REVISTA
CONTRIBUCIONES
A LAS CIENCIAS
SOCIALES

Principal component analysis for the selection of sunflower genotypes in 4 environments of the Cerrado of the Central Plateau

Análise de componentes principais para a seleção de genótipos de girassol em 4 ambientes do Cerrado do Planalto Central

Análisis de components principals para la selección de genotipos de girasol en 4 ambientes del Cerrado de la Planalto Central

DOI: 10.55905/revconv.18n.2-188

Originals received: 01/10/2025
Acceptance for publication: 02/03/2025

João Victor Pinheiro Melo
Master's Student in Agronomy
Institution: Universidade de Brasília
Address: Brasília – Distrito Federal, Brazil
E-mail: joaovictormelo29@gmail.com
Orcid: https://orcid.org/0000-0001-9115-5726

Renato Fernando Amabile
Doctor in Agronomy, with an emphasis on Genetics and Plant Breeding
Institution: Embrapa Cerrados
Address: Planaltina – Distrito Federal, Brazil
E-mail: renato.amabile@embrapa.br
Orcid: https://orcid.org/0000-0002-6943-5817

Arlini Rodrigues Fialho
Doctoral Student in Agronomy
Institution: Universidade de Brasília
Address: Brasília – Distrito Federal, Brazil
E-mail: arlinirf@gmail.com
Orcid: https://orcid.org/0000-0002-8658-6131

Gustavo Barbosa Cobalchini Santos
Master's Student in Agronomy
Institution: Universidade de Brasília
Address: Brasília – Distrito Federal, Brazil
E-mail: gustavocobalchini@gmail.com
Orcid: https://orcid.org/0009-0007-5706-9892

REVISTA
CONTRIBUCIONES
A LAS CIENCIAS
SOCIALES

Claudio Guilherme Portela de Carvalho
Doctor in Agronomy, with an emphasis on Genetics and Plant Breeding
Institution: Embrapa Soja
Address: Londrina – Paraná, Brazil
E-mail: portela.carvalho@embrapa.br
Orcid: https://orcid.org/0000-0003-0394-3752

Barbara Soares Aires França
Doctoral Student in Agronomy
Institution: Universidade de Brasília
Address: Brasília – Distrito Federal, Brazil
E-mail: baahf7@gmail.com
Orcid: https://orcid.org/0009-0002-0980-5318

Marcelo Fagioli
Doctor in Agronomy
Institution: Universidade de Brasília
Address: Brasília – Distrito Federal, Brazil
E-mail: mfagioli@unb.br
Orcid: https://orcid.org/0000-0002-4924-5004

Uirá do Amaral
Doctor in Agronomy, with an emphasis on Plant Sciences
Institution: Instituto Federal Goiano
Address: Urutaí – Goiás, Brazil
E-mail: uira.amaral@ifgoiano.edu.br
Orcid: https://orcid.org/0000-0003-2907-5586

ABSTRACT

The genetic improvement of sunflower is essential for the selection of adapted genotypes and the expansion of the crop. The aim of this study was to select adapted genotypes using morpho-agronomic characteristics, by means of principal component analysis (PCA), in four environments in the Cerrado of the Central Plateau. The experimental design used was a randomized block design with four replications, evaluating 8 sunflower genotypes in four Cerrado environments, at Embrapa Cerrados (CPAC), at the Plant Genetics Innovation Center (CIGV), at the experimental farm of the University of Brasilia (UnB), and in Cristalina-GO (CRIST). The following characteristics were evaluated: plant height, chapter diameter, grain yield and weight of a thousand achenes. The data obtained was submitted to the R program, a multivariate analysis was carried out using the PCA method and the means were grouped using the Scott-Knott test at 5% probability. The results of the PCA for the environments in relation to the morpho-agronomic characteristics show that the correlations between the variables indicated the existence of significant interactions between them. Among the genotypes, SYN 045 stood out in all the environments for yield, and for almost all of them for height, showing genetic superiority. The BRS G323, G57 and G46 genotypes also stood out in terms of height and weight of a thousand achenes, especially in the CIVGV, UnB and CRIST environments. The genotypes evaluated showed wide genetic variability, establishing their potential for integration into sunflower breeding programs. The most outstanding genotypes were SYN 045, BRS G323 and G57.

