Agronomy Science and Biotechnology

Research Article

Multivariate approach applied to phenotypic traits as a function of the selection of soybean cultivars

Eduardo Ely Foleto[1], Alexandre Kaue Foguesatto Ottonelli[1], Ivan Ricardo Carvalho[1,*], José Antonio Gonzalez Da Silva[1], Gerusa Massuquini Conceição[1], Willyan Júnior Adorian Bandeira[1], Gabriel Mathias Weimer Bruinsma[1] and Jaqueline Piesanti Sangiovo[1]

[1]Universidade Regional do Noroeste do Rio Grande do Sul, Departamento de Estudos Agrários, Avenida do Comércio, nº 3.000, Bairro Universitário, Ijuí, RS, Brazil, CEP 98700-000.
*Corresponding author, E-mail: carvalhoirc@gmail.com

Citation: Foleto, E. E., Ottonelli, A. K. F., Carvalho, I. R., Silva, J. A. G., Conceição, G. M., Bandeira, W. J. A., Bruinsma, G. M. W., & Sangiovo, J. P. (2024). Multivariate approach applied to phenotypic traits as a function of the selection of soybean cultivars. *Agronomy Science and Biotechnology*, 10, 1-16 https://doi.org/10.33158/ASB.r205.v10.2024

Received: July 15, 2024.
Accepted: July 31, 2024.
Published: October 7, 2024.

English by: Ivan Ricardo Carvalho

Abstract

The objective of this work was to identify superior soybean cultivars through a multivariate approach applied to phenotypic traits. This study was developed in the 2023/2024 agricultural harvest, in the experimental area of the Regional Institute for Rural Development, at UNIJUÍ. It is located in the municipality of Augusto Pestana, in the state of Rio Grande do Sul, Brazil. The experimental design used was randomized blocks with internal blocks, with the treatment being ten cultivars and five replications. The trial of ten cultivars in the northwest of the state of Rio Grande do Sul demonstrated that the soybean cultivar with the highest yield was NS5922IPRO, with 5235.3 kg ha^{-1}. The multivariate approaches formed two groups to explain the factors that influenced yield, where the first was discrepant for the variables *Euschistus heros*, phytotoxicity, *Fusarium solanie*, *Macrophomina phaseolina*, *Conyza bonariensis*, production zone area, number of total nodes in the branch, branch number, root length, number of vegetables with 4 grains, number of vegetables with 0 grains and vegetable grain weight of 2 grains. The second similar group for the variables *Diabrotica speciosa*, *Caliothrips brasiliensis*, *Euschisthus heros*, *Phakopsora pachyrhizi* and *Cercospora sojina*, area of production zone, number of vegetables with zero grains. The trial of ten cultivars in the northwest of the state of Rio Grande do Sul demonstrated that the soybean cultivar with the highest yield was NS5922IPRO, with 5235.3 kg ha^{-1}. The multivariate approaches formed two groups to explain the factors that influenced grain yield.

Keywords: *Glycine max*; compound of yield; vegetable protein; biofuel production; similarity indices; quantitative trait.

Introduction

Soybean (*Glycine max* L. Merrill) is one of the most important agricultural crops worldwide, being widely cultivated for its versatility and economic value. Belonging to the Fabaceae family, it is a species originating in China, domesticated for over three thousand years, and widely used worldwide for its protein content, with its production continuing to grow over the years, due to the increase in cultivated area in some countries and the yield per area, leveraged by the use of technological advances (Minuzzi, Frederico, & Silva, 2017). In 2022/2023 alone, the production of the world oilseed agricultural harvest was 322.8 million tons of grains, 18.4% higher compared to the previous harvest (Food and Agricultura Organization [FAO], 2023).

Seen as the second largest commodity in the world, behind only sugar cane, soy is the main crop of the Brazilian economy and is the main vegetable protein in the world (Gazzoni & Dall'agnol, 2018). Its leadership in Brazilian agriculture is due to the economic return and diverse use of grains, which can be occupied by industry, vegetable oil production or biofuel production, and mainly as a source of protein for animal husbandry (França-Neto et al., 2016).

Its so-called quantitative traits, such as grain yield, are greatly influenced by environmental factors on the phenotypic value of the individuals evaluated (Cruz, Carneiro, & Regazzi, 2014). The crop and its genotypes are grown in different environmental conditions, exposed to different soil characteristics, temperature, photoperiod and precipitation, and when compared, their performances may not be consistent. This relative change is called genotype x environment interaction (Destro, Carpentieri-Pípolo, Kiihl, & Almeida, 2001). Therefore, the behavior of the genotypes in the environments in which they will be cultivated must be analyzed in order to ensure correct positioning.

Among the possible analyzes of this behavior to be carried out, the biplot stands out, which is generated from the main component of the matrix, and where several variables and the study genotypes are considered simultaneously in a precise, robust and integrated analysis. (Maia et al., 2016). The dendrogram, where similarity indices can be read, corresponds to the Euclidean distances at which the observed points join together to form groups (Araújo, Uribe-Opazo, & Johann, 2013). K-means is a grouping method used for image segmentation, based on the similarity of the light intensity of the images (Takahashi, Bedregal, & Lyra, 2005). The Kohonen map is a neural network, whose main characteristic is that it is unsupervised (Paiva, Nagano, & Hongyu, 2019). Through this set of analyses, the objective of this work was to identify superior soybean cultivars through a multivariate approach applied to phenotypic traits.

Material and Methods

This study was developed in the 2023/2024 agricultural harvest, in the experimental area of the Regional Institute for Rural Development, at UNIJUÍ. It is located in the municipality of Augusto Pestana, in the state of Rio Grande do Sul, Brazil (28° 26' 30" S, 54º 00' 58" W, altitude of 400 meters). The soil in the area is classified as a typical dystroferric Oxisol, with a deep profile, well-drained, dark red in color, with high clay content. The climate is type Cfa according to the Köppen climate classification (Alvares, Stape, Sentelhas, Gonçalves, & Sparovek, 2014), characterized as a subtropical climate. Table 1 shows the chemical and physical analysis of the soil in the area, carried out prior to sowing. The cultivars used in the experiment and their qualitative and technological traits are shown in table 2.

Sowing was carried out in the first fortnight of November, with a seeder-fertilizer, using 250 kg ha^{-1} of N-P-K on 05-20-10, employing a density of 14 seeds per linear meter, sown on the eighth of November 2023 and with subsequent harvest on March

25, 2024. The experimental design used was randomized blocks with internal blocks, with the treatment being ten cultivars (Table 2) and five replications. The experimental unit was composed of seven lines 15 m long, spaced 0.5 meters apart. Phytosanitary management was carried out preventively (Table 3), in order to avoid the influence of abiotic effects on the results of the experiment.

Table 1. Chemical and physical analysis of the soil, where the experiment was carried out.

