Nonlinear Regression and Multivariate Analysis Used to Study the Phenotypic Stability of Cowpea Genotypes

Adriano dos Santos
Embrapa Agroenergia, Brasília, DF, Brazil

Francisco Eduardo Torres
Universidade Estadual de Mato Grosso do Sul, Aquidauana, MS, Brazil

Erina Vitório Rodrigues
Universidade de Brasília, Brasília, DF, Brazil

Ariane de André Pantaleão and Larissa Pereira Ribeiro Teodoro
Universidade Federal de Mato Grosso do Sul, Chapadão do Sul, MS, Brazil

Leonardo Lopes Bhering
Universidade Federal de Viçosa, Viçosa, MG, Brazil

Paulo Eduardo Teodoro
Universidade Federal de Mato Grosso do Sul, Chapadão do Sul, MS, Brazil

Additional index words. genotype × environment interaction, predictability, Vigna unguiculata

Abstract. This study aimed to evaluate the adaptability and phenotypic stability of cowpea genotypes using a nonlinear regression analysis and multivariate analysis. Experiments were performed at four sites in Brazil using a randomized blocks design with 20 treatments and four replications. The adaptability and stability of genotypes were evaluated by Toler nonlinear regression and genotype plus genotype × environment (GGE) biplot methodologies. Most of the genotypes revealed linear response patterns, with no differences regarding the favorable and unfavorable environments. Regarding the genotype classification for stability and adaptability, the Toler and GGE biplot methodologies are congruent. Genotypes MNC99-537F-4, MNC00-561G-6, MNC99542F-5, and Patativa have high overall adaptability and adequate yield. Therefore, they should be recommended for cultivation in the tested environments. Genotypes closer to the ideotype by the GGE biplot method are considered doubly desirable by the nonlinear method.

During the final stage of the cowpea [Vigna unguiculata (L.) Walp] breeding program, the genotypes must be evaluated in several environments and under similar conditions to those of actual cultivation (Freire Filho et al., 2011). Therefore, a new cultivar must have high yield and high performance of agronomic traits in a range of environmental conditions (Jarquin et al., 2017). At this final stage, the inconsistency in the genotype classification in multiple environments is one of the main challenges. This phenomenon is known as the genotypes × environment (G × E) interaction.

Therefore, the knowledge of the genotype × environment interaction is critical for making the correct decisions about cultivar recommendations. Studies of the genotypes × environment interaction is particularly essential for cowpea, because this crop is grown in a wide range of environments in Brazil. Therefore, it is necessary to study adaptability and yield stability of cowpea to obtain detailed information regarding the performance of genotypes under different environmental conditions (Freire Filho et al., 2011; Torres et al., 2016).

Several methods have been developed and applied by plant breeders to explain the genotype × environment interaction at the end of plant breeding programs. Among them, the Toler method (1990), which uses a nonlinear regression model in the parameters, offers alternatives to overcome the difficulties related to the estimate of the environmental index. Furthermore, it proposes a procedure that classifies genotypes into different groups according to the response pattern, statistically representing an improvement of the stability analyses in relation to the methods that use linear regression (Toler, 1990).

However, the genotype × environment interaction is too complex to be summarized by one or two stability parameters using univariate analyses. Therefore, multiplicative statistical approaches for interaction, such as the genotype plus genotype × environment (GGE) biplot method, are useful for studying genotypes × environments patterns and predicting the mean performance in specific environments (Karimizadeh et al., 2013). This study aimed to evaluate the adaptability and phenotypic stability of cowpea genotypes using the nonlinear regression analysis and multivariate analysis.

Materials and Methods

Experiments were performed in the municipalities of Aquidauana, Mato Grosso do Sul (lat. 20°27′12″ S; long. 55°40′06″ W; alt. 187 m asl.), Chapadão do Sul, Mato Grosso do Sul (lat. 18°46′25″ S; long. 52°37′24″ W; alt. 806 m asl.), Dourados, Mato Grosso do Sul (lat. 22°16′31″ S; long. 54°49′08″ W; alt. 407 m asl.), and Primavera do Leste, Mato Grosso (lat. 15°33′30″ S; long. 54°17′30″ W; alt. 636 m asl.). All sites are in the midwestern region of Brazil and belong to the Cerrado biome during the months of March and June.

