GMR

Identification of soybean genotypes with high stability for the Brazilian macro-region 402 via biplot analysis

E.U. Ramos Junior[1], R.L. Brogin[2], V.P.C. Godinho[3], F.J.E. Botelho[4], F.D. Tardin[5] and P.E. Teodoro[6]

[1]Embrapa Soja, Sinop, MT, Brasil
[2]Embrapa Soja, Vilhena, RO, Brasil
[3]Embrapa Rondônia, Vilhena, RO, Brasil
[4]Embrapa Rondônia, Porto Velho, RO, Brazil
[5]Embrapa Milho e Sorgo, Sinop, MT, Brazil
[6]Departamento de Ciências Agrárias,
Universidade Federal de Mato Grosso do Sul,
Chapadão do Sul, MS, Brasil

Corresponding author: F.D. Tardin
E-mail: flavio.tardin@embrapa.br

Genet. Mol. Res. 16 (3): gmr16039786
Received July 28, 2017
Accepted August 29, 2017
Published September 27, 2017
DOI http://dx.doi.org/10.4238/gmr16039786

ABSTRACT. Biplot analysis has often been used to recommend genotypes from different crops in the presence of the genotype x environment interaction (GxE). The objective of this study was to verify the association between the AMMI and GGE biplot methods and to select soybean genotypes that simultaneously meet high grain yield and stability to the environments belonging to the Edaphoclimatic Region 402, from Soybean Cultivation Region 4 (Mid-West), which comprises the Center North and West of Mato Grosso, and the southern region of Rondônia. Grain yield of 12 soybean genotypes was evaluated in seven competition trials of soybean cultivars in the 2014/2015 harvest. Significant GxE interaction revealed the need to use methods for recommending genotypes with adaptability and yield stability. The methods were complementary regarding the recommendation of the best genotypes. The AMMI analysis recommended MG/BR46 (Conquista) (G10) widely for all environments evaluated, whereas the BRY23-55012 (G9) and BRAS11-0149 (G2) were the most indicated genotypes by the GGE biplot method. However, the methods were concordant as to Porto Velho (PV1) environment that contributed least to the GxE interaction.

Key words: AMMI; GGE biplot; Genotype x environment interaction; *Glycine max*

INTRODUCTION

Brazil is currently the world's second largest producer of soybean, behind only the United States. This crop occupied, in the 2016/2017 harvest, an area of 33.85 million hectares, which totaled a production of 113 million tons, with average grain yield of 3338 kg/ha. In Mato Grosso, the largest national producer, this crop occupied an area of 9.32 million hectares, with a production of 30.5 million tons, with average grain yield of 3273 kg/ha (CONAB, 2017).

In Brazil, soybean is grown in several states, especially in the Midwest and South. Due to the great territorial extension of the country, different edaphoclimatic conditions are common among most of these sites. Besides, soybean has peculiarities, because it responds to the length of the day, which may interfere with its development depending on the latitude in which it is cultivated. In this sense, efforts are expended in the genetic breeding of the crop, aiming to establish cultivars adapted to the cultivation region (Mundstock and Thomas, 2005). As a way of validating new cultivars, since 1981, the Ministry of Agriculture, Livestock and Food Supply established the Brazilian System of Evaluation and Recommendation of Cultivars, and this has been improved, arriving at the way it was regionalized and how the value for cultivation and use assays of soybean lines were established. Mid-West region is formed by the Soybean Cultivation Macro-Region 4, where the Edaphoclimatic Region (REC) 402 comprises the Central North and West regions of Mato Grosso, as well as the southern region of Rondônia (Kaster and Farias, 2011).

This factor contributes to the occurrence of genotype x environment interaction (GxE), defined as the differential response of genotypes as a function of the environmental gradient. GxE interaction is one of the major bottlenecks of plant breeding, because it causes difficulties in recommending genotypes for a group of environment and/or region.

