Revista Ciência Agronômica, v. 56, e202493500, 2025
Centro de Ciências Agrárias - Universidade Federal do Ceará, Fortaleza, CE
www.ccarevista.ufc.br

Scientific Article
ISSN 1806-6690

Relationship between the plant, seeds and physiological characteristics in the soybean[1]

Jessica Deolinda Leivas Stecca[2], Maria Inês Diel[2], Alessandro Dal'Col Lúcio[2], Evandro Ademir Deak[2], Matheus Martins Ferreira[3], Glauber Monçon Fipke[4], Lucas Nascimento Brum[2], Thomas Newton Martin[2]

ABSTRACT - Canonical correlations estimate important relationships between sets of variables, making it easier to manage studies of a large number of characteristics. The aim of this study was to identify the canonical correlations between the seeds, physiological and plant characteristics of soybean, as well as the Pearson linear correlations within each set of characteristics. A randomised block design was used in the field and a completely randomised design in the laboratory, with four and three replications respectively. Canonical groups were defined between the set of plant characteristics (weight and number of nodules, shoot dry weight, number of pods, 1000-grain weight, and grain yield) and the set of seed variables (shoot and root length, plant dry weight, accelerated ageing, electrical conductivity, first count and germination), the set of physiological variables (protein, carotenoid and guaiacol peroxidase content of the grains and leaves) and the plant variables, and the physiological variables and the seeds. Electrical conductivity has a negative correlation with first count and germination. Grain yields are higher when the weight of the plants, the 1000-grain weight, and the number of vegetable parts per plant are greater, while carotenoids are correlated with the chlorophyll levels. Higher levels of protein in the grains and leaves are associated with a higher 1000-grain weight and higher grain yield. The canonical correlation analysis shows that the sets of plant, seed and physiological variables are not independent, and defines the characteristics that should be given priority in the evaluations.

Key words: Glycine max (L). Canonical correlations. Set of variables. Plant physiology. Nodulation.

DOI: 10.5935/1806-6690.20250067
Editor-in-Chief: Profa. Charline Zaratin Alves - charline.alves@ufms.br
*Author for correspondence
Received for publication 31/05/2024; approved on 06/02/2025
[1]Extracted from the doctoral thesis of the first author, presented to the Graduate Program in Agronomy at the Federal University of Santa Maria (UFSM). This study was partially funded by the Coordination for the Improvement of Higher Education Personnel - Brazil (CAPES) – Financial Code 001
[2]Department of Crop Sciences, Rural Sciences Centre, Federal University of Santa Maria (UFSM), 97105-900 Santa Maria-RS, Brazil, jessica.stecca@yahoo.com.br (ORCID ID 0000-0002-2939-1389), mariaines.diel@gmail.com (ORCID ID 0000-0002-7905-2166), adlucio@ufsm.br (ORCID ID 0000-0003-0761-4200), evandro.deak@hotmail.com (ORCID ID 0000-0003-0471-0707), lucasnascimentobrum@gmail.com (ORCID ID 0009-0009-2225-825X), martin.ufsm@gmail.com (ORCID ID 0000-0003-4549-3980)
[3]Centro Universitário Faema (UNIFAEMA), Ariquemes-RO, Brazil, math.ferreira10@yahoo.com.br (ORCID ID 0000-0002-4555-7852)
[4]Federal University of Pampa (Unipampa), 97650-000, Itaqui-RS, Brazil, glauberfipke@unipampa.edu.br (ORCID ID 0000-0001-9621-1678)

INTRODUCTION

Brazil is one of the world's leading producers of soybean, and although land capacity for expansion is limited, there is still potential for increasing yield per unit area. With the 2022/23 harvest, the country reached an average productivity of 3,508 kg ha^{-1} (CONAB, 2022). Brazil has introduced sustainable cultivation technologies, such as inoculation with nitrogen-fixing bacteria, which can provide up to 94% of the nitrogen (N) required for soybean growth and development (Hungria *et al*., 2005). The introduction of technological advances into the production chain, including the use of cultivars with high yield potential and the adoption of new management practices, has proved to be important for improving yields (Macholdt; Honermeier, 2017). However, expressing the productive potential of the crop depends on several factors, such as disease management (Gabriel *et al*., 2018; Knebel *et al*., 2019), climate conditions (temperature, sunlight and rainfall) (Leng *et al*., 2016) and fertiliser management (La Menza *et al*., 2017; Maiga *et al*., 2019). It is also important to determine a cultivar, since their genetic differences result in different responses to the treatments and environments in which they develop (Nascimento Junior *et al*., 2023). It is likely that high yield potentials are not economically sustainable (Hatfield *et al*., 2018).

Grain yield depends on a number of factors. First and foremost is the plant stand in the field, which determines seed quality, and which in turn is dependent on the conditions to which the parent plants were exposed (Kameswara Rao; Dullo; Engels, 2016; Roy *et al*., 2024; Wijewardana; Reddy; Bellaloui, 2019). For this reason, management practices implemented throughout the crop cycle are aimed at producing ideal conditions for growth and development in the field that will result in higher quality grain. We need to better understand the effects of technologies employed throughout the plant cycle, in order to identify possible relationships between the agronomic characteristics of the seeds, and the physiological and morphological characteristics of the plant, and adapt management practices according to end-use and the crop cultivars used.

Using multivariate statistical techniques, such as canonical correlation analysis, allows variables of interest to be grouped together in order to identify associations between the groups based on which agronomic characteristics can be selected for indirectly (Silva *et al*., 2019). This will facilitate further research, reducing the number of variables for analysis and providing the opportunity for studying the set of variables according to the availability of the material, ease of measurement and time of evaluation.