Keywords: *Helianthus annus* L., genetic variability, productivity, second crop.

RESUMO

O melhoramento genético do girassol é essencial para a seleção de genótipos adaptados e a expansão da cultura. O objetivo deste estudo foi selecionar genótipos adaptados utilizando características morfo-agronômicas, por meio da análise de componentes principais (PCA), em quatro ambientes do Cerrado do Planalto Central. O delineamento experimental utilizado foi o de blocos casualizados com quatro repetições, avaliando-se 8 genótipos de girassol em quatro ambientes de Cerrado, na Embrapa Cerrados (CPAC), no Centro de Inovação em Genética Vegetal (CIGV), na fazenda experimental da Universidade de Brasília (UnB) e em Cristalina-GO (CRIST). Foram avaliadas as seguintes características: altura da planta, diâmetro do capítulo, produtividade de grãos e peso de mil aquênios. Os dados obtidos foram submetidos ao programa R, foi realizada uma análise multivariada pelo método PCA e as médias foram agrupadas pelo teste de Scott-Knott a 5% de probabilidade. Os resultados da PCA para os ambientes em relação às características morfo-agronômicas mostram que as correlações entre as variáveis indicaram a existência de interações significativas entre elas. Entre os genótipos, o SYN 045 se destacou em todos os ambientes para produtividade e, em quase todos eles, para altura, mostrando superioridade genética. Os genótipos BRS G323, G57 e G46 também se destacaram em termos de altura e peso de mil aquênios, especialmente nos ambientes CIVGV, UnB e CRIST. Os genótipos avaliados apresentaram ampla variabilidade genética, estabelecendo seu potencial para integração em programas de melhoramento de girassol. Os genótipos que mais se destacaram foram SYN 045, BRS G323 e G57.

Palavras-chave: *Helianthus annus* L., variabilidade genética, produtividade, segunda safra.

RESUMEN

El mejoramiento genético del girasol es esencial para la selección de genotipos adaptados y la expansión del cultivo. El objetivo de este estudio fue seleccionar genotipos adaptados utilizando características morfoagronómicas, por medio del análisis de componentes principales (ACP), en cuatro ambientes del Cerrado de la Meseta Central. El diseño experimental utilizado fue un diseño de bloques al azar con cuatro repeticiones, evaluando 8 genotipos de girasol en cuatro ambientes de Cerrado, en Embrapa Cerrados (CPAC), en el Centro de Innovación en Fitogenética (CIGV), en la granja experimental de la Universidad de Brasilia (UnB), y en Cristalina-GO (CRIST). Fueron evaluadas las siguientes características: altura de la planta, diámetro del capítulo, rendimiento en grano y peso de mil aquenios. Los datos obtenidos se sometieron al programa R, se realizó un análisis multivariante mediante el método PCA y se agruparon las medias mediante el test de Scott-Knott al 5% de probabilidad. Los resultados del ACP para los ambientes en relación con los caracteres morfoagronómicos muestran que las correlaciones entre las variables indicaron la existencia de interacciones significativas entre ellas. Entre los genotipos, el SYN 045 destacó en todos los ambientes por su rendimiento, y en casi todos ellos por su altura, mostrando superioridad genética. Los genotipos BRS G323, G57 y G46 también se destacaron en altura y peso de mil aquenios, especialmente en los ambientes CIVGV, UnB y CRIST. Los genotipos evaluados mostraron una amplia variabilidad genética, estableciendo su potencial para la integración en programas de mejora de girasol. Los genotipos más destacados fueron SYN 045, BRS G323 y G57.

Palabras clave: *Helianthus annus* L., variabilidad genética, productividad, segundo cultivo.

1 INTRODUCTION

The sunflower (Helianthus annuus L.) is an annual plant, native to North America and widespread throughout the world. The species stands out for being a versatile plant with various uses, including the production of edible oil, seeds for human and animal consumption, and even in the production of biodiesel. In agriculture, as well as contributing to the production of food and renewable fuels, sunflower can be part of a system of nutrient cycling in a system of soil renewal plants, such as cover crops or green manure, improving the characteristics of the soil and diversifying agricultural practices (Lima Filho et al., 2014).