The experiment was performed in a randomized blocks design with 20 treatments and four replications. The experimental unit consisted of four 5-m rows, spaced at 0.50 m apart. The two central rows were considered as the useful area. The experiment used the genotypes from Embrapa Meio-Norte. Seeds were manually sown. Thinning was performed at 2 weeks after the emergence of seedlings, leaving eight plants per meter and totaling 160,000 plants/ha during all years at all locations.

Initially, an individual analysis of variance (ANOVA) was performed for each environment to verify the homogeneity of the residual variance. Then, a joint ANOVA was performed considering the effects of genotypes and environments (combination of locations and years) to determine the possible genotype × environment interactions.

The adaptability and stability of cowpea genotypes were evaluated with the bisegmented nonlinear regression model presented by Toler (1990) and GGE biplot. The parameters of the model of Toler (1990), $\alpha_i$, $\beta_i$, $\gamma_i$, and $\mu_i$, are estimated jointly by the least-squares iterative process (nonlinear) using the modified Gauss-Newton method. During this process, the initial values for the iteration are provided by the estimate of the environmental index ($*$), as conceived by Eberhart and Russell (1966). The GGE biplot is a multivariate method in which the genotypic effects are subtracted from the matrix of means and, only then, is the principal component analysis performed. In summary, during this method, the effects of genotypes and genotype × environment interactions are treated as multiplicative. Only the genotype and the genotype × environment interaction are essential during this method, and they must be considered simultaneously. All analyses were performed using R software version 3.5.3 (R Foundation for Statistical Computing; Vienna, Austria).
Table 1. Observed mean (Y_\text{i}), estimates of parameters of the nonlinear model of Toler (\alpha, \beta_1, \beta_2) for grain yield (kg ha^-1) in 20 uppirt and semi-uirip cowpea genotypes and their respective response patterns (group).

| ID   | Genotypes       | Y_\text{i} | \alpha  | \beta_1  | B_\text{i2} | \beta_1-\beta_2 | \beta_{\text{common}} | Group |
|------|-----------------|----------|--------|---------|------------|-----------------|-----------------------|-------|
| 1    | MNC99-537F-1    | 725.58   | 725.58 | 0.95**  | 1.26**     | 1.05**          | C                     |       |
| 2    | MNC99-537F-4    | 891.92   | 891.92 | 0.60    | 2.20**     | 1.60**          | A                     |       |
| 3    | MNC99-541-F5    | 716.75   | 716.75 | 1.43**  | 0.43*      | 1.31**          | C                     |       |
| 4    | MNC99-541-F8    | 651.01   | 651.01 | 1.15**  | 1.31**     | 1.96**          | B                     |       |
| 5    | IT93K-93-10     | 514.18   | 514.18 | 0.44*   | 0.50**     | 0.50**          | D                     |       |
| 6    | Pratinho        | 433.20   | 433.20 | 1.72**  | -0.58**    | -2.30**         | E                     |       |
| 7    | Fradinho-2      | 638.64   | 638.64 | 1.09    | 0.79**     | -0.30**         | C                     |       |
| 8    | MNC99-519D-1-1-5| 671.86   | 671.86 | 1.06**  | 0.82**     | 0.54**          | C                     |       |
| 9    | MNC00-544D-10-1-2-2| 602.69 | 602.69 | 1.09    | 1.13**     | 0.04**          | C                     |       |
| 10   | MNC00-544D-14-1-2-2| 722.08 | 722.08 | 1.45    | 0.79**     | 0.67**          | E                     |       |
| 11   | MNC00-553D-8-1-2-2| 641.91 | 641.91 | 1.00    | 0.90**     | -0.11**         | C                     |       |
| 12   | MNC00-553D-8-1-2-3| 650.44 | 650.44 | 1.20    | 0.83**     | -0.37**         | C                     |       |
| 13   | MNC00-561G-1    | 690.61   | 690.61 | 0.30**  | 2.32**     | 2.03**          | C                     |       |
| 14   | EV X 63-10E     | 682.57   | 682.57 | 1.17**  | 0.70**     | -0.47**         | B                     |       |
| 15   | MNC99542F-5     | 881.23   | 881.23 | 0.45**  | 2.60**     | 2.55**          | A                     |       |
| 16   | EV X 91-2E-2    | 722.23   | 722.23 | 1.17    | 1.19**     | 0.02**          | A                     |       |
| 17   | MNC99-557F-2    | 494.64   | 494.64 | 0.74**  | 0.80**     | 0.06**          | B                     |       |
| 18   | BRS Guariba     | 667.20   | 667.20 | 1.20    | 0.84**     | 0.67**          | C                     |       |
| 19   | Patativa        | 753.34   | 753.34 | 0.48*   | 2.26**     | 2.17**          | A                     |       |
| 20   | Vita-7          | 496.39   | 496.39 | 1.33    | -0.25**    | -1.58**         | A                     |       |
| Overall mean | | 662.47 | | | | | | |

