Artificial Neural Network and Multivariate Models Applied to Morphological Traits and Seeds of Common Beans Genotypes

I. R. Carvalho[1], V. J. Szareski[1], G. H. Demari[1], M. H. Barbosa[1], G. G. Conte[1], L. F. S. de Lima[1], T. da S. Martins[1], A. S. Uliana[1], M. T. Padilha[1], V. Q. de Souza[2], T. Z. Aumonde[1], F. A. Villela[1] & T. Pedó[1]

[1] Universidade Federal de Pelotas, Campus Capão do Leão, RS, Brazil

[2] Universidade Federal do Pampa, RS, Brazil

Corresponding author: I. R. Carvalho, Universidade Federal de Pelotas, Campus Capão do Leão, RS, Brazil. E-mail: carvalho.irc@gmail.com

Abstract

The aimed to characterize common beans genotypes utilizing multivariate models and artificial neural network thru the agronomic attributes and seeds dimensions. The experiment was conducted in the 2017/2018 crop season at the city of Tenente Portela - RS. The experimental design was expanded blocs, were 53 segregating F2 populations and ten cultivars considered checks, disposed in four repetitions. The accurate characterization of bean genotypes can be based in the reproductive period, cycle and seeds length. Genotypes with longer cycle increase the potential of ramifications, legume and seeds magnitude per plant and increase the seeds yield independent of the commercial group. The use of biometric approach allows revealing patterns to the genotype grouping, being the patterns magnitude dependent of the intrinsic premises to the Standardized Average Euclidian Distance, Tocher optimized grouping and Artificial Neural Network with non-supervised learning. It is defined that the Artificial Neural Network are determinant to define associative patterns, being these inferences indispensable to the common beans genotype selection that answer the agronomic attributes and seeds production.

Keywords: *Phaseolus vulgaris* L., computer learning, selection of genotypes, agronomic ideotype

1. Introduction

The common bean (*Phaseolus vulgaris* L.) is characterized as a leguminous species belonging to the Fabaceae family. Its importance is directed to human feeding and stands out as a major crop in the Brazilian food base, being cultivated in most regions of Brazil (Moura et al., 2013; Pedó et al., 2016). It is submitted to several technological levels from subsistence agriculture to proprieties with long extended area proprieties (Carvalho et al., 2016). Its cultivation results in social, economic and agricultural peculiarities, since it is a low cost source of protein, presents short cycle, and with adequate management, it can be used in seasonal periods in the same summer time of the crop season. It results in a sustainable system, in the other hand, there are necessities on the nutritional and hydric management as much for the seed production as for the grains (Farinelli & Lemos, 2010; Demari et al., 2015).

Many are the factors that can influence the characterization and selection of the most agronomic adequate genotypes, in order to increment the seeds productivity per plant and per area unit. Among these factors, the soil and climatic characteristics of the growing environment, crop season and sowing periods, management and technologies used, morphological and physiological plants characteristics, growth habit, commercial group and genotypes intrinsic genetic constitution (Torres et al., 2015; Troyjack et al., 2017). In this context, due to the varied growing conditions in which the common beans are submitted and to the large genetic variability available to this species, it is necessary to characterize the germplasm to the genetic similarity and dissimilarity with the commercial cultivars, based on the principal agronomic attributes of crops and seeds. In this way, it may be possible to select superior genotypes that attend the crop ideotype desired by the farmer.

With the object to reduce financial costs and expedite the selections to develop new common beans genotypes, an efficient alternative is to employ different biometric approaches that allow estimating and classifying genetic patterns utilizing phenotypic measurements (Carbonell et al., 2007; Szareski et al., 2017). In this manner, seeks

out to conciliate methods that reveal the similarity of the germplasm origin on the geographic distinctions, as well, its tendency of association conjugating the genetic dissimilarity and the development of Artificial Neural Network thru the computing learning. In this context, this work aimed to characterize common beans genotypes utilizing multivariate models and artificial neural network thru the agronomic attributes and seeds dimensions.

2. Material and Methods

The experiment was conducted in the 2017/2018 crop season at the city of Tenente Portela-RS, located at latitude 27°23′31.04″ S and longitude 53°46′50.71″ W, with an altitude of 420 meters. The climate is humid subtropical type *Cfa* according to Köppen classification, and the soil is characterized as typicalumino-ferric Red Latossol (STRECK et al., 2008). The used common beans genotypes were classified according to its germplasm geographic origin (Campos Borges-RS, Palmeira das Missões-RS, Santa Rosa-RS, Pejuçara-RS, Braga-RS, Cruz Alta-RS, Santo Antônio do Goiás-GO, São Paulo-SP, Uberlândia-MG and Londrina-PR), the commercial group characteristics (Roxo, Rajado, Preto and Carioca) and thru the genetic base (segregating populations and cultivars), the detailed information are exposed on Table 1.