Due to its specific characteristics and adaptability, sunflower is grown in all regions of Brazil (SANTOS *et al.*, 2019). The cultivation of this species has shown remarkable growth in the country's regions, with potential for expansion in the Cerrado as an off-season crop (Dalchiavon *et al.*, 2016). In the Cerrado, it is common to plant a second crop between February and March. Because it has a deep and vigorous root system, sunflower is less demanding in terms of water availability, making it ideal for dealing with periods of water stress.

The Cerrado biome has ample biodiversity and remarkable prospects for agricultural growth, making it an important region for Brazil's economic development. Demand for sunflower has grown in the industrial and commercial sectors, driving prospects for expanding the cultivated area, especially in the Midwest region (Carvalho *et al.*, 2018). According to CONAB (2023), approximately 65% of the area planted with sunflower in Brazil belongs to the Federal District and Goiás, highlighting the practice of planting it as a second crop, in succession to the soybean crop.

The successful integration of sunflower into the Brazilian production system is linked, among other factors, to the use of genotypes that are better adapted to the growing areas (Matssura *et al.*, 2015). These genotypes favour agricultural practices, minimize the risk of production losses and increase profitability for producers. Therefore, encouraging genetic improvement programs in Brazil is essential, as it is a viable tool for obtaining increasingly productive cultivars for farmers.

The aim of this study was to evaluate morpho-agronomic characteristics using principal component analysis and to select genotypes adapted to cultivation in the Cerrado.

2 METHODOLOGY

The experiments were conducted in 2019 at four locations in the Cerrado of the Central Plateau: in the experimental area of Embrapa Cerrados (CPAC), Planaltina – DF, located at 15°35'30” South latitude and 47°42'30” West longitude, at an altitude of 1. 007 m; at the Centro de Inovação em Genética Vegetal - Fazenda Sucupira (CIGV), Riacho Fundo II – DF, at 15°54'53” South latitude and 48°02'14” West longitude, at an altitude of 1. 254 m; at the Água Limpa Farm of the University of Brasilia (UnB), in the Vargem Bonita Nucleo Rural – DF, at 15°56'00” South latitude and 47°55'00” West longitude, with an average altitude of 1,100 m; and in the municipality of Cristalina – GO, at 16°08' South latitude and 47°29' West longitude, with an average altitude of 980 m.

The soil in the locations where the experiments took place was classified as typical dystrophic Latossolo argiloso, and the climate was classified as Tropical Seasonal (Aw), according to Köppen (Santos *et al.*, 2018). The base and top dressing fertilizations were made according to the interpretation of the soil analysis results and the crop recommendation given by Leite *et al.* (2005).

The experiment was conducted in a randomized block design with four replications. Eight genotypes were evaluated: G46, G53, G54, G55, G56, G57 and the control BRS G323 and SYN 045. The agronomic characteristics evaluated were: plant height (ALT) in cm, chapter diameter (DC) in cm, grain yield (REND) in kg ha-1, and weight of a thousand achenes (PMA) in g.

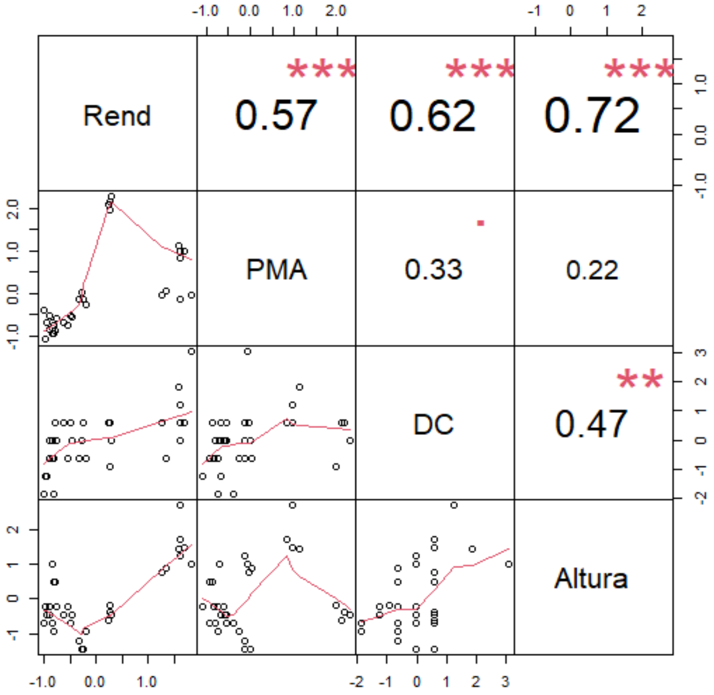
The data obtained was submitted to the R program (R Core Team, 2022) and multivariate analysis was carried out using the principal component analysis method and the means were compared using the Scott-Knott test at 5% probability.