Table 2. Number of cowpea genotypes, classified according to yield levels and response patterns following the methodology of Toler (1990).

| Yield classes | A | B | C | D | E |
|---------------|---|---|---|---|---|
| Low yield     | 0 | 0 | 2 | 1 | 2 |
| Average yield | 0 | 2 | 8 | 0 | 0 |
| High yield    | 3 | 0 | 1 | 0 | 0 |
| Total of genotypes | 4 | 2 | 10 | 1 | 2 |
| % of genotypes | 20 | 10 | 55 | 5 | 10 |

Results and Discussion

A significant effect (P < 0.01) was observed for all sources of variation (Supplemental Table 1), indicating a high contrast in environments. This significance also indicated that the genotypes are not composed of a genetically homogeneous set, resulting in different yields. The significance of the genotype x environment interaction (combination of years and locations) indicated the need to study the phenotypic stability because it showed a difference in the response pattern of genotypes for the edaphoclimatic variations of years and locations. This result suggested that some genotypes or groups of genotypes have specific adaptation to environments, whereas others may show overall adaptation to all of them. Significant genotype x environment interactions were verified by Torres et al. (2016) for cowpea genotypes grown in Brazilian Cerrado.

The experimental precision was evaluated using the estimate of selective accuracy (SA). This parameter reflects the quality of the information and procedures used for predicting genetic values (Pimentel et al., 2014). Resende and Duarte (2007) classified the accuracy values, with 0.70 to 0.85 corresponding to high precision. In this context, a significant value of accuracy (0.74) was observed, which represented high precision in the identification and the possibility of success in the selection of individuals with specific or broad adaptation.

Genotypes grown in the environments Aquidauana I (E1), Chapadão do Sul (E2), and Dourados I (E3) reached grain yield means that were higher than the overall mean. Therefore, these environments can be classified as favorable for the development of the evaluated genotypes. The climatic conditions of each location are provided in Supplemental Table 2. This means that the environmental conditions were favorable, and the genotypes expressed their genes for grain yield, resulting in an adequate performance according to the positive environmental indices (Supplemental Table 2).

Conversely, Aquidauana II (E4), Dourados II (E5), and Primavera do Leste (E6) showed means that were lower than the overall mean and, consequently, a negative environmental index. Therefore, these environments were classified as unfavorable for the study genotypes. This probably occurred due to oscillations in the quantity and distribution of rainfall and the differences in these environments regarding other abiotic factors such as soil and temperature, which did not favor the expression of the genes for grain yield.

The genotypes showed no differences in favorable and unfavorable environments; therefore, they are included in groups B, C, and D (Table 1). However, genotypes MNC99-537F-4, MNC00-561G-6, MNC99542F-5, and Patativa showed significant results for \( \beta_1-\beta_2 \), with nonlinear responses, resulting in a bisegmented model. Notably, these genotypes had a convex response pattern (group A), which is considered more desirable according to Toler (1990). In this case, these genotypes were well-adapted to favorable environments because they were responsive in high-quality environments. In contrast, genotypes Pratinho and Vita-7 had a concave response pattern (group E), which was considered more undesirable by the method.

The performance of the genotypes classified in group A satisfied the adaptability of the ideal genotype concept (Table 2). These genotypes demand environmental quality to express their full genetic yield potential. However, to reach their yield potential, these genotypes also require the use of high technologies because under adverse conditions, such as a negative environmental index (e.g., Aquidauana II, Dourados II, and Primavera do Leste), these genotypes have low responsiveness. Therefore, these genotypes should not be recommended to producers who do not use high technology.