Using canonical correlations and associations identified between sets of variables, higher soybean productivity was correlated with a greater number of pods, seeds with a high oil content, and a high germination percentage and rate of emergence (Pereira *et al*., 2017). To increase soybean productivity, it is essential to have an understanding of the physiological relationships between grain productivity and the canopy, dry matter accumulation, photoassimilate partitioning (Vogel *et al*., 2021), net photosynthesis, number of grains (Souza *et al*., 2022), number of pods and number of raceme nodes (Egli, 2013).

However, the relationships between the characteristics of the soybean have not been sufficiently studied in terms of canonical correlations between the seeds and the physiological and morphological characteristics of the plant. As such, the aim of the present study was to identify the canonical correlations between the seeds, and the physiological and morphological characteristics of the soybean.

MATERIAL AND METHODS

The trials were conducted in the experimental area of the Department of Crop Science of the Centre for Rural Sciences at the Federal University of Santa Maria (UFSM), at 29°43'2.81" S, 53°43'58.28" W and an elevation of 116 meters. According to the Köeppen classification, the climate in the region is 86.7% type Cfa and 13.3% type Cfb (Alvares *et al*., 2013).

The soil in the area is classified as an Ultisol (USDA, 2014), with the following chemical soil analysis in the 0-10 cm layer: pH (water, 1:1) = 5.7; organic matter (%, m/v) = 2.0; clay (%, m/v) = 21; phosphorus, P-Mehlich (mg dm^{-3}) = 35.9; potassium (cmol$_c$ dm^{-3}) = 0.153; H Al (cmol$_c$ dm^{-3}) = 3.9; CEC (pH 7, cmol$_c$ dm^{-3}) = 11.8; base saturation (%) = 67.0. Based on this analysis, the area was fertilised with P_2O_5 and K_2O for an expected yield of five tons of soybean per hectare.

The NS 5959 IPRO and TMG 7062 INTACTA cultivars were used (acquired from the seed trade) at a respective sowing density of 37 and 26 seeds m^{-2}. The treatments were applied in a randomised block design with four replications, using a (2 x 3) +1 factorial arrangement, and included the addition or lack of an osmoprotectant and different methods of Co and Mo application, either seed treatment (ST) or foliar application at stage V3, giving a total of six treatments and the control. Each experimental plot was 7.75 m x 2.25 m, with five rows spaced 0.45 m apart, for a total area of 17.4 m^2 and a working area of 6.75 m^2.

The chemical treatment of the seeds included Pyraclostrobin 25g a.i. L^{-1} + Methyl thiophanate 225 g a.i. L^{-1} + Fipronil 250g a.i. L^{-1} (Standak® Top). Mo and Co were applied to the seeds at doses of 20 mL ha^{-1}

and 2 mL ha^{-1}, respectively, with double the rate on the leaves during stage V3. Osmoprotectant (water, active metabolic bacteria, sugar complex and biopolymers) was applied at 1 mL kg^{-1} seeds. Coinoculation was carried out with liquid inoculant using the bacteria *Bradyrhizobium japonicum* at a concentration of 7 x 10^9 CFU mL^{-1} and *Azospirillum brasilense* at 2 x 10^8 CFU mL^{-1}. All the rates were as recommended by the manufacturers.

The following variables were measured in the field at phenological stage R2 (full bloom, open flower on one of the last two nodes of the stem containing one fully developed leaf) (Fehr; Caviness, 1977): number of nodules (NNP, $plant^{-1}$) - direct count of the nodules on the main and secondary roots of each plant; nodule dry weight per plant (NDW, mg $plant^{-1}$) - the nodules were washed, dried in a fan oven at 65 °C for 48 hours and weighed; shoot dry weight (SDW, g $plant^{-1}$) - shoots of the same plants used to determine NDW were dried in a fan oven at 65 °C for 48 hours and weighed; leaf protein (Leafprot, mg ml^{-1}) - determined as per Bradford (1976) using bovine serum albumin as the standard; guaiacol peroxidase enzyme activity (GPA, U mg^{-1} protein) (Zeraik *et al*., 2008); Carotenoids (Carot, mgg^{-1} FW) - extracted from the leaves as per Hiscox and Israelstam (1979) and estimated using the Lichtenthaler equation (1987).

The following variables were measured at full crop maturation (R8): number of pods per plant (NPP, $plant^{-1}$); grain yield (GY, kg ha^{-1}) and 1000-grain weight (TGW, g) (13% humidity). The total chlorophyll content of the plant (Chltotal) was determined using a digital chlorophyll meter (Falker chlorofiLOG CLF2060®) and the electrical conductivity of the seeds (EC) as per Vieira and Krzyzanowski (1999).

After harvesting the seeds, the replications for each treatment were grouped together to form a single sample and the following variables were then analysed in a completely randomised design: grain protein content (Grainprot), near-infrared spectroscopy (NIRS), seed germination (GE) and first count (FC) (Brasil, 2009), root (RL) and shoot (SL) length, plant dry weight (PDW), and accelerated ageing (AA) (Marcos Filho; Novembre; Chamma, 2001).

Data analysis

The data were first analysed using the Pearson linear correlation for each cultivar within each situation: plant variables (NNP, NDW, SDW, NPP, TGW, GY), seed variables (SL, RL, PDW, AA, EC, FC, GE) and physiological variables (Grainprot, Leafprot, Chltotal, GPA, Carot) (Steel; Torrie, 1997). The canonical correlations were analysed to identify the relationships between the sets of variables.

Estimations of the maximum correlation between linear combinations of variables in groups I and II (X_1 and Y_1) are: $X_1 = \alpha_1\chi_1 + \alpha_2\chi_2 + \cdots + \alpha_p\chi_p$ nd $Y_1 = b_1\gamma_1 + b_2\gamma_2 + \cdots + b_q\gamma_q$ where: $\alpha' = [\alpha_1\alpha_2 \cdots \alpha_p]$ = 1xp vector of the weights of the group I variables and $b' = [b_1b_2 \cdots b_q]$ = 1xq vector of the weights of the group II variables.