CLAY (%)	pH	Index SMP	P[1] (mg dm^{-3})	K (mg dm^{-3})	OM	Al (cmol dm^{-3})	Ca (cmol dm^{-3})	CTCph 7,0 (cmol dm^{-3})	H+Al	V (%)	M (%)	S	Zn	Cu	Mn
54 (Class 3)	5.2	5.8	39.5	328	3.3	0.1	5.3	14.5	5.5	62.3	1.2	16.1	4.6	12	35.1

[1]P: phosphorus; K: potassium; MO: organic matter; Al: aluminum; V%: base saturation; M%: aluminum saturation; Ca: calcium; S: sulfur; Zn: zinc; Cu: copper; Mn: manganese.

Table 2. Description of the qualitative traits of the cultivars used in the study. Information taken from the RNC (national registry of cultivars), ministry of agriculture, livestock and supply (MAPA).

CULTIVARS	Flower color	Hilo Color	Growth habit	Pubescence color	Vegetable Color	Seed Shape	Integument color	RMG
BMX FÚRIA (65K67RSF E)	White	Light brown	Undetermined	Gray	Dark gray	Spherical	Yellow	6.5
BMX TITANIUM (56IX58RSF I2X)	Purple	Imperfect black	Undetermined	Gray	Dark brown	Spherical	Yellow	5.6
BMX TORQUE (57IX60RSF I2X)	Purple	Imperfect nlack	Undetermined	Gray	Dark gray	Flattened Spherical	Yellow	5.7
BMX TROVÃO (51IX51RSF I2X)	White	Light brown	Undetermined	Light brown	Dark gray	Flattened Spherical	Yellow	5.2
NS 112262	NI[1]	NI	NI	NI	NI	NI	NI	NI
NS 6299 IPRO	Purple	Imperfect black	Undetermined	Gray	Dark gray	Flattened Spherical	Yellow	6.3
NS 5922 IPRO	White	Light brown	Undetermined	Gray	NI	NI	NI	5.9
NS 5252 I2X	White	Brown	Undetermined	Light brown	Dark brown	Flattened Spherical	Yellow	5.2
NEO 610 IPRO	Purple	Imperfect black	Undetermined	Gray	Light gray	Spherical	Yellow	6.1
NEO 590 I2X	Purple	Imperfect black	Undetermined	Gray	NI	NI	NI	5.9

[1]NI: No information.

To estimate the yield of these cultivars, ten plants were randomly collected in each experimental unit at the reproductive stage, when they were between R5.1 and R5.5, and again at the maturation stage. The variables analyzed in the reproductive period were: number of plants per linear meter (NPM, unit); plant height (PH, cm); area of productive zone (APZ, cm); insertion height of the first vegetable (IHFV, cm); number of total nodes on the main stem (NTNMS, unit); number of total nodes in branches (NTNB, unit); number of vegetables on the main stem (NVMS, unit);

number of vegetables on branches (NVB, unit); number of branches (NB, unit); branch length (BL, cm); root length (ROOTL, cm); main stem internode length (INTER_MS, cm); branching internode length (B_INT_LEN), cm); number of vegetables with 1 grain (NV1G, unit); number of vegetables with 2 grains (NV2G, unit); number of vegetables with 3 grains (NV3G, unit); number of vegetables with 4 grains (NV4G, unit); number of vegetables with 0 grains (NVOG, unit); Percentage of disease incidence (PDI, %); percentage incidence of pest insects (PIPI, %); percentage of incidence of invasive plants (PIIP, %);

On the plants collected after maturation, the following variables were analyzed: vegetable grain weight with one grain (VGW1G, grams); vegetable grain weight with two grains (VGW2G, grams); vegetable grain weight with three grains (VGW3G, grams); vegetable grain weight with four grains (VGW4G, grams); grain weight per plant (GWP, grams); grain yield (GY, grams). Satellite data on mean air temperature (Tmain, ºC), minimum air temperature (Tmin, ºC), maximum temperature (Tmax, ºC) and precipitation (Prec, mm) were obtained from the NASA POWER system (NASA Prediction of Worldwide Energy Resources [NASA POWER], 2023).

Table 3. Description of phytosanitary management carried out in the experiment.

Time of application	Active principles	Dose
Desiccation 30 days before sowing	Glyphosate	2.0 L ha^{-1}
	2,4-D	1.5 L ha^{-1}
	Clethodim	0.6 L ha^{-1}
	Mineral oil	0.25 L ha^{-1}
Sequential desiccation, pre-seeding	Diquate	2.5 L ha^{-1}
	Mineral oil	0.25 L ha^{-1}
V4	Bifenthrin	0.200 g ha^{-1}
	Diflubenzuron	0.060 g ha^{-1}
	Difenoconazole	0.25 L ha^{-1}
	Glyphosate	2.50 L ha^{-1}
	Mineral oil	0.25 L ha^{-1}
R1	Thiamethoxam + Lamba-Cialotrin	0.25 L ha^{-1}
	Benzovindiflupir + Ciproconazole + Difenoconazole	0.50 L ha^{-1}
	Mineral oil	0.25 L ha^{-1}
R3	Benzovindiflupir + Prothioconazole	0.50 L ha^{-1}
	Bifenthrin	0.20 kg ha^{-1}
	Mineral oil	0.25 L ha^{-1}
R5	Prothioconazole + Mancozeb	2.00 kg ha^{-1}
	Acetamiprid + Bifenthrin	0.30 kg ha^{-1}
	Mineral oil	0.25 L ha^{-1}

The data obtained was subjected to analysis to remove outliers. They were subsequently subjected to the assumptions of normality of errors and homogeneity of variances, using the Shapiro-Wilk and Barlett tests. With the assumptions met, analysis of variance was carried out in order to determine the effect of cultivars on the agronomic traits evaluated, at 5% probability using the F test. In the variables on which a significant effect was found, a Tukey mean comparison test was carried out at 5% probability. Principal component analysis was used to analyze the contribution of each variable to the explainability of the results obtained, as well as their relationship with the cultivars evaluated.

A dissimilarity dendrogram was obtained for the cultivars from the average Euclidean distance with the UPGMA (Unweighted Pair Group Method Using Arithmetic Averages) grouping method. In a complementary way, Kohonen self-organizing maps were used to group data that were similar to each other, forming classes and K-means clustering, using the basic algorithm to group n observations. The analyzes were carried out using the R software (R CORE TEAM, 2024). The packages used were: AGROR and METAN.