3 RESULTS AND DISCUSSIONS

The CPAC environment, for the principal component analysis of sunflower morpho-agronomic traits, showed a Vector value of 2.51 for Principal Component 1 (PC1) and 62.64% for the percentage explained; in Principal Component 2, the Vector value was 0.80 and the percentage explained was 19.99%. The highest loadings in the First Dimension are mainly associated with the variables analyzed, yield (0.938), weight of a thousand achenes (0.639), chapter diameter (0.775) and plant height (0.786), indicating that these variables contribute significantly to the direction and magnitude of the variations captured by PC1. The characteristic weight of a thousand achenes had a positive influence on the construction of the principal component 2 vectors, with a value of 0.751. The genotypes that stood out most in the formation of the vector for the yield trait were the control SYN 045 and G46.

All the characteristics that correlated with yield had a significant correlation. The interaction between yield and height (0.72) had a strong correlation, while the interactions with the weight of a thousand achenes (0.57) and chapter diameter (0.62) had moderate correlations (Figure 1).

Figure 1. Correlogram of Pearson's correlations for the CPAC environment.



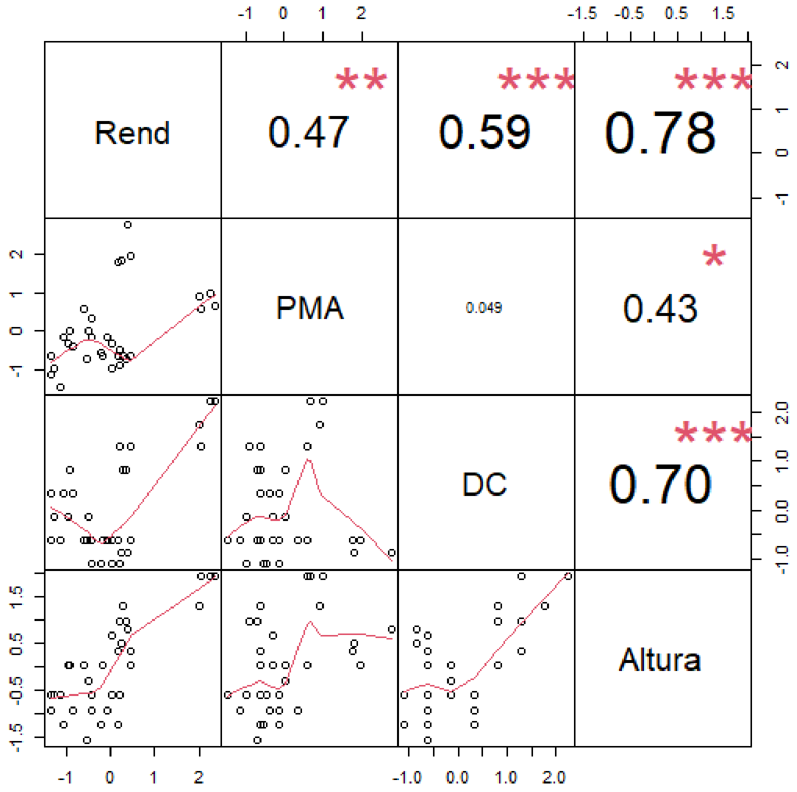
Fonte: Elaborada pelos próprios autores.