The first canonical correlation will therefore be the correlation that maximises the relationship between X_1 and Y_1. The X_1 and Y_1 functions form the first canonical pair associated with the canonical correlation expressed by: $r_1 = \frac{C\hat{o}v(X_1, Y_1)}{\sqrt{\hat{V}(X_1) \cdot \hat{V}(Y_1)}}$ where: $C\hat{o}v(X_1, Y_1) = \alpha' S_{12} b, \cdot \hat{V}(X_1) = \alpha' S_{11} \alpha \cdot \hat{V}(Y_1) = b' S_{22} b$, S_{11}= pxp covariance matrix between the group I variables; S_{22} = qxq covariance matrix between the group II variables; S_{12}= pxq covariance matrix between the variables in groups I and II (Cruz *et al*., 2012).

Groups were formed from the seed variables, plant variables and physiological variables. A multicollinearity diagnosis was carried out within each set of variables based on the condition number (CN), which is the ratio between the highest and lowest eigenvalue of the X'X correlation matrix (Montgomery; Peck; Vining, 1982). If the result of this division is CN <100, the multicollinearity is weak; if 100 <CN<1000, the multicollinearity is moderate to strong; and if CN> 1000, the multicollinearity is severe (Cruz *et al*., 2012). In each of the groups, multicollinearity was low.

Groups were then formed from the set of plant variables (NNP, NDW, SDW, NPP, TGW, GY) and the set of seed variables (SL, RL, PDW, AA, EC, FC, GE), the set of physiological variables (Grainprot, Chltotal, Leafprot, GPA, Carot) and the set of plant variables, and the physiological variables and the seed variables. A statistical analysis was run at 5% significance using the pracma (Borchers, 2018), faraway (Faraway, 2016) and yacca (Butts, 2018) packages, in the R software (R Core Team, 2019).

RESULTS AND DISCUSSION

The Pearson correlation analysis for the seed variables showed a positive significance for cultivar NS 5959 IPRO between AA x RL, FC x AA, GE x AA, GE x FC, and a negative significance between EC x AA, FC x EC, GE x EC. For cultivar TMG 7062, the significance was positive for PDW x SL, PDW x RL, FC x RL, FC x PDW, GE x RL, GE x FC, and negative for EC x RL, EC x PDW, FC x EC and GE x EC. For EC, lower values denote higher seed vigour, and a negative significance in relation to the other variables of interest (Table 1).

For the field variables, the correlations were significant and positive between NDW x NNP, GY x SDW, GY x NPP and GY x TGW for both cultivars, and between SDW x NDW for TMG 7062 only. Note that the variables that significantly influenced GY are responsible for the yield component of the crop. In terms of the correlation between the physiological variables, a high Chltotal was seen to correlate with a high Carot in both cultivars. For NS 5959, Grainprot x Leafprot showed a positive correlation, while for Chltotal x GPA the correlation was negative (Tables 2 and 3).

Table 1 - Pearson linear correlation between seed variables (shoot length (SL), root length (RL), plant dry weight (PDW), accelerated ageing (AA), electrical conductivity (EC), first count (FC) and germination (GE)) in the ND 5959 IPRO and TMG 7062 INTACTA cultivars

	ND 5959 IPRO								TMG 7062 INTACTA						
	SL	RL	PDW	AA	EC	FC	GE		SL	RL	PDW	AA	EC	FC	GE
SL	1	0.05	0.17	0.29	0.03	0.12	0.14	SL	1	-0.1	0.55*	0.14	0.04	0.02	-0.33
RL		1	-0.08	0.46*	-0.28	0.29	0.23	RL		1	0.50*	0.12	-0.73*	0.74*	0.63*
PDW			1	0.02	0.21	-0.14	0.03	PDW			1	0.06	-0.56*	0.34*	0.07
AA				1	-0.48*	0.55*	0.48*	AA				1	-0.1	0.11	0.07
EC					1	-0.66*	-0.51*	EC					1	-0.57*	-0.43*
FC						1	0.76*	FC						1	0.64*
GE							1	GE							1

*Significant by t-test (n = 28) at 5% probability

Table 2 - Pearson linear correlation between plant variables (number of nodules (NNP), dry nodule weight (NDW), shoot dry weight (SDW), 1000-grain weight (TGW), number of pods (NPP) and grain yield (GY)) in the ND 5959 IPRO and TMG 7062 INTACTA cultivars

	ND 5959 IPRO							TMG 7062 INTACTA					
	NNP	NDW	SDW	TGW	NPP	GY		NNP	NDW	SDW	TGW	NPP	GY
NNP	1	0.59*	0.16	0.13	0.22	0.26	NNP	1	0.6*	0.11	-0.15	0.23	-0.1
NDW		1	0.15	0.07	0.19	0.16	NDW		1	0.52*	0.12	0.26	0.04
SDW			1	0.16	0.19	0.36*	SDW			1	0.18	0.19	0.36*
TGW				1	0.35	0.75*	TGW				1	0.13	0.53*
NPP					1	0.57*	NPP					1	0.43*
GY						1	GY						1

*Significant by t-test (n = 28) at 5% probability

Table 3 - Pearson linear correlation between physiological variables (grain protein (Grainprot), leaf protein (Leafprot), total chlorophyll (Chltotal), guaiacol peroxidase enzyme activity (GPA) and carotenoid content (Carot)) in the ND 5959 IPRO and TMG 7062 INTACTA cultivars

	ND 5959 IPRO						TMG 7062 INTACTA				
	Leafprot	GPA	Chltotal	Carot.	Grainprot		Leafprot	GPA	Chltotal	Carot.	Grainprot
Leafprot	1	-0.14	0.22	0.01	0.70*	Leafprot	1	0.04	-0.22	-0.09	0.37
GPA		1	-0.45*	-0.21	-0.14	GPA		1	0.33	0.4	-0.13
Chltotal			1	0.48*	0.25	Chltotal			1	0.82*	0.05
Carot.				1	0.11	Carot.				1	-0.13
Grainprot					1	Grainprot					1