Table 1. Description of the germplasm geographic origin, commercial group characterization and genetic base of the used genotypes

Acronyms	Origin	Coordinates	Commercial groups	Genetic Base
CBR1; CBR2	Campos Borges-RS	28°52′31″S/53°00′55″W	Purple	Population
CBJ3; CBJ4; CBJ5; CBJ6; CBJ7; CBJ8; CBJ9			Rajado	
PMJ10; PMJ11	Palmeira das Missões-RS	27°53′19″S/53°18′19″W	Rajado	Population
PMP12; PMP13; PMP15; PMP16			Black	
PMC14			Carioca	
SRJ17; SRJ18; SRJ19; SRJ20; SRJ21; SRJ22	Santa Rosa-RS	27°52′16″S/54°28′55″W	Rajado	Population
SRJ23; SRJ24; SRJ25; SRJ26; SRJ27; SRJ28				
SRJ29; SRJ30; SRJ31; SRJ32; SRJ33; SRJ34				
SRP35; SRP36; SRP37			Black	
PJP38; PJP39; PJP40; PJP41; PJP42; PJP43	Pejuçara-RS	28°25′24′S/53°39′21″W	Black	Population
PJP44; PJP45				
BRP46; BRP47; BRP48	Braga-RS	27°37′16″S/53°44′17″W	Black	Population
CAP49; CAP50; CAP51; CAP52; CAP53	Cruz Alta-RS	28°38′22″S/53°36′22″W	Black	Population
SAP54	Santo Antônio-GO	16°29′01″S/49°18′40″W	Black	BRS Esplendor
CLC55			Carioca	BRS Estilo
SPC56	São Paulo-SP	23°32′56″S/46°38′20″W	Carioca	IAC Formoso
SPC57			Carioca	IAC Imperador
SA C58	Santo Antônio-GO	16°29′01″S/49°18′40″W	Carioca	BRS Perola
UBJ59	Uberlândia-MG	18°54′41″S/48°15′44″W	Rajado	BRSMG Realce
SA P60	Santo Antônio-GO	16°29′01″S/49°18′40″W	Black	BRS Supremo
LOC61	Londrina-PR	23°17′34″S/51°10′24″W	Carioca	IPR Tangara
LOP62			Black	IPR Tiziu
LOP63			Black	IPR Uirauru

The experimental design was expanded blocs (FEDERER, 1956), were 53 segregating F2 populations were used, being: CBR_1, CBR_2, CBJ_3, CBJ_4, CBJ_5, CBJ_6, CBJ_7, CBJ_8, CBJ_9, PMJ_{10}, PMJ_{11}, PMP_{12}, PMP_{13}, PMC_{14}, PMP_{15}, PMP_{16}, SRJ_{17}, SRJ_{18}, SRJ_{19}, SRJ_{20}, SRJ_{21}, SRJ_{22}, SRJ_{23}, SRJ_{24}, SRJ_{25}, SRJ_{26}, SRJ_{27}, SRJ_{28}, SRJ_{29}, SRJ_{30}, SRJ_{31}, SRJ_{32}, SRJ_{33}, SRJ_{34}, SRP_{35}, SRP_{36}, SRP_{37}, PJP_{38}, PJP_{39}, PJP_{40}, PJP_{41}, PJP_{42}, PJP_{43}, PJP_{44}, PJP_{45}, BRP_{46}, BRP_{47}, BRP_{48}, CAP_{49}, CAP_{50}, CAP_{51}, CAP_{52}, CAP_{53} and the cultivars BRS Esplendor, BRS Estilo, IAC Formoso, IAC Imperador, BRS Pérola, BRSMG Realce, BRS Supremo, IPR Tangará, IPR Tiziu and IPR Uirapuru being considered checks and disposed in four repetitions.

The experimental units were composed of five sowing lines with length of five meters each, spaced by 0. 45 meters. The used sowing density was 22 seeds per square meter (m²). The sowing was realized in the second fortnight of November, based on the no-till sowing system. For the fertilizing, it was used 250 kg ha⁻¹ of N, P_2O_5, K_2O in the formulation of (10-20-20) and broadcasted 90 kg ha⁻¹ of nitrogen from urea source (46% of N), in the V4 growth stage. The preventive practices were prioritized, in order to minimize the weed, pests and diseases effects that

could have influenced the experiment results. The agronomic characters were measured with 10 plants randomly harvested in the area (4.05 m²) of the experimental unit, being:

- *Duration of vegetative period (PVE)*: measured thru the counting of numberof days from the sowing to the complete flowering, results in units.

- *Duration of reproductive period (PRE)*: obtained thru the number of days between complete flowering and harvest, results in units.