Results and Discussion

During the experiment, the temperature in November (Figure 1) had a mean of 23ºC, a maximum of 28ºC and a minimum of 17ºC, the lower basal temperature of soybeans is 10°C and with reference to the minimum optimum temperature is 20°C (Neumaier et al., 2020). In relation to the month of December, the mean temperature was 24ºC, maximum 29ºC and minimum 18ºC, for January, when the crop was in the vegetative period, the mean temperature was 24ºC, maximum 29ºC and minimum 18ºC, in February the mean temperature was around 25ºC, a maximum of 30ºC and a minimum of 20ºC, finally, in March, the mean temperature was 23ºC, a maximum of 29ºC and a minimum of 18ºC. When analyzing the middle of January until the end of March, temperatures were around 27°C, remaining within the ideal temperature for growing soybeans, ranging between 20 and 30°C (Empresa Brasileira de Pesquisa Agropecuária [EMBRAPA], 2011).

The precipitation (Figure 1) obtained during the crop cycle in December was 80 mm, in November 80 mm, in January it was 50 mm, during flowering, in one of the periods when the crop most needs water (Ferrari, Paz, A. & Silva, 2015). In February, there were fluctuations in precipitation, reaching 80 mm, while in March, mild rain of 15 mm. According to Silva, Sediyama and Borém (2015), the ideal water demand for soybean cultivation varies between 450 and 850 mm. The availability of water for soybeans is essential for physiological processes, photosynthesis, nutrient transport and absorption, thermoregulation, plant growth and development (Campos, Santos, & Nacarath, 2021).

In the descriptive analysis (Table 4), for the variable length of the branching internode, the cultivar that presented the highest average was the cultivar NS 5922 IPRO, with 4.6 cm and the smallest BMX TORQUE (57IX60RSF I2X), with 1.4 cm.

The number of vegetables with 0 grains was higher in the NEO 590 I2X cultivar, with an average of 3.2 legumes. For the number of vegetables with 1 grain, it was observed more frequently in the cultivar NS 112262, with an average of 6 vegetables, and less frequently in the cultivar NS 5252 I2X. For the variable number of vegetables with 2 grains, the cultivar that stood out with the greatest quantity was NEO 610 IPRO with an average of 47.8 vegetables, and the lowest cultivar was BMX TORQUE (57IX60RSF I2X). Regarding the number of vegetables with 4 grains, the cultivar that obtained the highest average was BMX TITANIUM (56IX58RSF I2X) with 9.2 vegetables per plant. For the number of vegetables in the branch, NEO 610 IPRO was superior to the other cultivars, with an average of 77.2 vegetables. The number of legumes per plant, as well as the number of grains formed in these legumes, are important primary components of yield (Thomas & Costa, 2010).

In evaluating the percentage incidence of diseases manifested in working soybeans (Table 5), the cultivars BMX FÚRIA (65K67RSF E) and BMX TORQUE

(57IX60RSF I2X) presented the highest incidence of *Phakopsora pachyrhizi*, 12 and 11% respectively, in addition to the occurrence of phytotoxicity and *Cercospora sojina* in both. Similar behavior was observed for the cultivars BMX TITANIUM (56IX58RSF I2X) and BMX TROVÃO (51IX51RSF I2X) on the incidence of *Phakopsora pachyrhizi*, *Fusarium solani* and *Cercospora sojina*. The cultivars NEO 610 IPRO, NS 112262 and NS 5252 I2X stand out, which did not show an incidence of *Phakopsora pachyrhizi*.

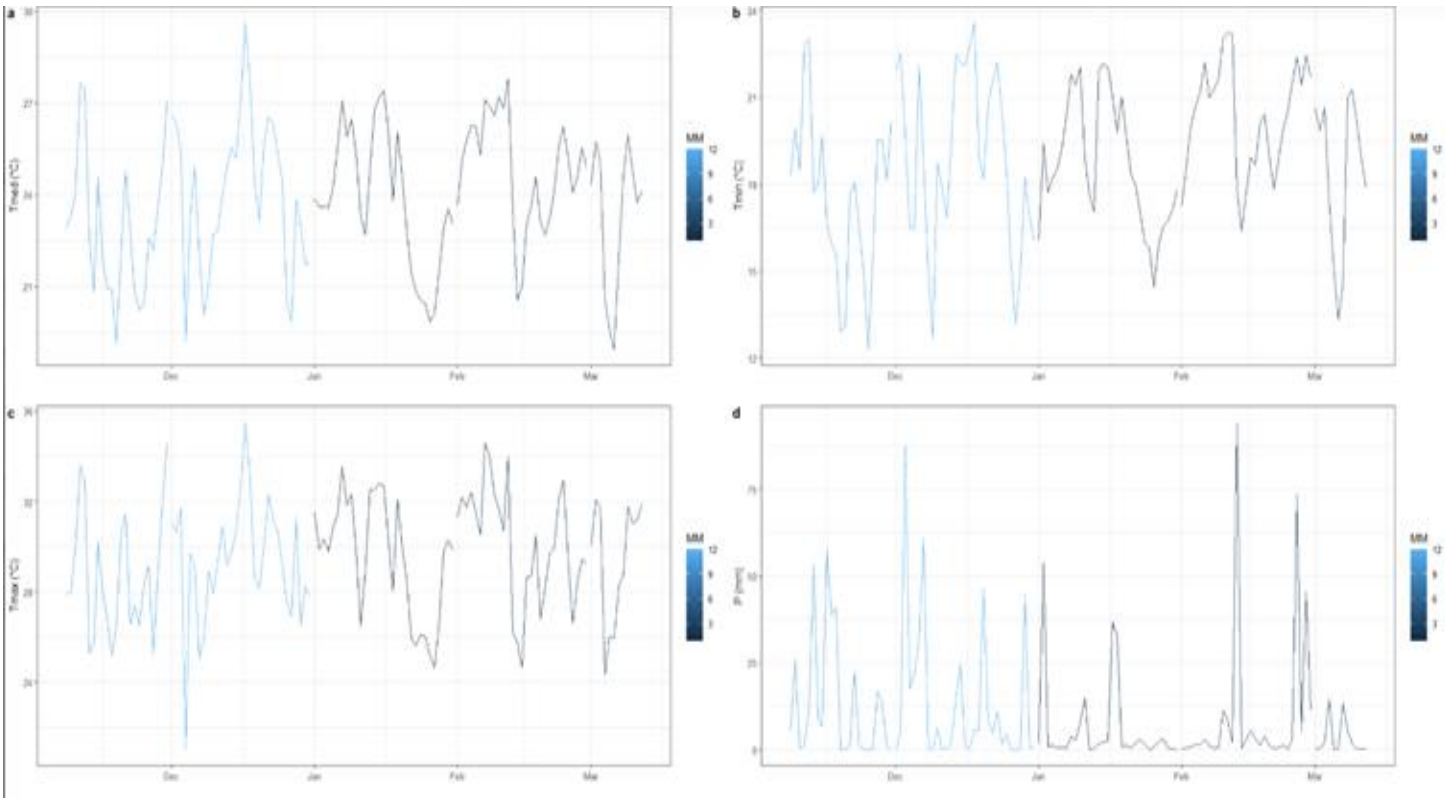

Figure 1. Meteorological data from the location and period during which the work was conducted: (A) - mean air temperature (Tmean, °C); (B) - minimum air temperature (Tmin, °C); (C) - maximum air temperature (Tmax, °C); (D) - precipitation (Prec, mm).