*Significant by t-test (n = 21) at 5% probability

The relationship between electrical conductivity, first count and germination is reversed. In seeds already showing degradation of the tegument tissues, vigour and germination are reduced (Coradi *et al*., 2020). Environmental and management conditions are related to seeds production and should be observed to reduce degradation of the seed membranes. The correlation between the number and weight of the nodules is important as it directs research in genetic improvement to cultivars having greater uniformity of nodule size. In the future, it may be possible to establish a relationship between nodule size and/or weight and the efficiency of biological nitrogen fixation (BNF), where this is the main form of nitrogen uptake by the plants. Each ton of grain requires approximately 83 kg of nitrogen, of which 65 are exported (Hungria *et al*., 2005). In terms of productivity components, the most important are the number of vegetable parts and the 1000-grain weight, which directly affect soybean productivity. To minimise biotic and abiotic environmental stress and reduce flower and pod abortion, the rate of biomass accumulation in the soybean should be high with a long period for grain filling (Van Roekel; Purcell; Salmerón, 2015). Managements, environments and cultivars that produce a greater number of pods and grains contribute to higher grain yields. The number of pods increases in proportion to the number of nodes (racemes), the maximum number of nodes on the plant being reached when the pod to node ratio is 70% (Egli, 2013).

Carotenoids are present in the chloroplasts, their levels remaining relatively stable throughout the vegetative phase of the plant, only degrading during senescence. The relationship between carotenoids and chloroplasts is important, since the breakdown of chlorophylls, which are responsible for photosynthesis, begins with the degradation of carotenoids. Management strategies that prevent carotenoid degradation will therefore invariably contribute to a higher rate of photosynthesis and greater dry matter production. Carotenoids also form non-covalent complexes with proteins, accumulating with alpha- and beta-carotene, beta-criptoxanthin, lutein, zeaxanthin, violaxanthin and neoxanthin (Uenojo; Maróstica Junior; Pastore, 2007).

For the canonical correlation analysis, multicollinearity was first tested and found to be low in each of the groups that were formed. In cultivar ND 5959, the condition numbers (CN) were 6.08, 22.13 and 16.91, respectively, for the sets of physiological, plant and seed variables, while for cultivar TMG 7062, the CN were 2.85, 12.21 and 25.83. The canonical correlations were first interpreted using the significance level; when significant, the magnitude of the canonical correlation was examined.

The correlations between the sets of plant and seed variables revealed the first high canonical correlation pairs (ND 5959). However only the first canonical pair was significant (0.96), showing that these groups are dependent and can be used to study the variables they contain. The canonical cross loadings of the first canonical pair revealed the relationship between plants with a higher SDW, TGW and GY. Group I variables are determinant in boosting seed vigour (characterised by the lowest EC and highest SL, AA and FC) and germination (by the highest value for GE) (Table 4).

Table 4 - Canonical correlations between the set of plant variables (number of nodules (NNP), nodule dry weight (NDW), shoot dry weight (SDW), 1000-grain weight (TGW), number of pods (NPP) and grain yield (GY)), and the set of seed variables (shoot length (SL), root length (RL), plant dry weight (PDW), accelerated ageing (AA), electrical conductivity (EC), first count (FC) and germination (GE)) in the ND 5959 IPRO and TMG 7062 INTACTA soybean cultivars

Variable	NS 5959 IPRO						TMG 7062 INTACTA					
	Canonical crossed loadings											
	1º	2º	3º	4º	5º	6º	1º	2º	3º	4º	5º	6º
	Plant variables											
NNP	-0.13	-0.3	0.11	0.14	0.05	0.28	0.12	0.42	0.24	0.29	0.07	0.04
NDW	-0.21	0.52	0.19	0.22	0.18	0.13	0.05	0.32	0.15	0.44	0.03	0.02
SDW	0.56	0.26	0.32	0.29	0.09	0.1	0.47	0.28	0.41	0.18	0.03	0.02
TGW	0.47	0.48	0.28	0.04	0.22	0.03	0.31	0.31	0.2	0.12	0.17	0.06
NPP	0.72	0.39	0.02	0.22	0.11	0.01	0.17	0.28	0.22	0.09	0.27	0.05
GY	0.76	-0.3	0.01	0.11	0.09	0.13	0.67	0	0.09	0.02	0.26	0.01

Continuation Table 4

	Seed variables											
SL	0.55	0.43	0.15	0.29	0.14	0.01	0.34	0.43	0.18	0.04	0.04	0.07
RL	0.4	0.15	0.43	0.08	0.23	-0.1	0.25	0.15	0.07	0.34	0.11	0.02
PDW	0.07	0.23	0.38	0.21	0.02	0.16	0.77	0.12	-0.1	0.17	0.08	0.02
AA	0.66	0.31	0.06	0.11	0.25	0.01	0.19	0.5	-0.3	0.23	0.04	0.02
EC	-0.71	0.17	0.07	0.24	0.01	0.15	0.46	0.02	0.11	0.16	0.28	0.04
FC	0.71	0.16	0.09	0.16	0.06	0.08	0.25	0.02	0.18	0.28	0.06	0.06
GE	0.7	0.26	0.27	0.12	0.15	0.05	0	0.32	0.12	0.39	0.04	0.05
Canonical correlation	0.96	0.82	0.76	0.6	0.42	0.32	0.87	0.69	0.66	0.54	0.43	0.11
Degrees of freedom	42	30	20	12	6	2	42	30	20	12	6	2
p-value	0.01	0.26	0.42	0.63	0.68	0.5	0.58	0.83	0.8	0.85	0.84	0.92

In the canonical correlation analysis between the sets of physiological and plant variables and the physiological and seed variables for both cultivars, and between the sets of seed and plant variables for TMG 7062, the canonical pairs were not significant (Tables 4, 5 and 6).