- *Cylce duration (CIC)*: valued thru the period between sowing and harvest, results in units.

- *Crop height at flowering (ALF)*: measured the extension from the soil level to the plant top, results expressed in centimeters (cm).

- *Crop height at harvesting (ALM)*: obtained at the harvest time by the measurement of the extension from the soil level to the plant top, results in cm.

- *First legume insertion height (IPL)*: measured the extension between the soil level and the first crop viable legume, results expressed in cm.

- *Number of legumes per plant (NPL)*: valued the magnitude of viable legumes per plant, resulting in units.

- *Number of seeds per legume (NSL)*: obtained thru the reason between mass of seeds per plant and total legumes number, results in grams (g).

- *Number of ramifications per plant (NR)*: obtained thru the counting of ramifications with more than ten centimeters of extension and legumes presence, results in units.

- *Seeds mass per plant (MSP)*: after trashing the plants, the seeds were cleaned and submitted to the measurement of seed mass with a precision scale, results in grams.

- *Seeds length (COS) and width (LAS)*: it were measured 100 seeds per experimental units, using a digital caliper ruler, results expressed in millimeters (mm).

The geographic coordinates were used to define the exact location of the germplasm origin environments, these physic distances were utilized in the analysis of main coordinates with the goal of graphically express the environments dispersion thru the geographic distances matrix. Afterwards, the phonotypical data were submitted to the normality and homogeneity of variances, aiming to identify the discrepant observations. Then, it was submitted to analysis of standardized Average Euclidian distance and the matrix construction of the genotypes genetic dissimilarity, this matrix was applied to the grouping method *Unweighted Pair Group Method with Arithmetic Mean* (UPGMA) where the dendogram was constructed. Later, was realized the Tocher optimized grouping method prioritizing the homogeneity inside the group and the heterogeneity between groups. In addition, it was realized the relative character contribution by the Singh method (1981) with the aim of understanding which characters were determinant to the distinction of the tested genotypes.

With the objective of understanding the linear associations between characters and create a casual diagram, it was realized the linear correlations, where the coefficients significance level was obtained by the *t* test with 5% probability. Aiming to identify the patterns for the selection and determine which measured characters associated to each determined pattern, it was used the Artificial Neural Network (*RNAs*) basing its estimates on the non-supervised computational learning, the associated centroids and neurons topologic definition were obtained using the method of *Kohonen* Mapping, using 200 thousand interactions. The statistical analysis were realized using the softwares Genes (Cruz, 2009) and R program (R Core Team, 2015).

3. Results and Discussion

Due to the variability of the germplasm origin used in genotype differentiation studies, it is crucial to employ biometric models that allow to evidence this variation in a comprehensive way. In this manner, the use of main coordinates it is based on the independency between the abscissas axis, that support the schematic representation of the non-linear associations, and its estimates are based in the similarity of the used observations. In this context, it was used the geographic localization of the studied genotypes origin (F2 populations and cultivars) to the construction of the distances matrix. After, the auto values that defined the relative scores (X and Y), to each origin, were estimated.

The main coordinates graphic (Figure 1) represents the variation of nine distinct origins of the common bean genotypes seeds. Where it is found high similarities among the origins of Cruz Alta (CA), Palmeira das Missões-RS (PM) and Pejuçara-RS (PJ), even by being located in the same state as Campos Borges-RS (CB) and Santa Rosa-RS (SR), they diverge from this group, because they are not contrasting environments and are

geographically distant. When it refers to the more discrepant origins, this situation is attributed to check cultivars that are originated from breeding programs located in the South, Southeast and Middle West. It is noted that when situated at the same quadrant São Paulo-SP (SP) and Uberlândia-MG (UB) presented different about its location, having a similar behavior as the ones obtained from Londrina-PR (LO) and Santo Antônio de Goiás-GO (SA).
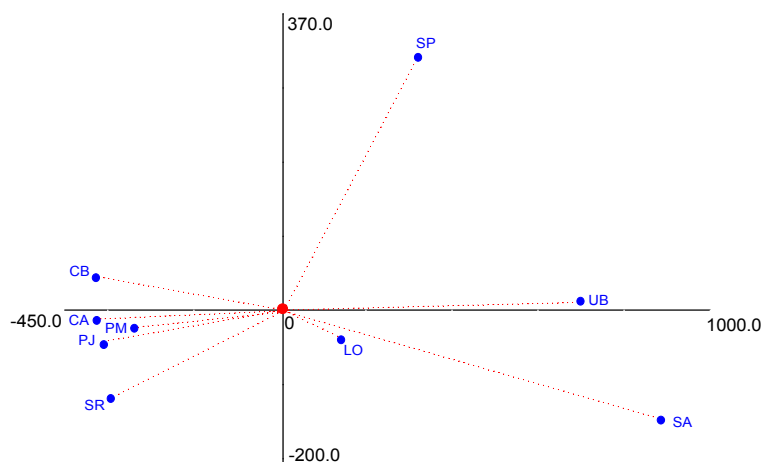