Table 4. Descriptive analysis of the agronomic traits of ten soybean cultivars. (INTER R, cm): Internode length of the branch; (NV0G, unit): Number of vegetables with 0 grains; (NV1G, unit): Number of vegetables with 1 grain; (NV2G, unit): Number of vegetables with 2 grains; (NV4G, unit): Number of vegetables with 4 grains; (NVB, unit): Number of vegetables in the branches.

CULTIVAR	INTER_R (cm)	NV0G (unit)	NV1G (unit)	NV2G (unit)	NV4G (unit)	NVB (unit)
BMX FÚRIA (65K67RSF E)	2.9	0.6	3.8	15.6	8.8	54.2
BMX TITANIUM (56IX58RSF I2X)	2.5	0	2.4	14.4	9.2	68.6
BMX TORQUE (57IX60RSF I2X)	1.4	0	4	12.6	4.4	45.8
BMX TROVÃO (51IX51RSF I2X)	2.7	0.2	4.8	24	1.4	36.8
NEO 590 I2X	2.7	3.2	5.6	26.4	0.4	31.8
NEO 610 IPRO	1.5	2	4.8	47.8	1.2	77.2
NS 112262	2.1	1	6	42	0.2	40.4

NS 5252 I2X	1.6	0.8	1.2	13.2	1.8	43.6
NS 5922 IPRO	4.6	0.8	2.8	15.4	2.6	13.4
NS 6299 IPRO	3.0	0.4	4.4	14.4	1.2	23.8

For diseases that compromise the plant stand during the crop cycle, *Phytophthora soyae* is observed, only when cultivating NEO 610 IPRO, the highest incidences of *Fusarium solani* in NS 5252 I2X (10%), and *Macrophomina phaseolina* in cultivars NEO 610 IPRO, NS 5922 IPRO and NS 5252 I2X, with 8, 10 and 12% respectively, the latter being the cultivar that showed incidence only of these root system diseases, without the incidence of others observed in the area. *Sclerotina sclerotiorum* was only identified in plants of the cultivar NS 112262, with 2%, but this is the cultivar with the lowest incidence of diseases, but with greater sensitivity to the products used, with 12% phytotoxicity. The economic importance of each disease varies from year to year and from region to region, depending on the climatic conditions of each harvest, where some can cause losses of up to 100% (Amorim, Rezende, Bergamin Filho, & Camargo, 2016).

Table 5. Incidence of measured diseases and phytotoxicity that were present in cultivars.

CULTIVAR	*Pseudomonas savastanoi* (%)	*Phakopsora pachyrhizi* (%)	*Cercospora sojina* (%)	*Phytophthora sojae* (%)	*Fusarium solani* (%)	*Macrophomina phaseolina* (%)	*Sclerotina sclerotiorum* (%)	Phytotoxicity (%)
BMX FÚRIA (65K67RSF E)	0	12	6	0	0	0	0	7
BMX TITANIUM (56IX58RSF I2X)	0	9	10	0	4	0	0	0
BMX TORQUE (57IX60RSF I2X)	0	11	7	0	0	0	0	8
BMX TROVÃO (51IX51RSF I2X)	0	4	6	0	6	0	0	2
NEO 590 I2X	0	6	1	0	0	1	0	0
NEO 610 IPRO	0	0	0	4	0	8	0	3
NS 112262	1	0	3	0	0	0	2	12
NS 5252 I2X	0	0	0	0	10	12	0	0
NS 5922 IPRO	5	3	13	0	0	10	0	0
NS 6299 IPRO	0	4	3	0	4	3	0	2

For insect pests (Table 6), there was a greater presence of *Diabrotiva speciosa* in the cultivar NS 112262, and its absence in NS 5252 I2X. For *Euschistus heros*, the cultivar that had the greatest presence was BMX TROVÃO (51IX51RSF I2X), and the cultivars BMX TITANIUM (56IX58RSF I2X), NEO 590 I2X, NEO 610 IPRO, NS 5252 I2X and NS 5922 IPRO demonstrated absence. For *Caliothrips braziliensis*, the cultivar that showed the highest presence of the pest was NS 112262 and the one that had the lowest presence was the cultivar NEO 610 IPRO. The occurrence of insect pests can be an important factor in the loss of productive potential in all types of crops, for which it is of great advantage to adopt more sustainable and efficient management, such as IPM (Integrated Pest Management) itself (Vasconcellos, Corassa, Pitta, & Rolim, 2023).

In the scenario of invasive plants present in the area (Table 6), the plots with the highest presence of *Conyza bonariensis* were the cultivar BMX FÚRIA (65K67RSF E), with 24%, and with the lowest presence in the cultivars NS 5922 IPRO and NS 6299 IPRO, however infesting plots of all cultivars. In an experiment conducted in the field over two consecutive harvests, a 12% decrease in soybean yield was observed, equivalent to around 560 kg per hectare, in the presence of just one *Conyza bonariensis* plant per square meter. For *Amaranthus sp*. the cultivar that had the greatest presence was BMX TORQUE (57IX60RSF I2X), and the cultivars that were absent of this invader were BMX FÚRIA (65K67RSF E), NS 112262 and NS 5252 I2X. In relation to *Ipomoea triloba*, only the cultivars BMX FÚRIA (65K67RSF E) and NS 5922 IPRO were present. Finally, for *Digitaria ciliaris*, only the cultivar NS 6299 IPRO was present. Some invasive plants, regardless of the botanical family, explore the same ecological niches and compete for nutrients with soybean crops (Fortes et al.,, 2017).

Table 6. Insect pests and invasive plants measured in the evaluation of cultivars in the experimental area.

CULTIVAR	***Conyza bonariensis* (%)**	***Amaranthus sp.* (%)**	***Ipomoea triloba* (%)**	***Digitaria ciliaris* (%)**	***Diabrotica speciosa* (%)**	***Euschistus heros* (%)**	***Caliothrips braziliensis* (%)**
BMX FÚRIA (65K67RSF E)	24	0	20	0	3	5	8
BMX TITANIUM (56IX58RSF I2X)	11	7	0	0	17	0	39
BMX TORQUE (57IX60RSF I2X)	15	11	0	0	6	6	30
BMX TROVÃO (51IX51RSF I2X)	6	3	0	0	12	11	17
NEO 590 I2X	7	2	0	0	0	0	7
NEO 610 IPRO	4	5.4	0	0	0	0	5
NS 112262	4	0	0	0	21.6	3	24.6
NS 5252 I2X	7	0	0	0	0	0	12
NS 5922 IPRO	3	3	3	0	6	0	22
NS 6299 IPRO	3	3	0	1	15	5	18

From the summary of the analysis of variance (Table 7), a significant effect was found at 5% probability using the F test, of the cultivars on the variables PHl, IHFV, NTNMS, INTER, NVMS, NB, INTER MS LEN, ROOTL FINAL , NV3G, VGW1G, VGW2G, VGW3G, GWP and GY, and no significant effect for APZ and BL. Comparison of means of variables measured on cultivars is showed in Table 8.