The genotypes classified in groups B and C do not differ statistically from 1, indicating that their performance follows the environment mean. They begin with low yield and reach satisfactory levels at the end. They also appropriately adjust the model to the data set, which results in high predictability. These groups of genotypes can adapt to different environments due to their phenotypic plasticity. Therefore, these genotypes are expected to show an average yield and low variability, even when cultivated in different environments, that is, they have high phenotypic stability.

Figure 1A shows the polygon connecting the farthest genotypes, whereas lines beginning at the origin of the biplot divided this polygon into five sectors. The environments grouped within these sectors were considered similar in relation to the environmental influence generated on the genotypes (Yan and Tinker, 2006). Genotype G15 (MNC99542F-5) was located in one of the vertices of the polygon in mega-environment 1, which was the most favorable for that group of environments, as well as genotype G11 (MNC00-553D-8-1-2-2) in sector 3 and genotype G14 (EV X 63-10E) in sector 1. Therefore, these genotypes have the highest yield in at least one of the environments and are among the genotypes with superior performance in the other environments of the group. Genotypes were also observed in sectors without any grouped environment. In this case, those
genotypes were considered unfavorable to the groups of tested environments, revealing low responsiveness and yield.

Genotype stability was evaluated using average environment coordination (AEC) (Fig. 1B). With this method, an ideal environment is defined from the mean of the scores of the principal components of all environments, represented by the small circle. The line that passes through the origin of the biplot and the ideal environment is known as the axis of the ideal environment and represents the AEC abscissa. The ordinate of the AEC is the line that passes through the origin of the biplot and is perpendicular to the AEC abscissa and indicates the higher effect of the genotypes × environments interactions and lower stability. This axis also separates the genotypes that are below and above the mean (Yan and Tinker, 2006; Yan et al., 2007).

The most productive genotypes farther away from the origin on the positive side of the AEC abscissa were G15 (MNC99542F-5), G2 (MNC99-537F-4), G19 (Pativativa), G13 (MNC00-561G-6), and G1 (MNC99-537F-1). Genotypes G3 (MNC99-541-F5), G16 (EV X 91-2E-2), and G10 (MNC00-544D-14-1-2-2) also showed a mean yield higher than the overall mean. Regarding genotype stability, as previously mentioned, the higher the projection in the ordinate of the AEC, the higher the instability of the genotype. Therefore, genotype G11 (MNC00-555D8-1-2-2) was the least stable for grain yield, with a higher projection of the AEC abscissa, followed by G14 (EV X 63-10E), G5 (IT92K-93-10), and G18 (BRS Guariba). However, these genotypes showed yield means close to the overall mean. These results corroborated with those obtained by the nonlinear method for these genotypes (classified in group C), that is, the simple linear response did not deviate from the mean response.

Conversely, genotypes G2 (MNC99-537F-4), G15 (MNC99542F-5), G3 (MNC99-541-F5), and G16 (EV X 91-2E-2) were relatively stable, showing lower projection in the AEC abscissa. The yield of these genotypes is higher than the overall mean; therefore, they are excellent choices for the study environments. Genotypes G20 (Vita-7) and G6 (Pretinho) also revealed high stability, which was associated with low yield. According to Toler (1990), they have a concave response and are more undesirable.

The ideal genotype must have a high mean yield and maintain this yield in all environments. This ideal genotype is graphically defined by the longest vector in the first
principal component (PC1) and without projections in the second principal component (PC2), represented by the arrow in the center of the concentric circles (Fig. 1C). Genotypes G15 (MNC99542F-5) and G2 (MNC99-537F-4) are contained in the second concentric circle and are the closest to the ideal genotype. Therefore, they can be considered desirable genotypes for yield and stability.

According to Yan and Tinker (2006), the ideal test environment (Fig. 1D) should discriminate the superior genotypes and be representative of all test environments. Therefore, the environment must have a high score for PC1 because this variable has greater discrimination capacity for the genotypes and a null score for PC2. Therefore, they should be recommended for cultivation in the tested environments. Genotypes closer to the ideotype using the GGE biplot method are considered most desirable by the nonlinear method.

Conclusions

The Toler and GGE biplot methodologies are congruent regarding the genotype classifications for stability and adaptability. Genotypes MNC99-537F-4, MNC00-561G-6, MNC99542F-5, and Patativa have high overall adaptability and adequate yield. Therefore, they should be recommended for cultivation in the tested environments. Genotypes closer to the ideotype using the GGE biplot method are considered most desirable by the nonlinear method.