In the literature, there are several methodologies to investigate the GxE interaction, defined as the differential response of genotypes as a function of environmental variation (Cruz et al., 2014). Recently, two methodologies have been gaining prominence and are widely used by researchers: the AMMI model proposed by Gauch and Zobel (1988) and the GGE biplot model developed by Yan et al. (2000). Both analyses, based on biplot graphs, represent a data matrix.

The GGE biplot model jointly analyzes the effect of genotypes (G) with the effect of the GxE interaction, while the AMMI model separates G from the GxE interaction and, in the final step of the analysis, biplot graphs are built for the two methodologies. However, this separation is not capable of conferring superiority to the AMMI analysis (Gauch et al., 2008). Yan (2000) indicates the use of the GGE biplot model for identifying mega-environments, selecting representative and discriminating environments and indicating genotypes more adapted and stable to specific environments. On the other hand, the AMMI analysis can be used with efficiency in the identification of special environmental conditions for the farm (selection of cultivation sites) and superior average performance genotypes (Gauch et al., 2008).

The objective of this study was to verify the association between the AMMI and GGE biplot methods and to select soybean genotypes that simultaneously meet high grain yield and stability to the environments of the Central North and West regions of Mato Grosso and the southern region of Rondônia, known as REC 402.

MATERIAL AND METHODS

Seven competition trials of soybean cultivars were conducted in the 2014/2015 crop, whose soil and climatic characteristics are shown in Table 1. Experimental design was a randomized block design with 12 genotypes (BRAS10-0022, BRAS11-0149, BRS7481, BRS8581, BRSMG752S, BRSMG753C, BRSMG772, BRSMG812CV, BRY23-55012, BR46 (Conquista), TMG132RR, and TMG4182), with four replicates. Experimental unit consisted of four rows of 5.0 m long, spaced 0.50 m, and with a density of 15 plants/m. In each experimental unit, grain yield was evaluated in the two central rows, correcting the moisture to 13% and transforming to kg/ha.

Table 1. Characteristics of the seven environments belonging to REC 402 evaluated.

Environments	Abbreviation	Altitude (m)	Latitude	Longitude	Sowing time
Castanheiras	CA1	234	11°33'57"	61°55'33"	11/21/2014
Castanheiras	CA2	234	11°33'57"	61°55'33"	12/12/2014
Nova Mutum	NM	480	13°50'14"	55°41'44"	11/14/2014
Porto Velho	PV1	87	08°47'34"	63°50'53"	11/25/2014
Porto Velho	PV2	87	08°47'57"	63°51'02"	12/10/2014
Vilhena	VI1	607	12°47'16"	60°05'35"	11/06/2014
Vilhena	VI2	607	12°47'16"	60°05'35"	12/02/2014

Grain yield data were submitted to analysis of individual variance, considering the effect of treatments as fixed and the other effects as random. It was found that the relationship between the largest and the smallest mean square error (MSE) of the individual analysis of variance did not exceed the 7:1 ratio, thus allowing the joint analysis of the trials (Banzatto and Kronka, 2006). Subsequently, data were submitted to the adaptability and stability analysis through AMMI and GGE biplot methodologies.

To AMMI biplot analysis, we considered effects of genotypes and environments as fixed and the model according to Equation 1:

$$Y_{ij} = \mu + g_i + a_j + \sum_{k=1}^{n} \lambda_k \gamma_{ik} \alpha_{jk} + \rho_{ij} + \bar{\varepsilon}_{ij} \qquad \text{(Equation 1)}$$

where Y_{ij} is the mean response of the i-th genotype (i = 1,2, ..., G genotypes) at *j-th* environment (j = 1,2, ..., E environments); μ is the overall mean of the trials; g_i is the *i-th* genotype effect; a_j is the *j-th* environment effect; λ_k is the *k-th* singular value (scalar) of the original interaction matrix (denoted by GE); γ_{ik} is the element corresponding to the *i-th* genotype in the *k-th* singular column vector of the GE matrix; α_{jk} is the element corresponding to the *j-th* environment at *k-th* singular line vector of the GE matrix; ρ_{ij} the noise associated with the term $(ge)_{ij}$ of the classical i genotype with j environment interaction; $\bar{\varepsilon}_{ij}$ is the mean experimental error.