Figure 1. Multivariate method of main coordinates obtained thru the geographic distances of the genotypes origin, being: Campos Borges-RS (CB), Cruz Alta-RS (CA), Palmeira das Missões-RS (PM), Pejuçara-RS (PJ), Santa Rosa-RS (SR), São Paulo-SP (SP), Uberlândia-MG (UB), Londrina-PR (LO) and Santo Antônio do Goiás-GO (SA)

Even though comprehending the differentiations of the germplasm origin, the commercial group and intrinsic genetic base of the studied genotype, many of them presents phenotypically similar, resulting in the necessity to differentiate them efficiently. As follows, some statistic alternatives were developed to support this differentiation, where, first it is established which characters are essential to the differentiation. The characters relative contribution is an efficient method in order to establish which measured character presents relevance (Figure 2). In the circumstance, higher polymorphism between the genotypes was expressed in the seeds length (14.5%) with 8.5 to 14.1 mm amplitude, reproductive period duration (13.3%) comprehending from 38 to 74 days and cycle duration (13.3%) between 68 to 116 days. It was established that these were the more contrasting characters among the genotypes, and so, must be considered to differentiate the genotypes classified as similar thru its origin, commercial group or genetic base.

After learned which characters that had more relevance were, it is indispensable to define which of these presents linearly associated. That is because it is clear that the main components of common beans yield are defined thru the joint association of several characters and these combinations, many times, presents inverted directions. In this manner, it was realized the linear correlations analysis with the goal of establish the tendency of the measured characters and develop a casual diagram using the associations with 5% probability significance by the t test. It were correlated 13 characters measured in the 63 common bean genotypes, where 78 linear associations were established with only 10 significant correlation coefficients (Figure 3) and defined as efficient to develop the casual diagram.
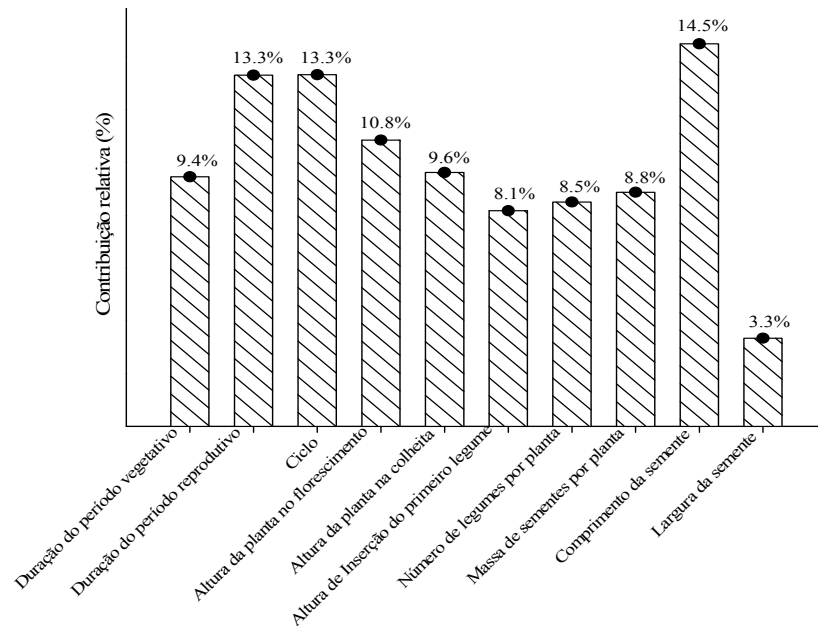
Figure 2. Relative contribution of Singh (1981) to the measured character in 63 common bens genotypes
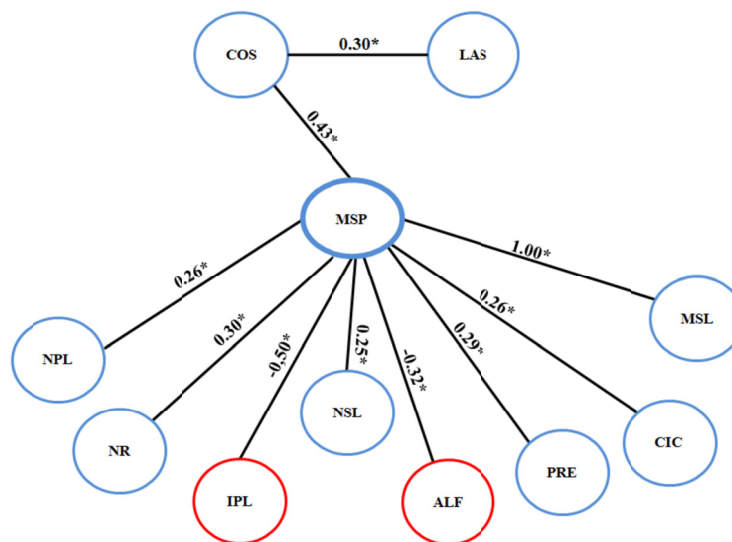

