indicated for mega-environment I, which was composed of the environments AtoTaquari in 2011/2012, AtoTaquari in 2012/2013, Lucas do Rio Verde in 2012/2013, and Sinop in 2011/2012. For mega-environment II, genotypes G3, G7, G9, G14, MSOY 7211, and UFUS Guarani were the most adapted. In mega-environment III, the most adapted genotypes were G2, G4, G13, and UFUS Riqueza, and for mega-environment IV genotypes, G1, G11, G12, G16, UFUS Carajás, and MSOY 6101 were the most adapted.

A comparison of the results obtained from all methods revealed that the AMMI model was the most efficient at retaining the greatest amount of the variation in the first two principal components (61.46%) followed by the GGE biplot model (57.90%), and factor analysis (54.12%) (Table 7).

**Table 7.** Autovalue, percentage value, and accumulated percentage in factor analysis performed with 27 early cycle soybean genotypes, cultivated in seven environments.

Autovalue	Percentage value	Accumulated percentage
2.25	32.17	32.17
1.54	21.95	54.12
1.20	17.18	71.30
0.77	11.08	82.39
0.58	8.37	90.77
0.45	6.51	97.29
0.18	2.70	100.00

The composition of the clusters obtained by all of the methods was similar for some environments, the models formed the set of corresponding environments from the same locale, however with distinct seasons. The most stable environment and the one that had the greatest yield was Lucas do Rio Verde in the 2012/2013 season, and genotype G5 was the most stable considering AMMI analysis, factor analysis, and GGE biplot.

The G x E interaction for grain yield in soybean was found to be complex. The AMMI analysis, factor analysis, and GGE biplot clustered environments in a similar fashion. The AMMI method was better than the GGE biplot and factor analysis models in explaining the greatest proportion of the sum of squares of the G x E interaction. Genotype G5 presented an elevated yield and high adaptability and stability in the AMMI, factor analysis, and GGE biplot methodologies. Lucas do Rio Verde in season 2012/2013 presented the greatest stability and the greatest grain yield.

### Conflicts of interest

The authors declare no conflict of interest.

### ACKNOWLEDGMENTS

Research supported by the Brazilian funding agencies Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