The yield characteristic showed the formation of 5 groups, with SYN 045 (2,843.50 kg ha-1), G46 (2,775.12 kg ha-1) having the highest yields. The weight of a thousand achenes resulted in the formation of 5 similarity groups and the highest values for this characteristic were found in the BRS G323 (56.33 g) and SYN 045 (48.00 g) controls. There were no statistically significant differences in the chapter diameter data, so SYN 045 (16.75 cm) and G46 (16.25 cm) were the materials with the largest chapter sizes. The materials with the highest plant height values were SYN 045 and G46 with 177.25 cm and 159.50 cm respectively, and the characteristic showed the formation of 5 similarity groups.

The CIGV environment showed a Vector value of 2.59 for principal component 1 (PC1) and 64.64% for the percentage explained, while in principal component 2 the Vector value was 0.96 and the percentage explained was 23.98%. The highest loadings in the First Dimension are mainly associated with the variables analyzed: yield (0.908), chapter diameter (0.768) and plant height (0.934), indicating that these variables contribute significantly to the direction and magnitude of the variations captured by PC1. The characteristic weight of a thousand achenes had a positive influence on the construction of the principal component 2 vectors, with a value of 0.805. The genotype that stood out most in the formation of the vector for the yield and height traits was the control SYN 045.

All the characteristics that correlated with yield had a significant correlation (Figure 2). The interaction between yield and height (0.78) had a strong correlation, while the interactions with the weight of a thousand achenes (0.47) and chapter diameter (0.59) had moderate correlations. The interaction between plant height and diameter showed a strong correlation of 0.70.

Figure 2. Correlogram of Pearson's correlations for the CIGV environment.



Fonte: Elaborada pelos próprios autores.

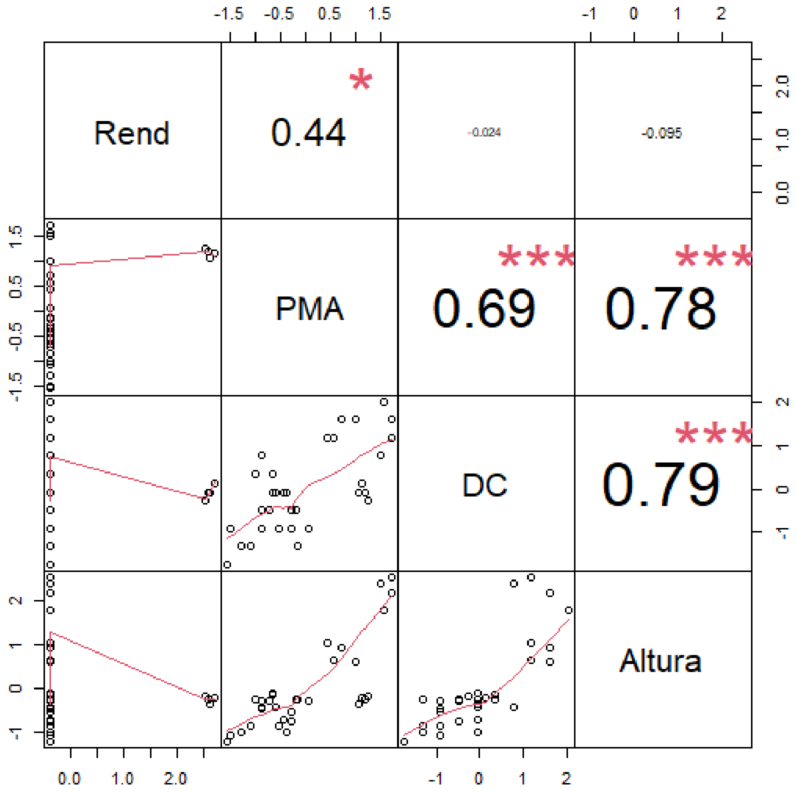
Analysis of the genotypes' characteristics in the CIGV environment highlighted the control SYN 045 as the material with the highest average for grain yield, shoot diameter and plant height, with 3,277.75 kg ha-1, 18.25 cm and 147.50 cm respectively. In the characteristic weight of a thousand achenes, the controls BRS G323 (61.56 g) and SYN 045 (53.75 g) had the best averages. As for the groups formed, yield showed the formation of 7 groups, thousand seed weight and chapter diameter 5 groups and plant height obtained 3 groups of similarity.