Figure 3. Casual diagram elaborated thru the linear associations between the measured characters, seeds width (LAS), seeds length (COS), seeds mass per plant (MSP), seeds mass per legume (MSL), cycle duration (ClC), duration of reproductive period (PRE), plant height at flowering (ALF), number of seeds per legume (NSL), first legume insertion height (IPL), number of ramifications per plant (NR), number of legumes per plant (NLP)

The casual diagram established that common bean plants with lower insertion height of the first legume (IPL) and crop height at flowering (ALF), as well, longer duration of the reproductive period (PRE), cycle (CIC), number of ramifications per plant (NR), number of legumes per plant (NPL), number of seeds per legume (NSL) and legume seeds mass (MSL) have a tendency to increment the mass of seeds per plant (MSP). In this way, it is defined that the ideal genotype for high seeds productivity per area must be based in the indirect identification of plants that increase the magnitude of reproductive units in the main stem and ramifications. In addition, the identification of late season genotypes or with superiority in the duration of grains filling period can intensify the production. The genotype superiority in relation to the seeds mass per plant (MSP) is justified not only by the seeds production magnitude but, according to the casual diagram, the seeds length (COS) and width (LAS) increment directly the MSP, once this characters are contrasting to the commercial group to which the common bean genotype belongs.

The phenotypic matrix of the characters was submitted to analysis of standardized Average Euclidian distance aiming to obtain the genetic distance matrix among the studied genotypes, this estimates were submitted to a grouping method that supported the confection of the dissimilarity dendogram. The genetic distances showed that the 63 studied genotypes could be classified in 14 subdivisions (Figure 4). Six of these are classified as isolated situations thru the populations (SRJ29, SRJ27, PJP42 and SRP36) and the cultivars BRS Supremo and BRS Esplendor, showing dissimilarity to the others genotypes. To neither possible groups there was a discrimination of the genotypes to the commercial group, because the phenotypic characteristics used did not allow the maximum stratification of the groups in order to exist a group exclusivity for a commercial group. In this manner, the resting eight subdivisions grouped genotypes from both commercial groups.

Figure 4. Dendogram based in the standardized Average Euclidian distance and average grouping method UPGMA, using the genotypes CBR1, CBR2, CBJ3, CBJ4, CBJ5, CBJ6, CBJ7, CBJ8, CBJ9, PMJ10, PMJ11, PMP12, PMP13, PMC14, PMP15, PMP16, SRJ17, SRJ18, SRJ19, SRJ20, SRJ21, SRJ22, SRJ23, SRJ24, SRJ25, SRJ26, SRJ27, SRJ28, SRJ29, SRJ30, SRJ31, SRJ32, SRJ33, SRJ34, SRP35, SRP36, SRP37, PJP38, PJP39, PJP40, PJP41, PJP42, PJP43, PJP44, PJP45, BRP46, BRP47, BRP48, CAP49, CAP50, CAP51, CAP52, CAP53, BRS Esplendor, BRS Estilo, IAC Formoso, IAC Imperador, BRS Pérola, BRS MG Realce, BRS Supremo, IPR Tangará, IPR Tiziu and IPR Uirapuru

Similarities were expressed among the genotypes PJP43, SRJ20, PJP44 and SRJ25 with the cultivar BRS MG Realce from the rajado commercial group. The cultivars IPR Uirapurú from the black commercial group and IPR Tangará from the carioca commercial group are phenotypically associated with the genotypes PJP45, PMP16, CAP49 and SRJ21. The cultivar IAC Imperador is characterized as carioca commercial group and the IPR Tiziu is from black commercial group, both presented similarities to the populations CBJ7, PMP12 and SRJ26. The cultivar IAC Formoso from the carioca commercial group grouped with the genotypes CBR2, CBJ5, SRJ30, CBJ4 and CAP50, and the cultivars of the carioca commercial group BRS Estilo and BRS Pérola are phenotypically similar to the populations CBJ9, PMJ11, SRJ34, SRJ35, CBJ8, CBR1, CAP52, PMP13, SRJ32, CBJ3, CAP51, BRP46, CBJ6 and CAP53.