The different responses of the cultivars are related to genetic factors that reflect the characteristics of the plants (size, architecture, branches and leaves), and the ability to translocate, absorb and accumulate nutrients, produce photoassimilates, and resist pests, i.e. induce physiological, morphological and phenological changes in the plant (Koester *et al.*, 2016; Todeschini *et al.*, 2019).

For the ND 5959 IPRO cultivar, the set of plant and seed variables showed a canonical correlation of high magnitude, indicating that SDW, TGW and GY in group I were determinant in boosting seed vigour, shown by the lower EC and higher SL, AA and FC, while the rate of germination was shown by the higher GE. This result highlights the importance of crop management throughout the growth cycle in obtaining healthy seeds of higher physical and physiological quality (Ludwig *et al*., 2021). In the Pearson linear correlation, the positive influence of SDW, TGW and NPP on GY could be seen in the set of plant variables, showing the importance of these variables in obtaining higher yields, i.e. management practices that enhance these variables should be prioritised. This correlation was also seen in TMG 7062 (Table 1). Todeschini *et al*. (2019) found that NPP has a greater influence on boosting GY. However, Wang *et al*. (2016) attributed increases in GY to TGW. These results corroborate those of the present study, where both variables were found to have a positive correlation with yield. Claims that high grain yields require a high number of nodes are partially correct; above a certain critical level, the number of nodes is not as important as the photosynthesis of the canopy and the amount of photoassimilates destined for the pods (Egli, 2013).

It was also found that higher values for SDW possibly boost photosynthetic activity in the plants, resulting in a greater accumulation of photoassimilates for subsequent grain filling (Ainsworth *et al.*, 2012; Kaschuk *et al.*, 2010). In addition, the Pearson correlation of the physiological variables (ND 5959 IPRO) showed that leaves with the highest concentration of protein correlated with grains with the highest protein content (Table 3). This shows that assimilates are transported from the leaves to the grains, the main type of photosynthate to be translocated being sucrose, a source of carbon and energy for the formation of proteins and oils (Smith; Rinne; Seif, 1989). Protein synthesis also depends on the availability of nitrogen for the seeds, which affects the protein and oil concentration (Assefa *et al.*, 2019). This characteristic of the grain is important with regard to animal feed. Furthermore, soybean are a low-cost source of protein, with the increasing use for human consumption boosting the food industry.

The canonical correlation for the amount of protein in the grain was not significant (ND 5959 IPRO). However, the high magnitude of the correlation of the first canonical pair leads to the conclusion that seeds with higher Grainprot show greater vigour (Table 6). According to Hao *et al*. (2020), this is related to the capacity to mobilise seed reserves during germination. This is because the seeds contain storage proteins, which are necessary as sources of metabolic energy that help in the formation of plant tissue in the new seedling (Goyoaga *et al*., 2011; Kim *et al*., 2011). This corroborates the findings of Carvalho and Nakagawa (2012), who found that seeds of higher physiological quality had a higher crude protein content.

The relationship between vigour and Grainprot was also seen when evaluating EC. A greater release of leachates in solution in the seeds resulted in less-structured membranes and less membrane selectivity

(Vieira *et al*., 2002) due to the low availability of proteins and enzymes responsible for reorganising and forming the cell membrane structures. More-vigorous seeds are able to produce seedlings that are more uniform and competitive in terms of space and natural resources (Müller *et al*., 2017).

Protein synthesis is mainly related to the ability of a plant to absorb nutrients such as nitrogen, which the soybean requires in greater quantity (Medic; Atkinson; Hurburgh Junior, 2014), and which is necessary for a number of growth and development processes. According to Terashima and Hikosaka (1995), nitrogen is essential for synthesising the main CO_2-fixing enzyme Rubisco, and for producing chlorophylls. It is also necessary for nitrogen fixation, which can increase the nutrient concentration in the leaves, thereby stimulating photosynthesis. This takes us back to the variable NPP (number of pods per plant), since nitrogen reduces flower and pod abscission, which is governed by a set of physiological and biochemical (Zhang *et al*., 2018) and hormonal (Kućko *et al*., 2019) factors. Management practices should ensure that the crop has a sufficient supply of nitrogen. The transfer of photoassimilates from the source (leaves) to the drain (grains) is controlled by a highly regulated and resource-dependent signalling network (Rossi; Bermudez; Carrari, 2015). However, source-drain activity is also affected by the environment, which stimulates photosynthetic activity in the source thereby increasing drain activity (tissue growth and storage) (Körner, 2015).

Increasing protein synthesis in the soybean depends on the management practices carried out during the experiment, including biological nitrogen fixation, which according to Zimmer *et al*. (2016), if implemented by inoculating with bacteria of genus *Bradyrhizobium*, results in a higher grain protein content compared to the application of nitrogen fertiliser. In addition, according to Kaschuk *et al*. (2010), inoculated plants have higher rates of photosynthesis and delayed leaf senescence, allowing more time for the production of photoassimilates. However, micronutrients such as Cobalt and Molybdenum are required to boost BNF efficiency, since they are involved in transforming atmospheric nitrogen into an assimilable form (Mus *et al*., 2016).

To achieve the above results, management was necessary throughout the crop cycle, with the aim of obtaining plants with a higher concentration of leaf proteins. This also results in a higher grain yield and protein content; if the aim is to produce seeds, these will be more vigorous and have a higher rate of germination.