The GGE biplot model used is represented by Equation 2:

$$Y_{ij} - y_j = y_1\varepsilon_{i1}\rho_{j1} + y_2\varepsilon_{i2}\rho_{j2} + \varepsilon_{ij} \qquad \text{(Equation 2)}$$

where y_{ij} represents the mean yield of the population of i order in the environment of j order; y_j is the overall mean of genotypes in the j environment; $y_1\varepsilon_{i1}\rho_{j1}$ is the first principal component (PC1); $y_2\varepsilon_{i2}\rho_{j2}$ is the second principal component (PC2); y_1 and y_2 are the eigenvalues associated with IPCA1 and IPCA2, respectively; ε_1 and ε_2 are the values of PC1 and PC2, respectively, of the i genotype; ρ_{j1} and ρ_{j2} are the values of PC1 and PC2, respectively, for the j environment; and ε_{ij} is the error associated with the i-th genotype and j-th environment (Yan et al., 2000). Data analysis was performed using the Agricolae and GGEBiplotGui packages implemented in the R software (R Development Core Team, 2014).

RESULTS AND DISCUSSION

Joint analysis of variance (Table 2) revealed a significant effect of genotypes (G), environments (E), and GxE interaction ($P \leq 0.01$), which indicates contrasts between the environments and the occurrence of differential genotype performance over the environments. The presence of GxE interaction can be attributed to predictable factors such as soil type, pest and disease management, and unpredictable factors such as precipitation, temperature, and humidity in each environment. This can be confirmed by observing the characteristics of each environment (Table 1), which show differences regarding altitude, latitude, and longitude, as well as the climatic effects of precipitation and temperature provided by the different sowing times in the same place. Similar results were obtained by Carvalho et al. (2002), Silva and Duarte (2006), Rangel et al. (2007), and Peluzio et al. (2008), which also observed the presence of a GxE interaction for soybean yield in different regions of Brazil. The existence of the GxE interaction suggests the need to use adaptability and stability analysis since edaphoclimatic factors are the ones that most influence the adaptability and stability of soybean genotypes.

Table 2. Summary of the joint analysis of variance for grain yield of 12 soybean genotypes evaluated in 7 environments belonging to REC 402, Brazil.

Sources of variation	Degrees of freedom	Mean square
Blocks/Environments	21	186,939.14
Genotypes (G)	11	1,324,716.84*
Environments (E)	5	23,157,770.87*
GxE	66	413,091.20*
Error	231	116,791.61
Coefficient of variation (%)	-	14.78

*Significant at 1% probability by the F-test.

In the AMMI1 biplot (Figure 1), stability is interpreted in the ordinate axis, being considered stable the genotypes and environments with score values close to zero; adaptability is interpreted in the abscissa axis, where the phenotypic means of genotypes and environment are plotted (Gauch and Zobel, 1988). The genotype BRY23-55012 (G9) was the one that most contributed to GxE interaction, being the most unstable (Figure 1) because it presents the highest score (in absolute values) in the interaction axis. The genotypes BRAS10-0022 (G1), BRAS11-0149 (G2), BRSMG772 (G7), BRSMG812CV (G8), MG/BR46 (Conquista) (G10), and TMG4182 (G12) were the most stable because they have the smallest coordinate for the PC1 axis. Among these, G10 achieved a performance above the overall mean (Table 3), demonstrating wide adaptation to all environments, and can be recommended for the set of environments evaluated.