The principal component analysis (PCA) (Figure 2) obtained in the study allowed explaining 43.4% of the total variability of the information (Dim1:23.3%; Dim2: 20.1%). The variable with the greatest contribution was the vegetable grain weight with 2 grains, with a value greater than 7.5%. The intermediate ones, which were found to be on average, were number of vegetables on the main stem, grain weight per plant, number of vegetables on the branch, branch length, number of total nodes on the branches and height of insertion of the first vegetable, with a value close to at 5.0%. It was observed that the variables that contributed least were vegetable grain weight with 1 grain, number of vegetables with 0 grains, number of legumes with 1 grain and number of total nodes on the main stem.

PCA Biplot made it possible to identify relationships between cultivars and the expression of certain traits. In this way, it was verified that the cultivars NEO 610 and NS 112262 presented proximity to the variables, grain weight per plant, number of vegetables in the main stem, grain weight with 3 legumes, number of legumes with 2 grains and length of the internode of the main stem. The cultivar BMX TORQUE (57IX60RSF I2X) showed differential performance for production zone area, root length and number of vegetables with 4 grains. The cultivar NS 5252 I2X demonstrated greater proximity to the grain yield trait. The cultivars BMX TITANIUM (56IX58RSF I2X), BMX FÚRIA (65K67RSF E), BMX TROVÃO (51IX51RSF I2X), NS 5922 IPRO and NS 6299 IPRO did not show specific proximity to any of the evaluated traits. (Figure 2).

Table 7. Summary of the analysis of variance for the effect of cultivars on soybean agronomic traits: PH: Plant height; IHFV: insertion height of the first vegetable; APZ: Production zone area; NTNMS: Number of total nodes on the main stem; NTNB: Number of total nodes in the branches; NVMS: Number of vegetables on the main stem; NB: Number of branches; BL: branch length ROOTL: Root length; HEI INTER MS: Internode length of the main stem; FINAL ROOTL: Final root length; NV3G: Number of vegetables with 3 grains; VGW1G: Vegetable grain weight with 1 grain; VGW2G: Grain weight of 2 vegetables; VGW3G: Vegetable grain weight with 3 grains; GWP: Grain weight per plant; GY: Grain yield.

VF[1]	**DF[2]**	**PH**	**IHFV**	**APZ**	**NTNMS**	**NTNB**	**NVMS**	**NB**	**BL**	**INTER_MS**	**ROOTL**	**NV3G**	**VGW1G**	**VGW2G**	**VGW3G**	**GWP**	**RG**
CULTIVAR	9	0.001*	8.008e-05*	0.373	3.003e-05*	0.0003*	0.0008*	0.0004*	0.054	0.0002*	0.0005*	0.002*	0.00005*	0.00000*	0.00000*	0.00000*	0.00000*
BLOCK	4	0.806	0.786	0.261	0.573	0.366	0.443	0.440	0.921	0.768	0.801	0.538	0.713	0.468	0.140	0.183	0.212
RESIDUAL	36																
CV%[3]		7.26	18.37	28.79	11.08	34.9	21.77	27.58	23.02	16.91	17.19	31.97	30.87	16.57	17.22	15.19	16.46

[1]VF: Variation factor; [2]DF: Degree of freedom; [3]CV: Coefficient of variation; *: significant at 5% probability using the F test.

Table 8. Comparison of means of variables measured on cultivars: PH: Plant height; IHFV: Insertion height of the first vegetable; APZ; Area of production zone; NTNMS: Number of total nodes on the main stem; NTNB: Number of total nodes in the branches; NVMS: Number of vegetables on the main stem; Number of branches; INTER MS: Internode length in the main stem; BL: Branch length; ROOTL: Root length; FINAL ROOTL: Final root length; NV3G: Number of vegetables with 3 grains; VGW1G: Vegetable grain weight with 1 grain; VGW3G: Vegetable grain weight with 3 grains; GWP: Grain weight per plant; GY: Grain yield.

CULTIVAR	PH (cm)	IHFV (cm)	APZ (cm)	NTNMS (unit)	NTNB (unit)	NVMS (unit)	NB (unit)	INTER_MS (cm)	BL (cm)	ROOTL (cm)	NV3G (unit)	VGW1G (g)	VGW2G (g)	VGW3G (g)	GWP (g)	GY (g)
NS 5922 IPRO	116.2 a	22.0 b	55.0	23.0 a	16.8 b	49.4 b	3.0 b	4.150 a	64.0	17.0 b	43.4 b	0.57 c	5.81 c	13.77 ab	21.41 ab	5235.3 a
NEO 590 I2X	113.8 a	24.6 a	87.2	25.4 a	30.2 b	56.8 a	5.4 a	3.436 b	79.6	15.8 b	50.2 b	0.79 bc	6.44 bc	10.29 bcd	17.95 abc	3191.3 bcd
NS 112262	110.2 a	17.0 c	74.0	23.8 a	40.4 a	58.0 a	6.4 a	3.992 a	85.8	24.4 ab	39.2 b	1.27 ab	10.44 a	12.02 abc	23.75 a	4222.5 ab
BMX TITANIUM (56IX58RSF I2X)	100.8 b	21.8 b	82.0	17.6 b	42.2 a	64.4 a	4.4 b	4.808 a	86.4	20.6 ab	83.4 a	1.04 abc	8.24 ab	12.64 abc	22.98 a	2553.9 d
NS 6299 IPRO	98.0 b	25.2 a	76.2	24.2 a	22.2 b	35.8 b	4.4 b	3.632 b	63.4	17.8 b	37.4 b	0.83 bc	5.03 cd	8.74 cde	15.94 bcd	4607.7 a
BMX FÚRIA (65K67RSF E)	97.8 b	28.0 a	82.8	26.0 a	49.2 a	37.2 b	7.0 a	3.193 b	81.8	25.6 a	48.4 b	0.64 c	3.23 d	6.99 de	12.59 cd	2798.0 cd
BMX TORQUE (57IX60RSF I2X)	97.2 b	27.4 a	71.6	20.6 b	55.6 a	42.8 b	6.2 a	3.568 b	78.4	17.8 b	47.6 b	0.77 bc	6.10 bc	14.25 ab	23.03 a	3071.8 bcd
NEO 610 IPRO	90.4 b	21.2 b	70.8	25.0 a	43.4 a	46.0 b	6.2 a	2.836 b	68.0	18.4 b	62.2 a	0.52 c	8.32 ab	14.35 a	24.16 a	4833.7 a
BMX TROVÃO (51IX51RSF I2X)	80.6 c	18.2 c	64.0	21.6 b	26.0 b	47.0 b	3.4 b	2.983 b	64.6	17.4 b	48.8 b	1.44 a	5.16 cd	4.98 e	11.65 d	2331.3 d
NS 5252 I2X	79.6 c	15.4 c	63.8	19.2 b	34.2 b	39.4 b	4.6 b	3.443 b	55.6	15.4 b	70.8 a	0.97 abc	6.53 bc	12.39 abc	20.38 ab	4077.7 abc