Table 5 - Canonical correlations between the set of physiological variables (grain protein (Grainprot), leaf protein (Leafprot), guaiacol peroxidase enzyme activity (GPA) and carotenoid content (Carot)) and the set of plant variables (number of nodules (NNP), nodule dry weight (NDW), shoot dry weight (SDW), 1000-grain weight (TGW), number of pods (NPP) and grain yield (GY)) in the ND 5959 IPRO and TMG 7062 INTACTA soybean cultivars

Variable	NS 5959 IPRO				TMG 7062 INTACTA			
	Canonical crossed loading							
	1º	2º	3º	4º	1º	2º	3º	4º
	Physiological variables							
Grainprot	0.68	0.37	-0.07	0	-0.26	0.47	0.05	-0.03
Leafprot	0.69	0.05	0.25	-0.08	0.57	0.29	0.01	-0.04
GPA	0.19	-0.55	-0.24	0.15	0.22	0.17	-0.17	0.07
Carot	-0.04	0.29	0.22	0.31	0.21	0.07	0.14	0.08
	Plant variables							
NNP	-0.07	0.16	-0.3	-0.09	0.03	-0.15	-0.18	0.05
NDW	-0.15	0.45	-0.32	0.07	-0.28	-0.43	-0.11	0
SDW	0.11	0.14	-0.15	-0.29	-0.59	-0.26	0.06	-0.02
TGW	0.48	-0.24	-0.27	0.05	-0.23	0.19	-0.07	-0.05
NPP	0.67	0.14	-0.22	0.01	-0.26	0.05	0.13	0.05
GY	0.64	0.07	-0.13	-0.13	-0.63	0.27	-0.01	0.01
Canonical correlation	0.8	0.71	0.53	0.39	0.79	0.54	0.29	0.09
Degrees of freedom	24	15	8	3	24	15	8	3
p-value	0.13	0.3	0.51	0.5	0.66	0.97	0.99	0.98

Table 6 - Canonical correlations between the set of physiological variables (grain protein (Grainprot), leaf protein (Leafprot), guaiacol peroxidase enzyme activity (GPA) and carotenoid content (Carot)) and the set of seed variables (shoot length (SL), root length (RL), dry weight of plant (PDW), accelerated ageing (AA), electrical conductivity (EC), first count (FC), germination (GE)) in the ND 5959 IPRO and TMG 7062 soybean cultivars

Variable	ND 5959 IPRO				TMG 7062 INTACTA			
	Canonical crossed loading							
	1º	2º	3º	4º	1º	2º	3º	4º
	Physiological variables							
Grainprot	-0.7	-0.22	-0.11	0.12	0.46	0.93	0.33	-0.04
Leafprot	-0.76	-0.2	0.24	0.03	0.56	-0.55	-0.61	-0.43
GPA	-0.21	0.72	-0.09	-0.04	0.09	-0.24	0.92	-0.56
Carot	0.26	0.08	0.21	0.2	0.61	-0.17	-0.09	0.89
	Seed variables							
SL	-0.37	-0.07	-0.39	-0.04	0.18	-0.55	1.47	0.01
RL	0	-0.15	-0.01	0.07	-0.38	-0.54	0.49	0.96
PDW	0.21	-0.29	0.05	-0.12	0.03	-0.65	-1.39	-0.2
AA	-0.62	-0.24	-0.05	0.09	-0.6	0.33	0.2	-0.13
EC	0.61	-0.27	-0.18	-0.02	0.16	-0.45	-0.58	1.25
FC	-0.54	-0.23	0.03	-0.05	0.03	0.18	-0.98	0.72
GE	-0.48	-0.45	0.15	-0.01	-0.35	-0.22	0.85	-0.61
Canonical correlation	0.89	0.77	0.53	0.23	0.92	0.72	0.41	0.32
Degrees of freedom	28	18	10	4	28	18	10	4
p-value	0.08	0.48	0.87	0.94	0.07	0.7	0.94	0.82

The production of quality seeds is essential, since more-vigorous seeds promote the uniform and rapid establishment of seedlings, affording better ground cover, a deeper root system, better use of the water, light and nutrients, and early onset of the photosynthetic process and biological N fixation (Caverzan *et al*., 2018; Müller *et al*., 2017). Better ground cover, due to a faster growth rate, also affords early shading of the surface, with less water evaporating from the soil so that it can be used for plant growth. Another advantage of ground cover is weed control, depriving weeds of the light they need to germinate, and preventing competition with the crop. In addition, less-vigorous seeds produce more-uneven crops and smaller stands, resulting in dominant and subservient plants, with a consequent reduction in GY.

With the growth in population, the need to increase the amount of protein means it is essential to understand the relationships between the variables involved in soybean cultivation. Research into the types of management and the relationships between variables, as well as their degree of interdependence, can optimise evaluations and direct efforts towards higher grain yields and more protein-rich grain.

CONCLUSIONS

1. The seeds, and the physiological and morphological variables are not independent, but rather interrelated;
2. Certain variables should be prioritised due to their cause and effect relationship. Higher levels of protein in the grains and leaves are associated with plants that produce a higher 1000-grain weight and higher grain yield.

REFERENCES

AINSWORTH E, A. *et al*. Accelerating yield potential in soybean: potential targets for biotechnological improvement. **Plant, Cell & Environment**, v. 35, p. 38-52, 2012.

ALVARES, C. A. *et al.* Köppen's climate classification map for Brazil. **Meteorologische Zeitschrift**, v. 22, n. 6, p. 711-728, 2013.

ASSEFA, Y. *et al.* Assessing variation in US soybean seeds composition (protein and oil). **Frontiers in Plant Science**, v. 10, e298, 2019.

BORCHERS, H. W. **Pracma**: practical numerical math functions. R package version 2.2.2. 2019. Available on:

https://cran.rproject.org/web/packages/pracma/pracma.pdf. Accessed on: Mar. 10, 2020.