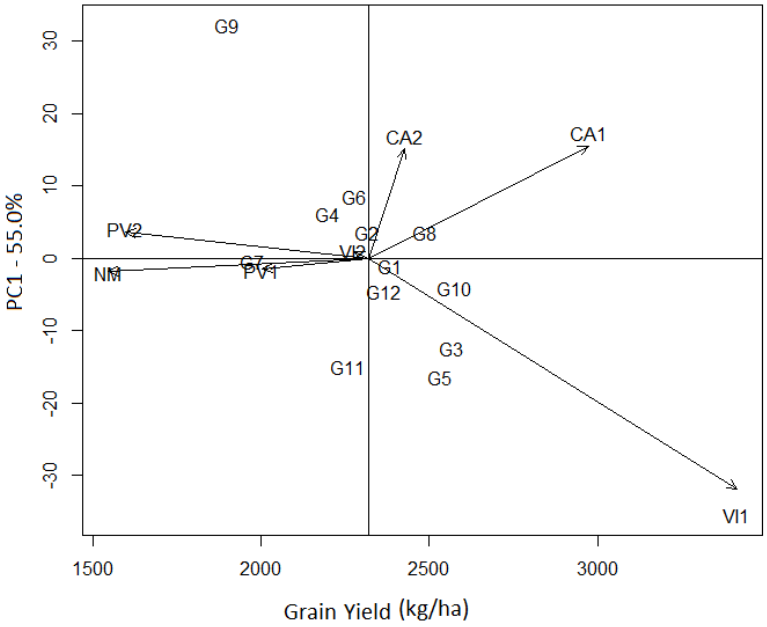

Figure 1. AMMI biplot for main effects and GxE interaction for grain yield of 12 soybean genotypes evaluated in 7 environments belonging to REC 402, Brazil.

Table 3. Means of grain yield (kg/ha) of 12 soybean genotypes for each tested environment and overall mean of genotypes and environments[1].

Genotype	ID	CA1	CA2	NM	PV1	PV2	VI1	VI2	Mean
BRAS10-0022	G1	2705.5	3204.3	1726.2	2424.8	2123.6	1412.5	1417.5	2144.9
BRAS11-0149	G2	2815.9	2268.0	1625.2	1786.9	3131.9	1899.2	1860.1	2198.2
BRS7481	G3	1820.8	2715.1	3237.3	2105.8	2697.5	1899.2	1657.3	2304.7
BRS8581	G4	2095.6	2530.7	2042.6	1501.9	1881.5	3506.1	3822.9	2483.1
BRSMG752S	G5	1436.2	1936.5	2869.5	2845.2	2099.0	2577.1	2480.3	2320.5
BRSMG753C	G6	3493.9	2319.7	2543.3	2496.4	1719.1	1742.2	3375.0	2527.1
BRSMG772	G7	2339.8	1774.1	1669.0	2668.6	3432.3	2516.3	2341.2	2391.6
BRSMG812CV	G8	3075.5	3990.8	2285.9	1778.0	2446.2	1754.7	1279.5	2372.9
BRY23-55012	G9	2374.5	2698.5	1727.9	1240.7	2983.5	3933.5	1264.1	2317.5
MG/BR46 (Conquista)	G10	1755.8	3402.6	4280.2	1561.3	2727.0	1998.3	1728.5	2493.4
TMG132RR	G11	2159.5	2220.3	2350.5	1455.2	880.4	2461.4	3847.3	2196.4
TMG4182	G12	1368.6	1132.8	2779.0	3001.3	1492.7	2096.5	2160.9	2004.5
Mean		2286.8	2516.1	2428.1	2072.2	2301.2	2316.4	2269.5	

[1] Environments detailed in Table 1.

The environment constituted by the first sowing time in Vilhena (VI1) is the most unstable (greater score magnitude in the PC1 axis), whereas the environments Nova Mutum (NM) and the two sowing times in Porto Velho (PV1 and PV2) are considered the most stable due to their lower coordinates for the PC1 axis. Environmental stability has great practical relevance for breeding programs, as it reports on genotype ordering reliability in a testing environment, regarding the classification for the tested environment means.

Figure 2 (AMMI2 biplot) allows identifying specific interactions between genotypes and environments and shows the stability zone that corresponds to the intercession point of the zero scores of the first and second interaction principal components (central region of the biplot). Genotypes and environments close to each other in any area of the graph represent a specific adaptation to the environment (Gauch and Zobel, 1988). In this context, specific interactions were observed between genotype BRAS10-0022 (G1) and the environments NM and VI2.