At UnB, principal component 1 (PC1) had a Vector value of 2.53 and 63.28% for the percentage explained, while in principal component 2, the Vector value was 1.17 and the percentage explained was 29.20%. The highest loadings in the First Dimension are mainly associated with the variables analyzed, weight of a thousand achenes (0.926), chapter diameter (0.889) and plant height (0.919), indicating that these variables contribute significantly to the direction and magnitude of the variations captured by PC1. The yield characteristic had a positive influence on the construction of the principal component 2 vectors, with a value of 0.973. The genotypes that stood out most in the formation of the vector for the yield trait were the control

BRS G323, while for the chapter diameter and plant height traits the SYN 045 and G57 materials stood out more than the others.

The interactions (Figure 3): weight of a thousand achenes and chapter diameter (0.69), weight of a thousand achenes and height (0.78) and between chapter diameter and height (0.79) were significant and strong. The correlation between yield and thousand-seed weight (0.48) was significant and moderate, but the other interactions with this characteristic were weak and not significant.

Figure 3. Correlogram of Pearson's correlations for the UnB environment.



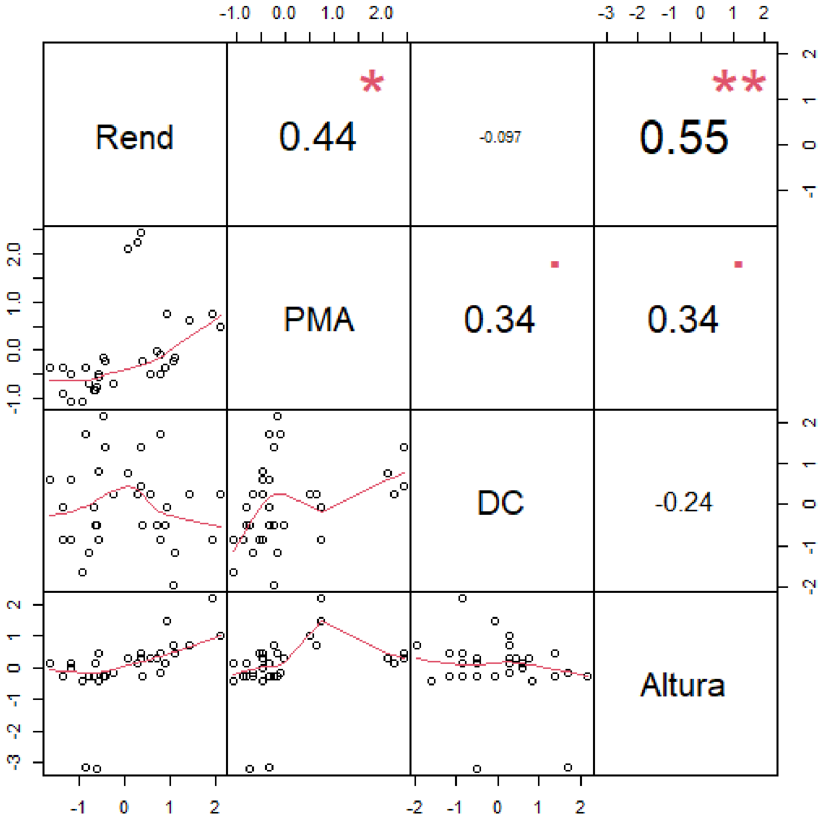
Fonte: Elaborada pelos próprios autores.

In the formation of the clusters, yield obtained 6 groups, weight of a thousand achenes presented 7 groups, while chapter diameter 3 and plant height 4. The analysis of the characteristics of the genotypes in the UnB environment highlighted the control SYN 045 as the material with the highest average for grain yield, weight of a thousand achenes, chapter diameter and plant height with 2,600.98 kg ha-1, 52.37 g, 17.50 cm and 188.00 cm respectively.

For CRIST, principal component 1 had a Vector value of 1.89 and an explained percentage of 47.25%. Principal component 2 showed values of 1.29 and 32.31% for the Vector value and percentage explained respectively. In the first dimension, the characteristics yield (0.855), weight of a thousand achenes (0.71) and height (0.809) made the greatest significant contribution to the direction and magnitude of the variations captured. Chapter diameter (0.934) had a positive influence on the construction of the vectors of principal component 2. The BRS G323 control had the greatest contribution to the vector of the thousand-seed weight variable, while the SYN 045 control contributed to the vector of the yield variable and the G53 genotype to the chapter diameter.