Literature Cited

Eberhart, S.A. and W.A. Russell. 1966. Stability parameters for comparing varieties. Crop Sci. 6:36–40.

Freire Filho, F.R., V.Q. Ribeiro, M. de M. Rocha, K.J.D. Silva, M.S.R. Nogueira, and E.V. Rodrigues. 2011. Feijão-caupi no Brasil: Produção, melhoramento genético, avanços e desafios. Embrapa Meio-Norte, Teresina, PI.

Jarquín, D., C. Lemes da Silva, R.C. Gaynor, J. Poland, A. Fritz, R. Howard, S. Battenfield, and J. Crossa. 2017. Increasing genome-enabled prediction accuracy by modeling genotype-environment interactions in Kansas wheat. Plant Genet. 10:1–15.

Karimizadeh, R., M. Mohammadi, N. Sabaghni, A.A. Mahmoodi, B. Roustami, F. Seyyedi, and F. Akbari. 2013. GGE biplot analysis of yield stability in multi-environment trials of lentil genotypes under rainfed condition. Not. Sci. Biol. 5:256–262.

Pimentel, A.J.B., J.F.R. Guimarães, M.A. Souza, M.D.V. Resende, L.M. Moura, J.R.A.S.C. Rocha, and G. Ribeiro. 2014. Estimação de parâmetros genéticos e predição de valor genético aditivo de trigo utilizando modelos mistos. Pesqui. Agropecu. Bras. 49:882–890.

Resende, M.D.V. and J.B. Duarte. 2007. Precisão e controle de qualidade em experimentos de avaliação de cultivares. Pesqui. Agropecu. Trop. 37:182–194.

Toler, J.E. 1990. Patterns of genotypic performance over environmental arrays. Clemson, Thesis (PhD.). Clemson University. Clemson, South Carolina.

Torres, F.E., P.E. Teodoro, E.V. Rodrigues, A. Santos, A.M. Corrêa, and G. Cecconi. 2016. Simultaneous selection for cowpea (Vigna unguiculata L.) genotypes with adaptability and yield stability using mixed models. Genet. Mol. Res. 15:1–11.

Yan, W., M.S. Kang, B. Ma, S. Woods, and P.L. Cornelius. 2007. GGE Biplot vs AMMI analysis of genotype-by-environment data. Crop Sci. 47:643–653.

Yan, W. and N.A. Tinker. 2006. Biplot analysis of multi-environment trial data: Principles and applications. Can. J. Plant Sci. 86:623–645.
Supplemental Table 1. Summary of the joint analysis variance for grain yield (kg·ha$^{-1}$) of cowpea cultivated in Mato Grosso do Sul and São Paulo.

| Source of variation     | Degrees of freedom | Mean square |
|-------------------------|--------------------|-------------|
| Block/environment       | 18                 | 243,127.94  |
| Genotype (G)            | 19                 | 328,041.29**|
| Environment (E)         | 5                  | 12,574,956.75**|
| G × E                   | 95                 | 150,564.77**|
| Error                   | 342                | 31,709.49   |
| Selective accuracy      | —                  | 0.74        |
| Mean                    | —                  | 662.47      |

**Significant at ($P < 0.01$) according to the F test.

Supplemental Table 2. Overall means and environmental indices of cowpea genotypes according to the Toler method and climatic conditions of each environment.

| Location                  | Environment | $y_{ij}$ | Temperature avg ($^\circ$C) | Rainfall accumulated |
|---------------------------|-------------|----------|----------------------------|----------------------|
| Aquidauana I, MS          | E1          | 1155.25  | 28.21                       | 440.20               |
| Chapadão do Sul, MS       | E2          | 910.62   | 26.13                       | 550.60               |
| Dourados I, MS            | E3          | 924.79   | 25.25                       | 480.50               |
| Aquidauana II, MS         | E4          | 218.74   | 28.64                       | 300.50               |
| Dourados II, MS           | E5          | 210.53   | 25.41                       | 310.40               |
| Primavera do Leste, MT    | E6          | 554.90   | 27.10                       | 320.60               |

MS = Mato Grosso do Sul; MT = Mato Grosso.