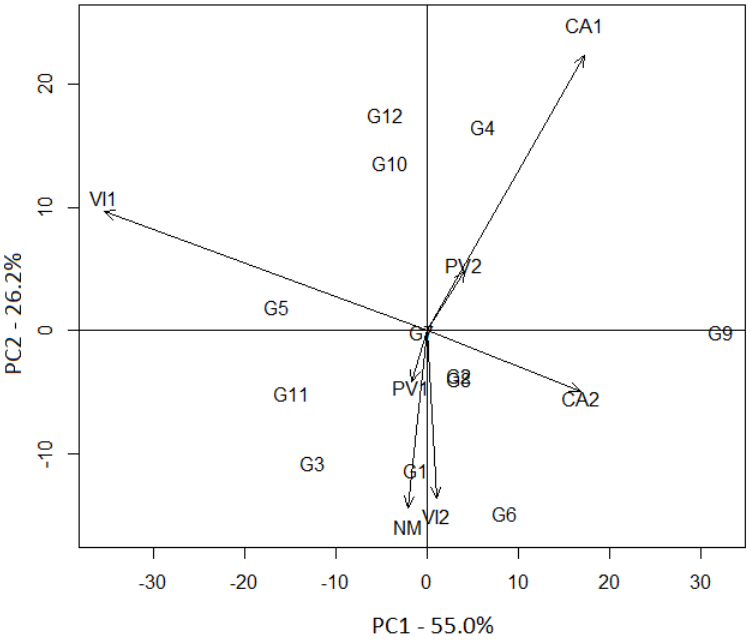

Figure 2. AMMI2 biplot for the first principal component of the interaction (PC1) x second principal component of the interaction (PC2) for grain yield of 12 soybean genotypes evaluated in 7 environments belonging to REC 402, Brazil.

Figure 3 shows a polygon connecting the genotypes BRY23-55012 (G9), BRS8581 (G4), TMG132RR (G11), TMG4182 (G12), MG/BR46 (Conquista) (G10), and BRSMG812CV (G8), which are furthest from the biplot origin. These genotypes have the largest vectors, which express the extent of genotype response to the tested environments. All other genotypes are contained within the polygon and have smaller vectors, that is, they are less sensitive concerning the interaction with the environments from each sector (Yan and Rajcan, 2002;

Karimizadeh et al., 2013). The vectors from the biplot center (0; 0), perpendicular to the polygon sides, divided the graph into six sectors. Similar results were observed by Mattos et al. (2013), who when evaluating the shoot yield of sugarcane genotypes via GGE biplot methodology observed the graph division into six sectors.

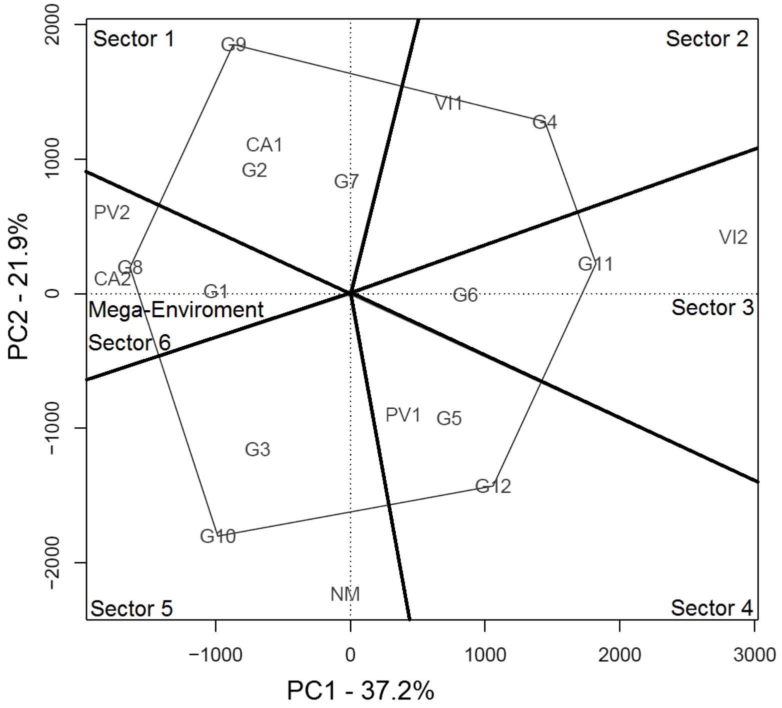

Figure 3. Sectors and mega-environments obtained by the GGE biplot model for grain yield of 12 soybean genotypes evaluated in 7 environments belonging to REC 402, Brazil.