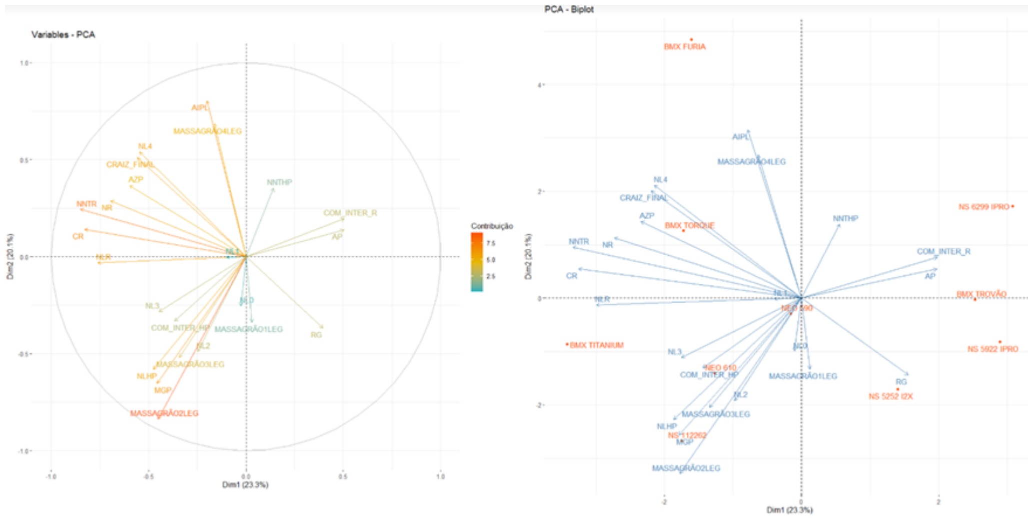

Figure 2. Principal components (Biplot), to explain the variables plant height, insertion height of the first vegetable, area of productive zone, number of total nodes on the main stem, number of total nodes on the branches, number of vegetables on the main stem, number of branches, branch length, main stem internode length, root length, number of vegetables with 3 grains, grain weight of vegetables with 1 grain, grain weight of 2 vegetables, grain weight of vegetable with 3 grains, grain weight per plant and grain yield.

The dendrogram (Figure 3) obtained by UPGMA (Unweighted Pair Group Method Using Arithmetic Averages) from the dissimilarity matrix, showed the formation of 2 large groups, group 1 with the cultivars NS 6299 IPRO, BMX TROVÃO (51IX51RSF I2X), NS 5252 I2X and NS 5922 IPRO, were discrepant when compared with the other group due to the presence of *Euschistus heros*, phytotoxicity, *Fusarium solanie*, *Macrophomina phaseolina*, *Conyza bonariensis*, production zone area, number of total nodes in the branch, number of branches, root length, number of vegetables with 4 grains, number of vegetables with 0 grains, vegetable grain weight with 2 grains. Group 2 was separated into 2 subgroups, the first 2-1 with the cultivars NEO 610 IPRO, NEO 590 I2X and NS 112262, similar due to the presence of *Diabrotica speciosa*, *Caliothrips brasiliensis*, *Euschistus heros*, *Phakopsora pachyrhiz*, *Cercospora sojina* and area of production zone, and subgroup 2-2 with cultivars BMX TITANIUM, BMX FURIA and BMX TORQUE similar to *Caliothrips brasiliensis*, *Cercospora sojina*, production zone area and number of vegetables with 0 grains.

When analyzing the first cluster of the Kohonen map (Figure 4), the variables that were grouped were: vegetable grain weight with 1 grain, number of vegetables with 1 grain, number of vegetables with 2 grains, number of vegetables with 4 grains and number of vegetables with 3 grains. According to Lima and Perluzio (2015), quantifying the genetic dissimilarity between individuals generates information about the degree of similarity, which allows the formation of groups, using grouping methods.

In the second cluster (Figure 4), the variables that were grouped were: number of vegetables with 4 grains, number of vegetables with 2 grains, number of vegetables with 1 grain, grain weight per plant, grain yield, vegetable grain weight with 4 grains, vegetable grain weight with 3 grains, vegetable grain weight with 2 grains, vegetable grain weight with 1 grain, number of vegetables with 3 grains. In relation to the third cluster, the variables that were grouped were: number of vegetables with 3 grains, number of vegetables with 4 grains, number of vegetables with 0 grains, vegetable grain weight with 3 grains, vegetable grain weight with 1 grain, vegetable grain weight with 2 grains, vegetable grain weight with 4 grains, grain weight per plant, grain yield. For Val, Júnior, Bizari, Mauro and Trevisoli (2014), measuring crop agronomic traits is a relevant strategy for selecting genotypes of great agronomic interest.

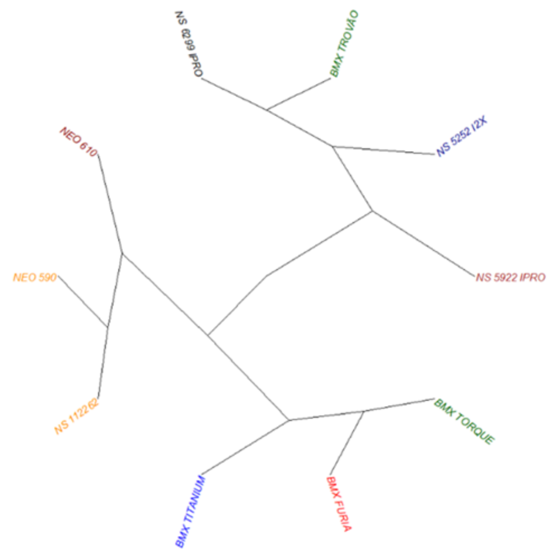

Figure 3. Dendrogram for the formation of groups with variable dissimilarity: plant height; insertion height of the first vegetable; production zone area; number of total nodes on the main stem; number of total nodes in the branches; number of vegetables on the main stem; number of branches; branch length; main stem internode length; root length; number of vegetables with 3 grains; vegetable grain weight with 1 grain; grain weight of 2 vegetables; vegetable grain weight with 3 grains; grain weight per plant; grain yield.