BRADFORD, M. M. A rapid and sensitive method for the estimation of microgram quantities of protein utilizing the principle of protein-dye binding. **Analytical Biochemistry**, v. 72, p. 248-254, 1976.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. **Regras para análise de sementes.** Brasília, DF: Mapa/ACS, 2009.

BUTTS, C. T. **Yacca**: yet another canonical correlation analysis package. 2018. Available on: https://CRAN.R-project.org/package=yacca. Accessed on: Mar. 13, 2020.

CARVALHO, N. M.; NAKAGAWA, J. **Sementes: ciência, tecnologia e produção**. Jaboticabal, SP: Editora FUNEP, 2012.

CAVERZAN, A. *et al*. How does seeds vigor affect soybean yield components? **Agronomy Journal**, v. 110, p. 1318-1327, 2018.

COMPANHIA NACIONAL DE ABASTECIMENTO (BRASIL). **Acompanhamento da safra brasileira de grãos**: 2021/2022. 2022. Available on: https://www.conab.gov.br/info-agro/safras/graos/boletim-dasafra-de-graos. Accessed on: Sep. 12, 2023.

CORADI, P. C. *et al.* Soybean seeds storage: Packaging technologies and conditions of storage environments. **Journal of Stored Products Research**, v. 89, e101709, 2020.

CRUZ, C. D. *et al*. **Modelos biométricos aplicados ao melhoramento genético**. Viçosa, MG: Editora UFV, 2012.

EGLI, D. B. The relationship between the number of nodes and pods in soybean communities. **Crop Science**, v. 53, p. 1668-1676, 2013.

FARAWAY, J. Functions and datasets for books by Julian Faraway. Package" faraway. **R package version**, v. 1, n. 7, 2016. Available on: http://CRAN. R-project. org/package= faraway. Accessed on: Mar. 12, 2020.

FEHR, W. R.; CAVINESS, C. E. **Stages of soybean development**. 1977. (*Special Report*, 87). Available on: https://lib.dr.iastate.edu/specialreports/87. Accessed on: Mar. 12, 2020.

GABRIEL, M. *et al.* Chemical control of asian soybean rust and its effect in the yield and quality of soybean seeds. **Journal of Agricultural Science**, v. 10, n. 10, 2018. p. 518-526.

GOYOAGA, C. *et al.* Content and distribution of protein, sugars and inositol phosphates during the germination and seedling growth of two cultivars of Vicia faba. **Journal of Food Composition and Analysis**, v. 24, p. 391-397, 2011.

HAO, Q. *et al.* Evaluation of seeds vigor in soybean germplasms from different eco-regions. **Oil Crop Science**, v. 5, p. 22-25, 2020.

HATFIELD, J. L. *et al.* Vulnerability of grain crops and croplands in the Midwest to climatic variability and adaptation strategies. **Climatic Change**, v. 146, p. 263-275, 2018.

HISCOX, J. D.; ISRAELSTAM, G. F. A method for the extraction of chlorophyll from leaf tissue without maceration. **Canadian Journal of Botany**, v. 57, p. 1332-1334, 1979.

HUNGRIA, M. *et al.* The Importance of Nitrogen Fixation to Soybean Cropping in South America. *In*: WERNER, D.; NEWTON, W. E. (ed.). **Nitrogen fixation in agriculture, forestry, ecology, and the environment**. Dordrecht: Springer, 2005. p. 25-42.

KAMESWARA RAO, N.; DULLO, M. E.; ENGELS, L. M. M. A review of factors that influence the production of quality seeds for long-term conservation in genebanks. **Genetic resources and crop evolution**, v. 64, p. 1061-1074, 2016.

KASCHUK, G. *et al*. Differences in photosynthetic behaviour and leaf senescence of soybean (Glycine max [L.] Merrill) dependent on N2 fixation or nitrate supply. **Plant Biology**, v. 12, p. 60-69, 2010.

KIM, H. T. *et al.* Mobilization of storage proteins in soybean seeds (Glycine max L.) during germination and seedling growth. Biochimica et Biophysica Acta (BBA). **Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics**, v. 1814, p. 1178-1187, 2011.

KNEBEL, D. F. *et al.* Economic efficiency and soybean yield due to the use of different fungicide combinations. **Plant Omics**, v. 12, p. 9-14, 2019.

KOESTER, R. P. *et al.* Has photosynthetic capacity increased with 80 years of soybean breeding? An examination of historical soybean cultivars. **Plant, Cell & Environment**, v. 39, p. 1058-1067, 2016.

KÖRNER, C. Paradigm shift in plant growth control. **Current Opinion in Plant Biology**, v. 25, p. 107-114, 2015.

KUĆKO, A.; WILMOWICZ, E.; OSTROWSKI, M. Spatio-temporal IAA gradient is determined by interactions with ET and governs flower abscission. **Journal of Plant Physiology**, v. 236, p. 51-60, 2019.

LA MENZA, N. C. *et al.* Is soybean yield limited by nitrogen supply? **Field Crops Research**, v. 213, p. 204-212, 2017.

LENG, G. *et al.* The role of climate covariability on crop yields in the conterminous United States. **Scientific Reports**, v. 6, e33160, 2016.

LICHTENTHALER, H. K. Chlorophylls and carotenoids: pigments of photosynthetic biomembranes. **Methods in Enzymology**, v. 148, p. 350-382, 1987.

LUDWIG, V. *et al.* Impact of controlled atmosphere storage on physiological quality of soybean seeds. **Journal of Stored Products Research**, v. 90, e101749, 2021.

MACHOLDT, J.; HONERMEIER, B. Impact of highly varying seeding densities on grain yield and yield stability of winter rye cultivars under the influence of delayed sowing under sandy soil conditions. **Archives of Agronomy and Soil Science**, v. 63, p. 1977-1992, 2017.