Only the interactions between yield and weight of a thousand achenes (0.44) and between yield and height (0.55) had moderate and significant correlations (Figure 4).

Figure 4. Correlogram of Pearson's correlations for the CRIST environment.



Fonte: Elaborada pelos próprios autores.

The results obtained by principal component analysis (PCA) reveal that the height variable showed the highest correlation with the yield of sunflower genotypes (correlation of 0.72, highly significant). This indicates that taller plants tend to be associated with higher grain yields. Mengistu and Tesema (2024), in their study, found a lower correlation than this (0.491), which was still significant, demonstrating that in addition to genetic and environmental factors, plant height influences grain yield. In contrast, the study by Duca *et al.* (2023) found a correlation ranging from 0.54-0.78 between plant height and seed yield under biotic and abiotic stress, suggesting that both the genetic characteristics of the hybrids studied and the environmental conditions influence this relationship.

Furthermore, the average yield found (which ranged from 2,436.54 kg.ha-1 to 3,277.75 kg.ha-1) in all environments is similar to the average observed by Neto *et al.* (2000), in studies with different sunflower genotypes in the Cerrado biome, in which the achene yield ranged from 2,507.8 kg.ha-1 to 3,138.7 kg.ha-1. Capone *et al.* (2012), in a study of five other sunflower cultivars in Tocantins, found that yields ranged from 2,300 kg.ha-1 to 3,000 kg.ha-1, highlighting the influence of the planting season on the yield of these genotypes in the region.

The chapter diameter variable also showed a significant correlation (0.62) with yield. This result was to be expected, since larger chapters tend to contain more seeds, which has a direct impact on yield. However, the moderate correlation suggests that other variables, such as sowing density or environmental conditions, may also influence this characteristic. Other studies with 40 and 100 sunflower genotypes have shown a robust relationship between chapter diameter and grain yield, indicating that this is a trait that contributes positively to productivity and should be evaluated by breeding programs (RANI *et al.*, 2016) (Pandya *et al.*, 2015).

On the other hand, the variable weight of a thousand achenes showed a low correlation with yield (r = 0.44 - 0.57) for two of the four environments studied. With regard to the other variables analyzed, only the UnB environment showed a significant correlation. This shows that the environment influenced the weight of the achenes. According to Yasin and Singh (2010), who carried out studies with 14 sunflower genotypes to understand the influence of certain characteristics on yield, the weight of a thousand achenes, the number of seeds and the diameter of the chapter have a high effect on yield, as observed in the UnB environment.

In general, the first two principal components explained most of the variability in the data, indicating that the variables studied are effective in discriminating between genotypes and environments. This shows that these characteristics, both in this study and in the others mentioned above, are effective indicators for improving sunflower performance and productivity.

4 CONCLUSION

With the data obtained from the morpho-agronomic characterization and through statistical analysis, it was possible to verify that there is genetic variability among the genotypes analyzed, making it possible to select genotypes with good adaptability and agronomic characteristics to be part of the sunflower improvement program.

Through Pearson's Correlation carried out on the sunflower genotypes in each environment, all the agronomic characteristics analyzed had an influence on the construction of the main components.

The genotypes that stood out most in the environments analyzed, especially in terms of grain yield and weight of a thousand achenes, were: SYN 045, BRS G323 and G57.

REFERENCES

CAPONE, A.; BARROS, H. B.; SANTOS; E. R. dos; FERRAZ, E. de C.; SANTOS, A. F. dos; FIDELIS. R. R. Influência de diferentes épocas de semeadura no desempenho agronômico de cultivares de girassol no cerrado tocantinense/Influence of different times of sowing on the agronomic performance of sunflower cultivars in cerrado tocantinense. **Bioscience Journal**, v.28, p.136-144, 2012.

CARVALHO, C.G.P.; CALDEIRA, A.; CARVALHO, L.M. de; CARVALHO, H.W.L. de; RIBEIRO, J.L.; MANDARINO, J.M.G.; RESENDE, J.C.F, de; SANTOS, A.R. dos; SILVA, N.R. da; ARRIEL, N.H.C. Fatty acid profile of sunflower achene oil from the brazilian Semi-arid region. **Journal of Agricultural Science**, v.10, p.144-150, 2018.