It is understood that the populations formed groups with some of the used commercial cultivars, and reveal that indirectly the phenotype is near to the agronomically desired to seeds production. In this manner, there are those populations that do not follow this tendency and formed groups with the exclusive presence of populations. Being the populations PJP38, SRJ22, SRJ28, SRJ30, PJP41, BRP47, BRP48, SRJ31, SRJ33, PMP15, SRJ19, SRP37,SRJ17, SRJ24, SRJ23,PJP40, PMJ10 and SRJ18 the ones that needed higher number of measured characteristics to define patterns with the commercial cultivars using this biometric approach. The Tocher optimized grouping method was realized in a multivariate way thru the utilization of the genetic distance matrix

from the measured characters, its interferences are based in the groups definition where internally it is prioritized the homogeneity and between groups the heterogeneity (Santo et al., 2014).

Using this method allowed the fragmentation of 63 genotypes in 10 groups, being that in the first (I) 73% of the genotypes were concentrated, where 11% are commercial cultivars. It is established that in this groups are united those genotypes with agronomic characteristics most close to the prioritized ones (Table 2), in this conditions gather PMP16, PJP45, CAP49, IPR Uirapuru, IPR Tangará, CBR2, SRJ30, CBJ5, CBJ4, CBJ6, IAC Formoso, PMC14, CBJ3, CAP51, CAP53, BRP46, CAP50, CBJ8, SRP35, CBR1, BRS Perola, PMP13, SRJ32, CBJ9, CAP52, SRJ31, SRJ34, BRS Estilo, SRJ23, SRJ33, BRS Esplendor, SRJ18, PJP40, SRJ19, PMJ11, CBJ7, PMP12, IAC Imperador, SRJ26, PMP15, SRP37, PMJ10, IPR Tiziu, SRJ24, BRP47 and SRJ28 defining this group with genotypes with less genetic distinctions. The group II gathered 7.9% of the studied genotypes being PJP39, PJP41, SRJ22, PJP38 and BRS MG Realce, characterizing it as belonging to the rajado commercial group. The remaining groups (III, IV, V, VI, VII, VIII, IX, and X) represents 13.1% of genotypes and its distinctions are based in small amplitudes in the average obtained results, and if there is interest in selecting specific characteristics to the next generations this groups can be prioritized.

Table 2. Optimized grouping method of Tocher to define groups thru the genotypes dissimilarities

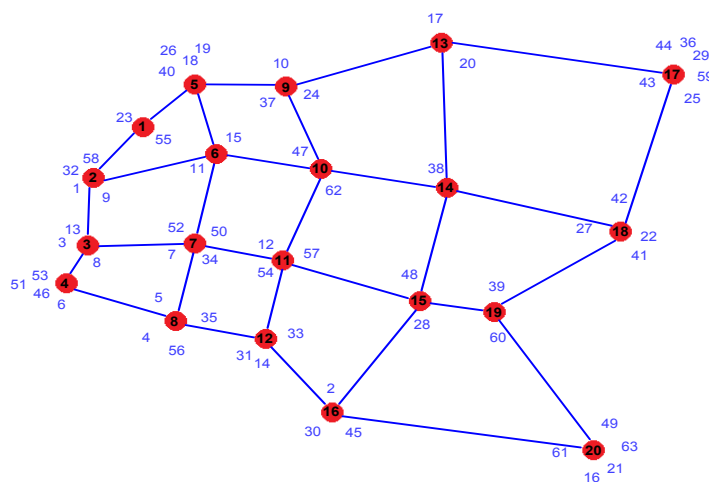| Groups	Optimized grouping of Tocher
I	PMP16, PJP45, CAP49, IPR Uirapuru, IPR Tangará, CBR2, SRJ30, CBJ5, CBJ4, CBJ6, IAC Formoso, PMC14, CBJ3, CAP51, CAP53, BRP46, CAP50, CBJ8, SRP35, CBR1, BRS Perola, PMP13, SRJ32, CBJ9, CAP52, SRJ31, SRJ34, BRS Estilo, SRJ23, SRJ33, BRS Esplendor, SRJ18, PJP40, SRJ19, PMJ11, CBJ7, PMP12, IAC Imperador, SRJ26, PMP15, SRP37, PMJ10, IPR Tiziu, SRJ24, BRP47, SRJ28,
II	PJP39, PJP41, SRJ22, PJP38, BRS MG Realce
III	SRJ20, PJP43, SRJ17, PJP44
IV	SRJ27, BRP48
V	BRS Supremo
VI	SRJ21
VII	PJP42
VIII	SRP36
IX	SRJ29
X	SRJ25