According to Yan and Rajcan (2002), mega-environments are those sectors that contain two or more similar environments regarding soil-climatic conditions. It is possible to verify in Figure 3 that the environments PV2 and CA2 were the only ones that formed a mega-environment, where the genotype BRSMG812CV (G8) obtained means above the overall mean of these sites (Table 3). It is important to highlight here the difference between the methods used for the specific GxE interaction, in addition to the most similar environments. These results suggest that AMMI and GGE biplot methods can be used in a complementary way.

By the GGE biplot method, grain yield and genotype stability were evaluated from the average environmental coordination (AEC) expressed in Figure 4. Yang et al. (2009) emphasize that the greater the genotype projection in the AEC ordinate axis, the greater the genotype instability, representing a greater interaction with the environments. In this sense, the genotypes BRY23-55012 (G9) and BRAS11-0149 (G2) stood by phenotypic stability (Figure 3) and have also been defined as the optimal genotypes for this group of environments. This ideal genotype is graphically defined by the longest vector in PC1 and PC2 without projections, represented by the arrow in the center of the concentric circles (Yan and Rajcan, 2002). Although this genotype is another representative model, it is used as a reference for evaluating the genotypes.

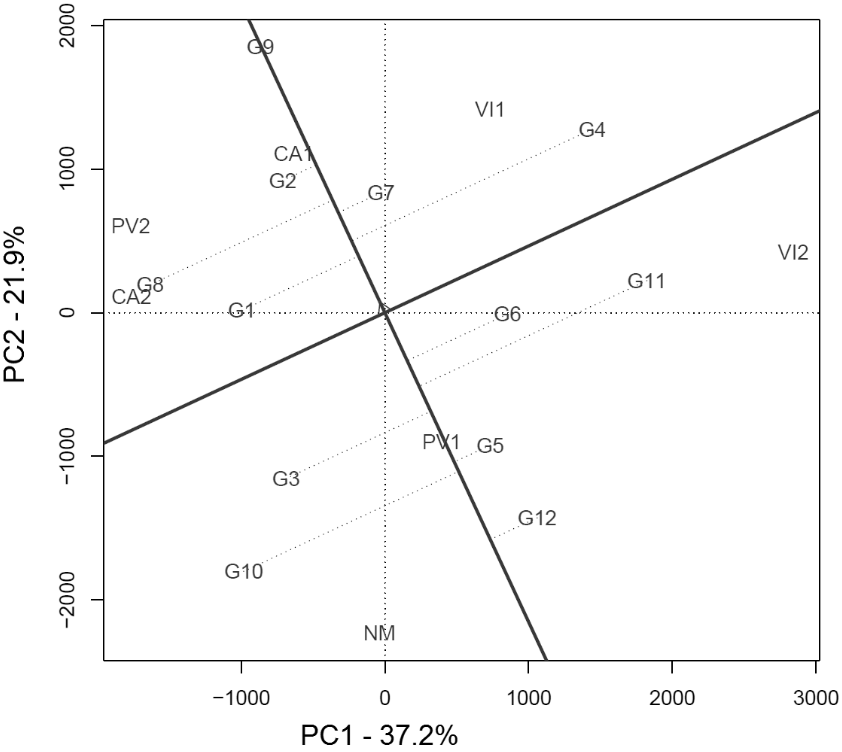

Figure 4. Mean versus stability according to the GGE biplot model for grain yield of 12 soybean genotypes evaluated in 7 environments belonging to REC 402, Brazil.