CONAB - COMPANHIA NACIONAL DE ABASTECIMENTO. **Acompanhamento de safra brasileiro – grãos.** Décimo primeiro levantamento, agosto 2023 - safra 2022/2023. Brasília: CONAB. 2023.

DALCHIAVON, F.C.; CARVALHO, C.G.P. de; AMABILE, R.F.; GODINHO, V. de P.C.; RAMOS, N.P.; ANSELMO, J.L. Características agronômicas e suas correlações em híbridos de girassol adaptados à segunda safra. **Pesquisa Agropecuária Brasileira**, v.51, p.1806-1812, 2016.

DUCA, M.; MUTU, A.; PORT, A.; CLAPCO, S. Genotype-environment interaction in the variability of yield associated indices under stress conditions in sunflower. **Helia**, v.46, n.79, p.201-214, 2023.

LEITE, R.M.V.B.de C.; BRIGHENTI, A.M.; CASTRO, C. **Girassol no Brasil**. Londrina: Embrapa Soja, 2005.

DE LIMA FILHO, O.F.; AMBROSANO, E.J.; ROSSI, F.; CARLOS, J.A.D. (Eds.). **Adubação verde e plantas de cobertura no Brasil: fundamentos e prática.** Brasília: Embrapa. 2014.

MATSSURA, M.I.S.F.; DIAS, F.R.T.; PICOLI, J.F.; LUCAS, K.R.G.; CASTRO, C.; HIRAKURI, M.H. Avaliação do ciclo de vida do sistema de produção soja-girassol no cerrado brasileiro, In: Embrapa Pantanal-Artigo em anais de congresso. In: **REUNIÃO NACIONAL DE PESQUISA DE GIRASSOL, 21; SIMPÓSIO NACIONAL SOBRE A CULTURA DO GIRASSOL**, nº 9, 2015. Londrina, **Anais,** Londrina: Embrapa Soja, 2015, 201p.

MENGISTU, B.; TESEMA, T. Analyzing sunflower trait patterns using MANOVA, PCA, and correlation across seasons and locations. **American Journal of Biological and Environmental Sciences**, v.10, n.3, p.87-95, 2024.

NETO, A.L. de F.; AMABILE, R.F.; AZEVEDO, J. De A.; FONSECA, C.E.L. Da; CASTIGLIONI, V.B. Avaliação de variedades de girassol nos cerrados do Distrito Federal. **Revista Ceres,** v.47, n.237, p.469-482, 2000.

PANDYA, M.M.; PATEL, P.B.; NARWADE, A.V. A study on correlation and path analysis for seed yield and yield components in Sunflower [Helianthus annuus (L.)]. **Electronic Journal of Plant Breeding**, Navsari, v.7, n.2, p.177-183, 2015.

R CORE TEAM. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Austria, 2022.

RANI, R.; SHEORAN, R.K.; CHANDER, S. Association analysis for yield and component traits in sunflower (Helianthus annuus L.). **Journal of Oilseeds Research**, Hisar, v.33, n.3, p.129-133, 2016.

SANTOS, H.G.; JACOMINE, P.K.T.; ANJOS, L.H.C.; OLIVEIRA, V.A.; LUMBRERAS, J.F.; COELHO, M.R.; DE ALMEIDA, J.A.; ARAÚJO FILHO, J.C.; OLIVEIRA, J.B.; CUNHA, T.J.F. **Sistema brasileiro de classificação de solos.** 5.ed. Brasília: Embrapa, 2018.

SANTOS, J.M.S.; PEIXOTO, C.P.; SILVA, M.R.; ALMEIDA, A.T.; CASTRO, A.M.P.B. Agronomic and productive characteristics of sunflower intercropped with forage in a crop-livestock integration system. **Revista Caatinga**, v.32, n.2, p.514-525, 2019.

YASIN, A.B.; SINGH, S. Correlation and path coefficient analyses in sunflower. **Journal of Plant Breeding and Crop Science**, v.2, n.5, p.129-133, 2010.