Figure 5. Artificial Neural Networks (RNAs) obtained by the Kohonen Map defining the centroids (red) and neurons (blue) of interest, and the synaptic connections (blue lines), being: CBR1, CBR2, CBJ3, CBJ4, CBJ5, CBJ6, CBJ7, CBJ8, CBJ9, PMJ10, PMJ11, PMP12, PMP13, PMC14, PMP15, PMP16, SRJ17, SRJ18, SRJ19, SRJ20, SRJ21, SRJ22, SRJ23, SRJ24, SRJ25, SRJ26, SRJ27, SRJ28, SRJ29, SRJ30, SRJ31, SRJ32, SRJ33, SRJ34, SRP35, SRP36, SRP37, PJP38, PJP39, PJP40, PJP41, PJP42, PJP43, PJP44, PJP45, BRP46, BRP47, BRP48, CAP49, CAP50, CAP51, CAP52, CAP53, BRS Esplendor (54), BRS Estilo (55), IAC Formoso (56), IAC Imperador (57), BRS Pérola (58), BRS MG Realce (59), BRS Supremo (60), IPR Tangará (61), IPR Tiziu(62) and IPR Uirapuru (63)

The Artificial Neural Network (RNAs) presents indispensable to define mathematic patterns thru the non-linear stochastic phenomenon. In this manner, the definition of the network topology is defined by n non-supervisedinteractive computational process, where the magnitude of the network entries correspond to the neurons number associated to the model explicability. These neurons when related allow the definition of a centroid that correspond to the medium point among the existing associations (Nascimento et al., 2013; Teodoro et al., 2015). For this study, it was prioritized the utilization of the *Kohonen* Mapping method using 63 entries (neurons) with more than 200 thousand interactions (Figure 5). The phenotypic matrix was submitted to interactive proceedings that defined a neural network with topology of 20 centroids that establish associative patterns among the tested genotypes (Figure 5), being necessary 31 synaptic connections to interconnect the centroids.

Patterns were established on the genotypes SRJ23 and BRS Estilo (centroid 1); CBR1,CBJ9, SRJ32 and BRS Pérola (centroid 2); CBJ3, CBJ8 and PMP13 (centroid 3); CBJ6, BRP46, CAP51 and CAP53 (centroid 4); PMJ18, SRJ19, SRJ26 and PJP40 (centroid 5); PMJ11 and PMP15 (centroid 6); CBJ7, SRJ34, CAP50 and CAP52 (centroid 7); CBJ4, CBJ5, SRP35 and IAC Formoso (centroid 8); PMJ10, SRJ24 and SRP37 (centroid 9); BRP47 and IPR Tiziu (centroid 10); PMP12, BRS Esplendor and IAC Imperador (centroid 11); PMC14, SRJ31 and SRJ33 (centroid 12); PMJ17 and SRJ20 (centroid 13); PJP38 (centroid 14); SRJ28 and BRP48 (centroid 15); CBR2, SRJ30 and PJP45 (centroid 16); SRJ25, SRJ29, SRP36, PJP43, PJP44 and BRS MG Realce (centroid 17); SRJ22, SRJ27, PJP41 and PJP42 (centroid 18); PJP39 and BRS Supremo (centroid 19) and PMP16, SRJ21, CAP49, IPR Tangará and IPR Uirapuru (centroid 20). Based on the approach of the computational learning, 20 genetic patterns are defined, and thus the selections can be directed and determined agronomic aspects creating the possibility to identify the centroid of interest and consequently the genotype group or corresponding neurons.

The expressed tendencies in this study presented essential to identify which biometric approach is more adequate to the genotypes characterization thru the phenotypic measurements. It was defined that the common bean seeds length is the most contrasting character to differentiate the genetic constitutions, as well, the seeds productivity is directly related to genotypes with higher duration of reproductive period and cycle, because this attributes increase the yield compounds. The multivariate approaches based in the genetic distance estimate and define different genotypes patterns where the dendogram allows expressing 14 groups, the Tocher optimized method suggests 10 subdivisions, but the used of Artificial Neural Networks with non-supervised learning that better stratified the genotypes basing its inferences in estimates of 20 centroids established by 63 neurons and 31 synaptic associations.

4. Conclusions

The accurate characterization of bean genotypes can be based in the reproductive period, cycle and seeds length. Genotypes with longer cycle increase the potential of ramifications, legume and seeds magnitude per plant and increase the seeds yield independent of the commercial group.The use of biometric approach allows revealing patterns to the genotype grouping, being the patterns magnitude dependent of the intrinsic premises to the Standardized Average Euclidian Distance, Tocher optimized grouping and Artificial Neural Network with non-supervised learning. It is defined that the Artificial Neural Network are determinant to define associative patterns, being these inferences indispensable to the common beans genotype selection that answer the agronomic attributes and seeds production.