It was observed that in the fourth cluster (Figure 4), the variables that were grouped were vegetable grain weight with 1 grain, vegetable grain weight with 3 grains, vegetable grain weight with 4 grains, grain weight per plant, grain yield, number of vegetables with 1 grain and number of vegetables with 4 grains. Self-organizing maps, SOM (Self-Organizing Maps) or Kohonen Maps are neural network models whose artificial neurons are arranged in a usually two-dimensional grid and are interconnected.

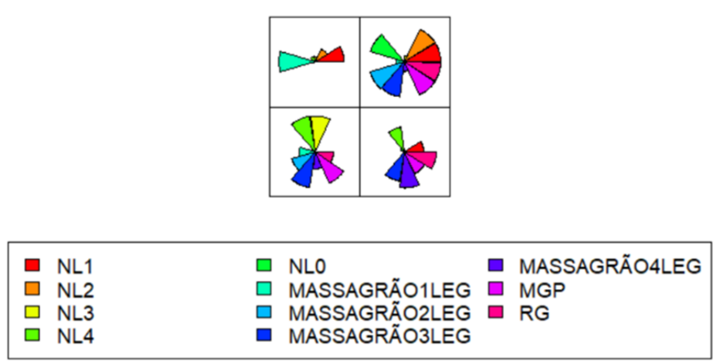

Figure 4. Kohonen map, based on unsupervised learning, to explain which variables contributed to the separation of clusters. Number of vegetables with 0 grains; number of vegetables with 1 grain; number of vegetables 3 grains; number of vegetables with 4 grains; vegetable grain weight with 1 grain; grain weight of 2 vegetables; vegetable grain weight with 3 grains; vegetable grain weight with 4 grains; grain weight per

plant; grain yield.

K-means clustering (Figure 5) consists of an iterative classification process, which finds statistically similar groups during the analysis (Vianna et al., 2013). In this study, K-means clustering allowed explaining 45.3% of the information variability (Dim1:26.1%; Dim2:19.3%), with the formation of two clusters. The first grouped the cultivars BMX TORQUE (57IX60RSF I2X), NEO 610 IPRO, NS 112262, similar due to the presence of *Diabrotica speciosa*, *Caliothrips brasiliensis*, *Euschisthus heros*, *Phakopsora pachyrhizi* and *Cercospora sojina*, APZ, NV0, and the cultivars BMX FURIA (65K67RSF E), NEO 590 I2X and BMX TITANIUM (56IX58RSF I2X), with the occurrence of these factors, but with less intensity, insufficient to enter the group. In the second cluster (Figure 5). The cultivars NS 6299 IPRO, BMX TROVÃO (51IX51RSF I2X), NS 5922 IPRO and 5252 I2X were found, which formed the group due to the great discrepancy between them for the following considerations: *Euschistus heros*, phytotoxicity, *Fusarium solanie*, *Macrophomina phaseolina*, PMLF, APZ, NTNB, NB, BL_I, NV4G, NV0 and VGW2G.

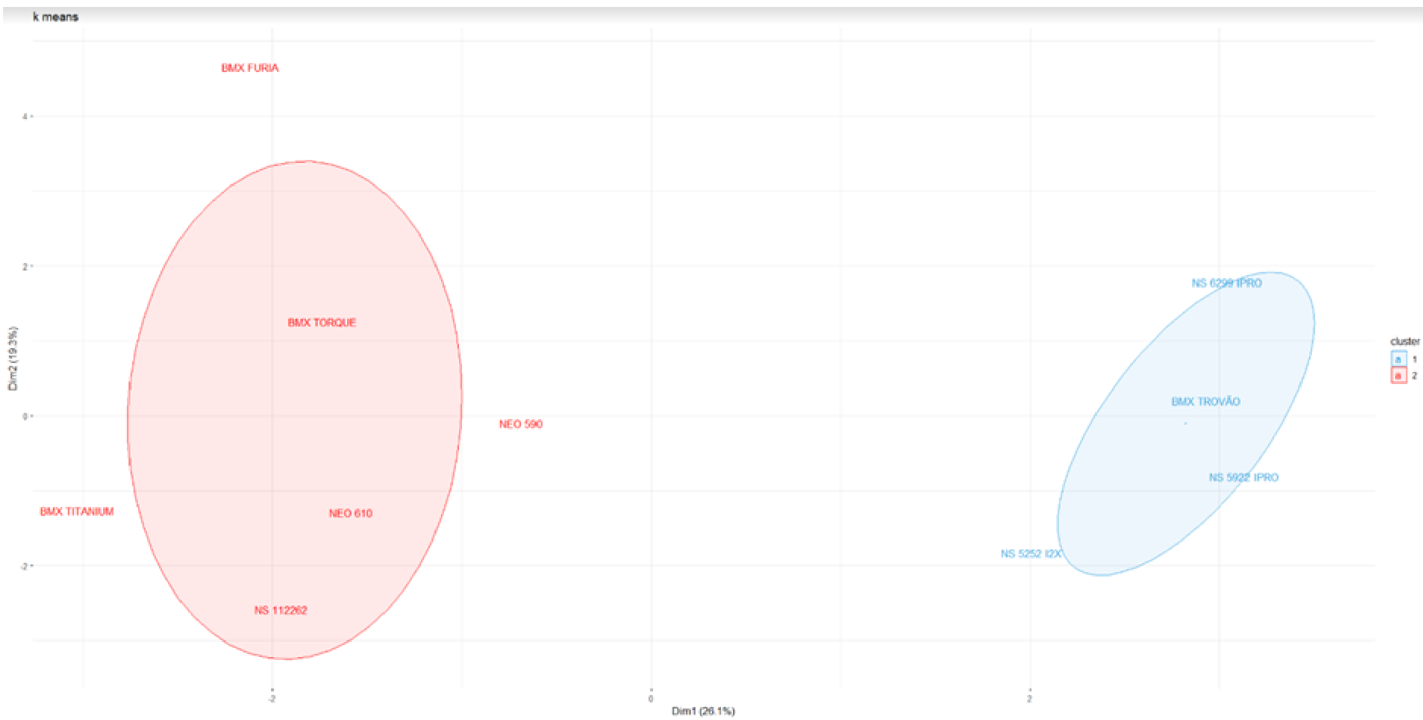

Figure 5. K-means to explain information variability. The first grouped the cultivars BMX TORQUE (57IX60RSF I2X), NEO 610 IPRO, NS 112262, BMX FÚRIA (65K67RSF E), NEO 590 I2X and BMX TITANIUM (56IX58RSF I2X), the second cluster with the cultivars NS 6299 IPRO, BMX TROVÃO (51IX51RSF I2X), NS 5922 IPRO and 5252 I2X.