Figure 5 shows the relationship between grain yield and the stability from the vectorial viewpoint of the environments by GGE biplot analysis, where the environments are connected by vectors with the origin of the biplot. In environments with small vectors, yield stability is high (Yang et al., 2009). Thus, the first sowing time in Porto Velho (PV1) contributed less to the GxE interaction, agreeing with the results obtained by the AMMI method. According to Yang et al. (2009), an ideal environment should have a high PC1 score (greater genotype discriminating power regarding genotype main effects) and zero scores for PC2 (greater representativeness of all other environments). Likely the ideal genotype, ideal environment is only a theoretical concept and serves as a reference for site choice for multi-environment trials (Figure 6). Thus, the environment constituted by the second sowing time in Vilhena (VI2) is the one with the highest capacity to discriminate genotypes, favoring the selection of superior genotypes.

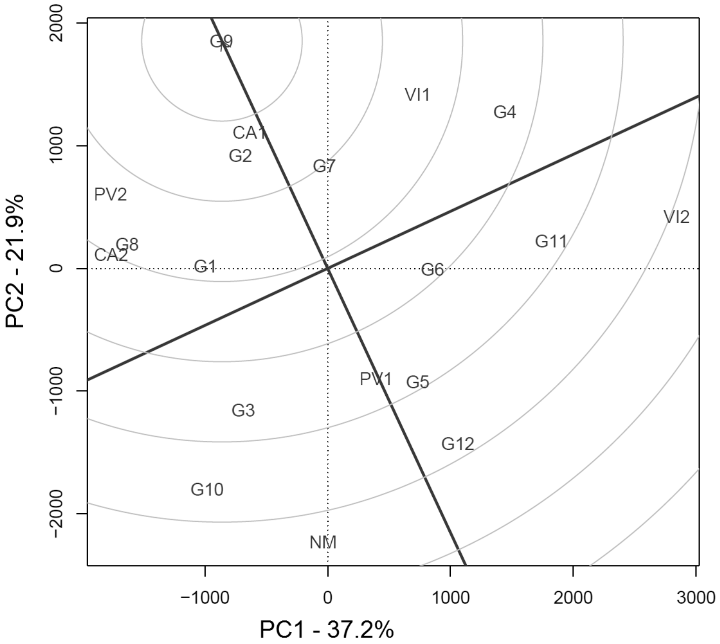

Figure 5. Classification of 12 soybean genotypes evaluated in 7 environments belonging to REC 402 according to GGE biplot model.

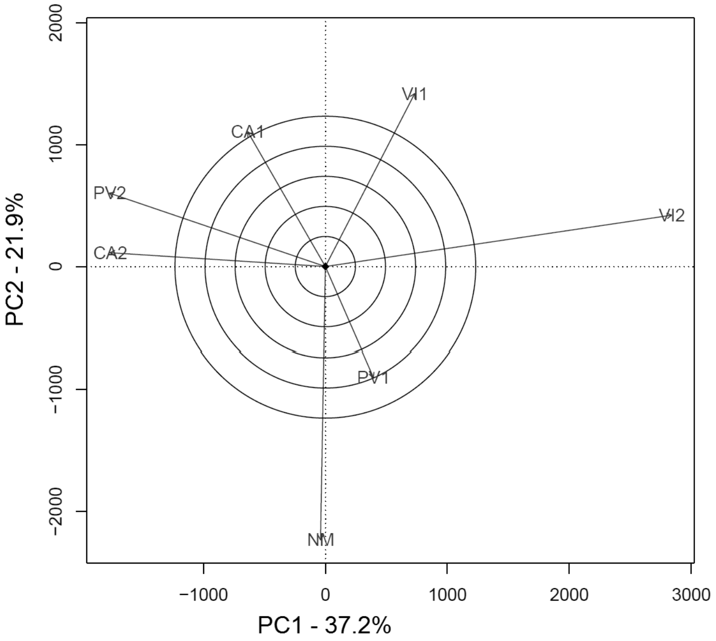

Figure 6. Relationship between the environments obtained by the GGE biplot model based on grain yield of 12 soybean genotypes evaluated in 7 environments belonging to REC 402, Brazil.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

Research supported by grants from Embrapa (Empresa Brasileira de Pesquisa Agropecuária) and CNPq (National Counsel of Technological and Scientific Development). F.D. Tardin is the recipient of a research fellowship from CNPq.