MAIGA, A. *et al.* Responses of soil organic carbon, aggregate stability, carbon and nitrogen fractions to 15 and 24 years of no-till diversified crop rotations. **Soil Research**, v. 57, p. 149-157, 2019.

MARCOS FILHO, J.; NOVEMBRE, A. D. C.; CHAMMA, H. M. C. P. Testes de envelhecimento acelerado e de deterioração controlada para avaliação do vigor de sementes de soja. **Scientia Agricola**, v. 58, p. 421-426, 2001.

MEDIC, J.; ATKINSON, C.; HURBURGH JUNIOR, C. R. Current knowledge in soybean composition. **Journal of the American Oil Chemists' Society**, v. 91, p. 363-384, 2014.

MONTGOMERY, D. C.; PECK, E. A.; VINING, G. G. **Introduction to linear regression analysis**. New York: J. Wiley, 1982. 504 p.

MÜLLER, M. *et al.* Grain yield differences of soybean cultivars due to solar radiation interception. **American Journal of Plant Sciences**, v. 8, p. 2795-2810, 2017.

MUS, F. *et al*. Symbiotic nitrogen fixation and the challenges to its extension to nonlegumes. **Applied and Environmental Microbiology**, v. 82, p. 3698-3710, 2016.

NASCIMENTO JUNIOR, V. C. *et al*. Fixação biológica de nitrogênio em soja com diferentes níveis de tolerância à seca. **Revista Brasileira de Ciências Agrárias**, v. 18, p. 3268-3268, 2023.

PEREIRA, E. M. *et al.* Canonical correlations between agronomic traits and seeds physiological quality in segregating soybean populations. **Genetics and Molecular Research**, v. 16, p. 1-11, 2017.

R CORE TEAM. **R**: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2019.

ROSSI, M.; BERMUDEZ, L.; CARRARI, F. Crop yield: challenges from a metabolic perspective. **Current Opinion in Plant Biology**, v. 25, p. 79-89, 2015.

ROY, D. *et al.* Climate change on seeds physiology. *In*: PATHAK, H. *et al*. (ed). **Climate change impacts on soil-plant-atmosphere continuum**. Singapore: Springer Nature Singapore, 2024. p. 347-382. (Advances in Global Change Research, 78).

SILVA, F. L. *et al.* Predictor model and canonical interrelationships based on morphological, bromatological and grain yield characteristics of dual purpose wheat. **Genetics and Molecular Research**, v. 18, n. 2, 2019.

SMITH, A. J.; RINNE, R. W.; SEIF, R. D. Phosphoenolpyruvate carboxylase and pyruvate kinase involvement in protein and oil biosynthesis during soybean seeds development. **Crop Science**, v. 29, p. 349-353, 1989.

SOUZA, A. P. *et al.* Soybean photosynthesis and crop yield are improved by accelerating recovery from photoprotection. **Science**, v. 377, p. 851-854, 2022.

STEEL, R. G.; TORRIE, J. H. **Principles and procedures of statistics**: a iometrical approach. 3rd ed. New York: McGraw Hill, 1997. 358 p.

TERASHIMA, I.; HIKOSAKA, K. Comparative ecophysiology of leaf and canopy photosynthesis. **Plant, Cell & Environment**, v. 18, n. 10, p. 1111-1128, 1995.

TODESCHINI, M. H. *et al*. Soybean genetic progress in South Brazil: physiological, phenological and agronomic traits. **Euphytica**, v. 215, p. 1-12 2019.

UENOJO, M.; MARÓSTICA JUNIOR, M. R.; PASTORE, G. M. Carotenóides: propriedades, aplicações e biotransformação para formação de compostos de aroma. **Química Nova**, v. 30, p. 616-622, 2007.

UNITED STATES. DEPARTMENT OF AGRICULTURE. **Keys to soil taxononmy**. 12 th. ed. Washington: USDA, 2014. 372 p.

VAN ROEKEL, R. J.; PURCELL, L. C.; SALMERÓN, M. Physiological and management factors contributing to soybean potential yield. **Field Crops Research**, v. 182, p. 86-97, 2015.

VIEIRA, R. D. *et al.* Condutividade elétrica e teor de água inicial das sementes de soja. **Pesquisa Agropecuária Brasileira**, v. 37, p. 1333-1338, 2002.

VIEIRA, R. D.; KRZYZANOWSKI, F. C. Teste de condutividade elétrica. *In*: KRZYZANOWSKI, F. C.; VIEIRA, R. D.; FRANÇA NETO, J. B. (ed.). **Vigor de sementes**: conceitos e testes. Londrina: Abrates, 1999. v. 4, p. 1-26.

VOGEL, J. T. *et al.* Soybean yield formation physiology–a foundation for precision breeding based improvement. **Frontiers in Plant Science**, v. 12, e719706, 2021.

WANG, C. *et al.* Seventy-five years of improvement of yield and agronomic traits of soybean cultivars released in the Yellow-Huai-Hai river valley. **Crop Science**, v. 56, p. 2354-2364, 2016.

WIJEWARDANA, C.; REDDY, K. R.; BELLALOUI, N. Soybean seeds physiology, quality, and chemical composition under soil moisture stress. **Food Chemistry**, v. 278, p. 92-100, 2019.

ZERAIK, A. E. *et al.* Desenvolvimento de um spot test para o monitoramento da atividade da peroxidase em um procedimento de purificação. **Química Nova**, v. 31, p. 731-734, 2008.

ZHANG, Q. *et al.* Pod anatomy, morphology and dehiscing forces in pod dehiscence of soybean (Glycine max (L.) Merrill). **Flora**, v. 248, p. 48-53, 2018.

ZIMMER, S. *et al.* Effects of soybean variety and *Bradyrhizobium* strains on yield, protein content and biological nitrogen fixation under cool growing conditions in Germany. **European Journal of Agronomy**, v. 72, p. 38-46, 2016.