Conclusions

The trial of ten cultivars in the northwest of the state of Rio Grande do Sul demonstrated that the soybean cultivar with the highest yield was NS5922IPRO, with 5235.3 kg ha^{-1}. The multivariate approaches formed two groups to explain the factors that influenced yield, where the first was discrepant for the variables *Euschistus heros*, phytotoxicity, *Fusarium solanie*, *Macrophomina phaseolina*, *Conyza bonariensis*, production zone area, number of total nodes in the branch, branch number, root length, number of vegetables with 4 grains, number of vegetables with 0 grains and vegetable grain weight of 2 grains. The second similar group for the variables *Diabrotica speciosa*, *Caliothrips brasiliensis*, *Euschisthus heros*, *Phakopsora pachyrhizi* and *Cercospora sojina*, area of production zone, number of vegetables with 0 grains.

References

Alvares, C. A., Stape, J. L., Sentelhas, P. C., Gonçalves, J. L. M., & Sparovek, G. (2014). Köppen's climate classification map for Brazil. *Meteorologische Zeitschrift,* 22(6), 711–728. https://doi.org/10.1127/0941-2948/2013/0507

Amorim, L., Rezende, J. A. M., Bergamin Filho, A., & Camargo, L. E. A. (2016). Phytopathology Manual:*Diseases of Cultivated Plants*, (vol.2, 5th ed.). São Paulo, SP: Editora Agronômica Ceres.

Araújo, E. C., Uribe-Opazo, M. A., & Johann, J. A. (2013). Cluster analysis of the spatial variability of soybean productivity and agrometeorological variables in the western region of Paraná. *Engenharia Agrícola*, 33(4), 782-795. https://doi.org/10.1590/S0100-69162013000400018

Campos, A. J. M., Santos, S. M., & Nacarath, I. R. F. F. (2021). Water stress in plants: a review**.** *Research, Society and Development,*10(15), e311101523155-https://doi.org/10.33448/rsd-v10i15.23155

Cruz, C. D., Carneiro, P. C. S., & Regazzi, A. J. (2014). *Biometric models applied to genetic improvement.* (3rd ed,). Viçosa, MG: Editora da UFV.

Destro, D.; Carpentieri-Pípolo, V.; Kiihl, R. A. S. & Almeida, L. A. (2001). Photoperiodism and genetic control of the long juvenile period in soybean: a review**.** *Crop Breeding and Applied Biotechnology,* 1(1), 72-92.

EMBRAPA - Empresa Brasileira de Pesquisa Agropecuária. (2011). Soybean production technologies - Central region of Brazil 2012 and 2013. Sistemas de Produção, 15. Londrina, PR: Embrapa Soja.

FAO – Food and Agriculture Organization. (2023). *FAOSTAT: Agricultural production*. Rome, Italy: FAOSTAT. Access in: May, 2024. Available in: http://faostat.fao.org/.

Ferrari, E., Paz, A., & Silva, A. C. (2015). Water deficit in soybean metabolism in early sowing in Mato Grosso. *Nativa3*(1), 67-77. https://doi.org/10.14583/2318-7670.v03n01a12

Fortes, C. T., Basso, F. J. M., Galon, L., Agazzi, L. R., Nonemacher, F., & Concenço, G. (2017). Competitive ability of transgenic soybean cultivars coexisting with weeds. *Brazilian Journal of Agricultural Sciences*. 12(2), 185-193.

França-Neto, J. D. B.; Krzyzanowski, F. C.; Henning, A. A.; Pádua, G. P.; Lorini, I.; Henning, F. A. (2016). High quality soybean seed production technology. Brasília, DF: Infoteca/Embrapa.

Gazzoni, D. L., & Dall'agnol, A. (2018). *A saga da soja: de 1050 aC a 2050 dC*. Londrina, PR: Embrapa Soja.

Lima, M. D., & Peluzio, J. M. (2015). Genetic dissimilarity in soybean cultivars focusing on the fatty acid profile aiming to produce biofuel. *Brazilian Journal of Agricultural Sciences,* 10(2), 256-261.

Maia, M. C. C., Araújo, L. B. D., Dias, C. T. D. S., Oliveira, L.C.D., Vasconcelos, L. F. L., Carvalho, E.V. D., Simeão, M., & Bastos, Y. G. M. (2016). Selection of mango osa genotypes in a breeding population using the multivariate-biplot method**.** *Rural Science, 46*(10), 1689-1694. https://doi.org/10.1590/0103-8478cr20130722

Minuzzi, R. B., Frederico, C. A., & Silva, T. G. F. (2017). Estimation of soybean agronomic performance in climatic scenarios for Southern Brazil. *Revista Ceres*, 64(6), 567-573. https://doi.org/10.1590/0034-737X201764060002

NASA POWER – NASA Prediction of Worldwide Energy Resources. (2023). *Prediction of Worldwide Energy Resource Applied Science Program.* Hampton, VA, USA: NASA POWER. Available at: < https://power.larc.nasa.gov/docs/>

Neumaier, N., Farias, J. R. B., Nepomuceno, A. L., Mertz-Henning, L. M., Foloni, J., Moraes, L. A. C., & Gonçalves, S. (2020). *Soybean ecophysiology.* In: SEIXAS, C. D. S., Neumaier, N., Balbinot Junior, A. A., Krzyzanowski, F. C., Leite, R. M. V. B. C. (Eds.). Tecnologias de produção de soja. (pp. 33-54). Londrina, PR: Embrapa Soja.

Paiva, M. S., Nagano, M. S., & Hongyu, K. (2019). Analysis of soybean expansion: An application of multivariate analysis and artificial neural networks. *Sigmae,* 8(2), 554-563.

R Core Team - *A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing.

Silva, F., Sediyama, T., & Borém, A. (2015). *Soybeans from planting to harvesting***.** Viçosa, MG: Editora UFV.

Takahashi, A., Bedregal, B. R. C., & Lyra, A. (2005). An Interval Version of the Image Segmentation Method Using K-means. *Trends in Computational and Applied Mathematics*, 6(2), 315-324.

Thomas, A. L., & Costa, J. A. (2010). Development of the soybean plant and grain yield potential. *Soy: management for high grain productivity*. p. 13-33.

Val, B. H. P., Júnior, J. A. F., Bizari, E. H., Mauro, A. O., & Trevisoli, S. H. U.(2014). Genetic diversity of soybean genotypes through agromorphological characters. *Science & Technology*, v.6, p.72-83.

Vasconcellos, M. C., Corassa, J. D. N., Pitta, R. M., & Rolim, G. G. (2023). Pest control strategies in soybeans and their implications for the arthropod community and crop profitability. *Nativa, 11*(1), 28-43.

Vianna, V. F., Trevisoli, S. H. U., Desidério, J. A., Santiago, S. D., Charnai, K., Júnior, J. A. F., & Mauro, A. D. (2013). The multivariate approach and influence of characters in selecting superior soybean genotypes. *African Journal of Agricultural Research, 8*(30), 4162-4169. https://doi.org/10.5897/AJAR2013.7064