REFERENCES

Banzatto DA and Kronka SN (2006). Experimentação agrícola. 4a. ed. FUNEP, Jaboticabal.

Carvalho CGP, Arias CAA, Toledo JFF, Almeida LA, et al. (2002). Interação genótipo x ambiente no desempenho produtivo da soja no Paraná. *Pesqui. Agropecu. Bras.* 37: 989-1000. https://doi.org/10.1590/S0100-204X2002000700013

CONAB. Companhia Nacional de Abastecimento (2017). Acompanhamento de safra brasileiro - grãos: Oitavo levantamento, maio 2017 - safra 2016/2017. Companhia Nacional de Abastecimento Brasília. Available at [http://www.conab.gov.br/OlalaCMS/uploads/arquivos/17_05_12_10_37_57_boletim_graos_maio_2017.pdf]. Accessed May 22, 2017.

Cruz CD, Carneiro PCS and Regazzi AJ (2014). Modelos biométricos aplicados ao melhoramento genético. 3a. ed. Editora UFV, Viçosa.

Gauch HG, Jr. and Zobel RW (1988). Predictive and postdictive success of statistical analyses of yield trials. *Theor. Appl. Genet.* 76: 1-10. PubMed https://doi.org/10.1007/BF00288824

Gauch HG, Piepho HP and Annicchiarico P (2008). Statistical analysis of yield trials by AMMI and GGE: Further considerations. *Crop Sci.* 48: 866-889. https://doi.org/10.2135/cropsci2007.09.0513

Karimizadeh R, Mohammadi M, Sabaghni N, Mahmoodi AA, et al. (2013). GGE biplot analysis of yield stability in multienvironment trials of lentil genotypes under rainfed condition. *Not. Sci. Biol.* 6: 256-262.

Kaster M and Farias JRB (2011). Regionalização dos testes de Valor de Cultivo e Uso e da indicação de cultivares de soja: terceira aproximação. Embrapa Soja, Londrina. *Documentos* •••: 330.

Mattos PHC, Oliveira RA, Bespalhok Filho C, Daros E, et al. (2013). Evaluation of sugarcane genotypes and production environments in Paraná by GGE biplot and AMMI analysis. *Crop Breed. Appl. Biotechnol.* 13: 83-90. https://doi.org/10.1590/S1984-70332013000100010

Mundstock CM and Thomas AL (2005). Soja: fatores que afetam o crescimento e o rendimento de grãos. Evangraf/UFRGS, Porto Alegre.

Peluzio 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. *Ceres* 55: 34-40.

R Development Core Team (2014). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. Available at [http://www.R-project.org].

Rangel MAS, Minuzzi A, Braccini AL, Scapim CA, et al. (2007). Efeito do genótipo, ambiente e da interação no rendimento de grãos e nos teores de proteína de cultivares de soja. *Acta Sci. Agron.* 29: 351-354. https://doi.org/10.4025/actasciagron.v29i3.280

Silva WCJ and Duarte JB (2006). Métodos estatísticos para estudo de adaptabilidade e estabilidade fenotípica em soja. *Pesqui. Agropecu. Bras.* 41: 23-30. https://doi.org/10.1590/S0100-204X2006000100004

Yan W and Rajcan I (2002). Biplot evaluation of test sites and trait relations of soybean in Ontario. *Crop Sci.* 42: 11-20. PubMed https://doi.org/10.2135/cropsci2002.1100

Yan W, Hunt LA, Sheng QL and Szlavnics Z (2000). Cultivar evaluation and mega-environment investigation based on the GGE Biplot. *Crop Sci.* 40: 597-605. https://doi.org/10.2135/cropsci2000.403597x

Yang RC, Crossa J, Cornelius PL and Burgueño J (2009). Biplot analysis of genotype x environment interaction: proceed with caution. *Crop Sci.* 49: 1564-1576. https://doi.org/10.2135/cropsci2008.11.0665