References

Carbonell, S. A. M., Chioratto, A. F., Resende, M. D. V., Dias, L. A. S., Beraldo, A. L. A., & Perina, E. F. (2007). Stability of common bean cultivars and lines in different environments in the state of São Paulo. *Bragantia, 66*, 193-201. https://doi.org/10.1590/S0006-87052007000200003

Carvalho, I. R., Nardino, M., Ferrari, M., Pelegrin, A. J., Demari, G., Szareski, V. J., … Maia, L. C. (2016). Genetic variability among common black bean (*Phaseolus vulgaris* L.) accessions in southern Brazil. *Aust J Crop Sci, 10*, 1474-1483. https://doi.org/10.21475/ajcs.2016.10.10.p7786

Cruz, C. D. (2013). Genes-a software package for analysis in experimental statistics and quantitative genetics. *Acta Scientiarum. Agronomy, 35*, 271-276. https://doi.org/10.4025/actasciagron.v35i3.21251

Demari, G., Souza, V. Q., Carvalho, I. R., Nardino, M., & Follmann, D. N. (2015). Bean in season not preferred subject to nitrogen doses and its impact on traits agronomic. *Enciclopédia Biosfera, 11*, 1102-1112.

Farinelli, R., & Lemos, L. B. (2010). Yield, agronomic efficiency, technological and nutritional characteristics of the common bean with nitrogen fertilization in no-tillage and conventional tillage. *Bragantia, 69*(1), 165-172. https://doi.org/10.1590/S0006-87052010000100021

Federer, W. T. (1956). Augmented (hoonuiaku) designs. *HawPlan Rec, 55*, 191-208.

Moura, M. M., Carneiro, P. C. S., Carneiro, J. E. S., & Cruz, C. D. (2013). Potential of characters for evaluating common bean plant architecture. *Pesquisa Agropecuária Brasileira, 48*, 417-425. https://doi.org/10.1590/S0 100-204X2013000400010

Nascimento, M., Peternelli, L. A., Cruz, C. D., Nascimento, A. C. C., Ferreira, R. de P., Bhering, L. L., & Salgado, C. C. (2013). Artificial neural networks for adaptability and stability evaluation in alfalfa genotypes. *Crop Breeding and Applied Biotechnology, 13*, 152-156. https://doi.org/10.1590/S1984-7033201 3000200008

Pedo, T., Koch, F., Delias, D. S., Aisenberg, G. R., Szareski, V. J., Carvalho, I. R., … Aumonde, T. Z. (2016). Antioxidant enzyme activity and initial growth of wheat, rye and bean under soil flooding. *International Journal of Current Research, 8*, 6635-36642.

R Core Team. (2015) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Singh, D. (1981). The relative importance of characters affecting genetic divergence. *Ind J Gen Plant Breed, 41*, 237-245.

Santos, J. A. S., Teodoro, P. E., Correa, A. M., Soares, C. M. G., Ribeiro, L. P., & Abreu, H. K. A. (2014). Agronomic performance and genetic divergence among cowpea genotypes in Cerrado/Pantanalecotone. *Bragantia, 73*, 377-382. https://doi.org/10.1590/1678-4499.0250

Streck, E. V., Kämpf, N., Dalmolin, R. S., Klamt, E., Nascimento, P. C. do, Schneider, P., … Pinto, L. F. S. (2008). *Solos do Rio Grande do Sul* (2nd ed., p. 222). Porto Alegre, Emate.

Szareski, V. J., Carvalho, I. R., Kehl, K., Levien, A. M., Nardino, M., Demari, G., … Aumonde, T. Z. (2017). Univariate, multivariate techniques and mixed models applied to the adaptability and stability of wheat in the Rio Grande do Sul State. *Genetics and Molecular Research, 16*, 1-13. https://doi.org/10.4238/gmr 16039735

Teodoro, P. E., Barroso, L. M. A., Nascimento, M., Torres, F. E., Sagrilo, E., Dos Santos, A., & Ribeiro, L. P. (2015). Artificial neural networks to identify semi prostrate cowpea genotypes with high phenotypic adaptability and stability. *Pesq. Agropec. Bras., 50*(11), 1054-1060. https://doi.org/10.1590/S0100-204X 2015001100008

Torres, F. E., Teodoro, P. E., Sagrilo, E., Ceccon, G., & Correa, A. M. (2015). Genotype × environmentinteraction in semiprostradecowpeagenotypes via mixedmodels. *Bragantia, 74*, 255-260. https://doi.org/10.1590/1678-4499.0099

Troyjack, C., Dubal, I., Koch, F., Szareski, V. J., Pimentel, J. R., Carvalho, I. R., … Pedo, T. (2017). Attributes of growth, physiological quality and isoenzymatic expression of common bean seeds produced under the effect of gibberellic acid. *Aust J Crop Sci, 11*, 1116-1122. https://doi.org/10.21475/ajcs.17.11.09.pne531