## REFERENCES

- Amira JO, Ojo DK, Ariyo OJ, Oduwaye OA et al. (2013). Relative discriminating powers of GGE and AMMI models in the selection of tropical soybean (*Glycine max* L. Merr.) genotypes. *Afr. Crop Sci. J.* 21: 67-73.
- Asfaw A, Alemayehu F, Gurum F and Atnaf M (2009). AMMI and SREG GGE biplot analysis for matching varieties onto soybean production environments in Ethiopia. *Sci. Res. Essays* 4:1322-1330.
- Barros HB, Sediyaama T, de Melo AV, Fidelis RR, et al. (2012). Adaptability and stability of soybean genotypes using univariate and multivariate methods. *J. Biotechnol. Biodiv.* 3: 49-58.
- Branquinho RG, Duarte JB, Souza PIM, Silva Neto SP, et al. (2014). Environmental stratification and optimization of a multi-environment trial net for soybean genotypes in Cerrado. *Pesq. Agropec. Bras.* 49: 783-795.
- Cavalcanti AK, Hamawaki OT, Hamawaki RL, Sousa LB et al. (2014). Fenotopic adaptability and stability in soybean genotypes in Porto Alegre do Norte, State of Mato Grosso. *Bioscie. J.* 30: 942-949.
- CONAB - Companhia Nacional de Abastecimento (2013). Acompanhamento da safra brasileira de grãos - nono levantamento. Brasília: CONAB, junho de 2013. Available at [[http://www.conab.gov.br/OlalaCMS/uploads/arquivos/13\\_06\\_06\\_09\\_09\\_27\\_boletim\\_graos\\_junho\\_2013.pdf](http://www.conab.gov.br/OlalaCMS/uploads/arquivos/13_06_06_09_09_27_boletim_graos_junho_2013.pdf)] Accessed July 1, 2014..
- CONAB - Companhia Nacional de Abastecimento (2015). Acompanhamento da safra brasileira de grãos - quinto levantamento. Brasília: CONAB, fevereiro de 2015, Available at [[http://www.conab.gov.br/OlalaCMS/uploads/arquivos/15\\_02\\_13\\_10\\_34\\_06\\_boletim\\_graos\\_fevereiro\\_2015.pdf](http://www.conab.gov.br/OlalaCMS/uploads/arquivos/15_02_13_10_34_06_boletim_graos_fevereiro_2015.pdf)]. Accessed February 22, 2015.
- Correa WR (2007). Adaptabilidade e estabilidade de genótipos de soja em Minas Gerais. Master's thesis. Universidade Federal de Uberlândia, Uberlândia.
- Cruz CD and Castoldi FL (1991). Decomposição da interação genótipos x ambientes em partes simples e complexa. *Rev. Ceres* 38: 422-430.
- Cruz CD (2013). GENES-a software package for analysis in experimental statistics and quantitative genetics. *Acta Sci.* 35: 271-276.
- Cruz CD, Regazzi AJ and Carneiro PCS (2012). Modelos biométricos aplicados ao melhoramento genético (volume 1). 4th edn. Universidade Federal de Viçosa, Viçosa.
- Cruz CD, Regazzi AJ and Carneiro PCS (2014). Modelos biométricos aplicados ao melhoramento genético (volume 2). 3rd edn. Universidade Federal de Viçosa, Viçosa.
- Embrapa - Empresa Brasileira de Pesquisa Agropecuária. (2011). Tecnologias de produção de soja - região central do Brasil 2012 e 2013. Embrapa Soja, Londrina.
- Duarte JB and Vencovsky R (1999). Interação genótipos x ambientes - uma introdução à análise "AMMI". Sociedade Brasileira de Genética, Ribeirão Preto.
- Fehr WR and Caviness CE (1977). Stages of soybean development: Iowa State University of Science and Technology Ames, Iowa.
- Ferreira DF (2002). Programa Estabilidade. Available at [[www.dex.ufpa.br](http://www.dex.ufpa.br)]. Accessed June 10, 2013.
- Gauch HG and Zobel RW (1996). AMMI analysis of yield trials. In: Genotype-by-environment proved and under what conditions this can be most environment interaction. (Kang MS and Gauch Júnior HG, eds), CRC, Boca Raton.
- Glaser JS, Sediyaama T, Cruz CD, Matsuo E, et al. (2015). Diversity of agronomic and molecular traits of the soybean cultivars with different levels of phakopsora. *Biosci. J.* 31: 25-36.
- Johnson RA and Wichern DW (1992). Applied multivariate statistical analysis. Englewood Cliffs, New Jersey.
- Maia MCC, Vello NA, Araújo LB, Dias CTS, et al. (2014). Genotype by environment interaction in soybean using three-mode principal component analysis. *Revista Agro@ambiente* 8: 104-111.
- Malosetti M, Ribaut JM and van Eeuwijk FAV (2013). The statistical analysis of multi-environment data: modeling genotype-by-environment interaction and its genetic basis. *Front. Physiol.* 4:1-17.
- Mandel JA (1971). A new analysis of variance model for non-additive data. *Technometrics* 13: 1-18.
- Mendonça O, Carpentieri-Pipolo V, Garbuglio DD and Fonseca Junior NS (2007). Factor analysis and environmental stratification in the assessment of soybean adaptability and stability. *Pesq. Agropec. Bras.* 42: 1567-1575.
- Meotti GV, Benin G, Silva RR, Beche E, et al. (2012). Épocas de semeadura e desempenho agrônomo de cultivares de soja. *Pesq. Agropec. Bras.* 47: 14-21.
- Murakami DM and Cruz CD (2004). Proposal of methodologies for environment stratification and analysis of genotype adaptability. *Crop Breed. Appl. Biotechnol.* 4: 7-11.
- Novais R, Ribeiro A, Guimarães P and Alvarez V (1999). Soja. In: Recomendações para uso de corretivos e fertilizantes em Minas Gerais - 5º Aproximação. (Ribeiro A, Guimarães P and Alvarez V, eds.). Editora UFV, Viçosa.
- Pelúzio JM, Fidelis RR, Giongo P, Silva JC, et al. (2008). Adaptabilidade e estabilidade de cultivares de soja em quatro épocas de semeadura no sul do Estado do Tocantins. *Rev. Ceres* 55: 34-40.

- Pinheiro LCM, GoodGod PIV, Faria VR, Oliveira AG, et al. (2013). Relationship in the selection for productivity and oil and protein contents in soybean using mixed models. *Pesq. Agropec. Brasil.* 48: 1246-1253.
- Polizel AC, Juliatti FC, Hamawaki OT, Hamawaki RL, et al. (2013). Phenotypical adaptability and stability of soybean genotypes in the State of Mato Grosso. *Biosci. J.* 29: 910-920.
- Ramalho MAP, Abreu AFB, Santos JB and Nunes JAR (2012). Aplicações da Genética quantitativa em plantas autógamas. UFLA, Lavras.
- Santi AL, Amado TJC, Cherubin MR, Martin TN, et al. (2012). Principal component analysis of soil chemical and physical attributes limiting grain yield. *Pesq. Agropec. Brasil.* 47: 1346-1357.
- Souza EC (2010). Os métodos biplot e escalonamento multidimensional nos delineamentos experimentais. Doctoral thesis. Escola Superior de Agricultura "Luiz de Queiroz", Piracicaba.
- Vencovsky R and Barriga P (1992). Genética biométrica no fitomelhoramento. Ribeirão Preto: Sociedade Brasileira de Genética, Ribeirão Preto.
- Yan W (2001). GGEbiplot: a Windows application for graphical analysis of multi-environment trial data and other types of two-way data. *Agron. J.* 93: 1111-1118.
- Yan W and Tinker NA (2006). Biplot analysis of multi-environment trial data: principles and applications. *Can. J. Plant Sci.* 1: 623-645.
- Yan W, Hunt LA, Sheng Q and Szlavnic Z (2000). Cultivar evaluation and mega-environment investigation based on the GGE biplot. *Crop Sci.* 40: 597 - 605.
- Yokomizo GK, Duarte JB, Vello NA and Unfried JR (2013). AMMI analysis of grain yield in soybean lines selected for resistance to Asian rust. *Pesq. Agropec. Bras.* 48: 1376-1384.
- Zobel RW, Wright MJ and Gauch HG (1988). Statistical analysis of a yield trial. *Agron. J.* 80: 388-393.