Heritability profiles defined by hierarchal models and artificial neural networks for dual-purpose wheat attributes

I.R. Carvalho[1], J.A.G. Da Silva[1], L.L. Ferreira[2], V.E. Bubans[1], M.H. Barbosa[3], R.B. Mambrin[1], S.M. Fachi[4], G.G. Conte[3] and V.Q. de Souza[5]

[1] Universidade Regional do Noroeste do Estado do Rio Grande do Sul, Ijuí, RS, Brasil
[2] Centro Universitário de Mineiros, Mineiros, GO, Brasil
[3] Universidade Federal de Pelotas, Capão do Leão, RS, Brasil
[4] Universidade Federal de Santa Maria, Santa Maria, RS, Brasil
[5] Universidade Federal do Pampa, São Gabriel, RS, Brasil

Corresponding author: I.R. Carvalho
E-mail: carvalho.irc@gmail.com

ABSTRACT. Dual purpose wheat could be a good alternative for helping overcome the need to import this cereal in Brazil. To achieve this, development of cultivars with high yield is necessary. The contribution of genetics in defining traits is very important for directing breeding programs for the development of cultivars that provide the desired agronomic ideotype. We estimated heritability for 36 characters of agronomic importance in dual-purpose wheat. The inheritable genetic patterns were examined using linear trends, a Euclidean algorithm, factor analysis and artificial neural networks. The study was carried out during the crop seasons of 2011, 2012 and 2013. The experimental design was randomized block, arranged in a factorial scheme with three growing seasons (2011, 2012 and 2013) and five dual-purpose wheat genotypes (BRS Tarumã, BRS Umbu, BRS Figueira, BRS Guatambu and BRS 277) x three cuttings (first cutting, second cutting and third cutting), with three replicates. Deviance analysis or maximum likelihood was significant for the 36 characters. The length of the head of the main plant, plant height

before the first second cutting and dry mass of the seedlings showed high variability. The 36 characters expressed linear genetic dependence based on the Euclidean Algorithm; similar to what was found with the Tocher Optimized Clustering and Artificial Neural Networks K-means methods. Similar genetic trends for heritability profiles were obtained with factor analysis and Artificial Neural Networks by the Kohonem method. The use of Artificial Neural Networks through the Kohonem method gave the greatest efficacy in the definition of the genetic profiles needed to develop the recommended agronomic ideotype for the improvement of dual-purpose wheat.

INTRODUCTION

Wheat (*Triticum aestivum*) is one of the cold season species with greatest importance for agriculture worldwide. It is a key crop in small to large Brazilian proprieties, and generally integrates crop rotation systems and/or succession with crops such as soybean and maize, in no-till systems, as well assisting in the sustainability of cultivated areas and economic viability for producing units (Valério et al., 2009).

In Brazil, wheat is grown mainly in the states of Rio Grande do Sul and Paraná. In the 2018 season, production was 5.4 million tons, with a grain yield greater than two and a half tons per hectare (Conab, 2019). However, Brazil is still not self-sufficient for this crop, and imports are necessary. The use of dual-purpose wheat to increase the production of this cereal can be an alternative to supply the demand. Its use for both forage and grain enables planting in systems that integrate crop and livestock production. In this sense, the aim is to enhance the competitiveness of Brazilian wheat production, with the need to increase yield potential, by developing new cultivars that are responsive to management techniques and to growing environments (Smanhotto et al., 2006).

From the point of view of genetic improvement, knowledge on the genetic and phenotypic variances as well as the associations between the components of grain production and its heritability are of great interest for the development of strategies and suitable selections (Kenga et al., 2006). Heritability shows the proportion of genetic effects in the character expression; this proportion comprises the portion whose expression is not influenced by the environment. It can be explored in breeding programs in order to increase genetic gains, which may or may not increment the production potential of the crops (Hartwig et al., 2007; Carvalho et al., 2018).

For obtaining a superior cultivar, it is necessary to select many attributes together. In this context, there is a need to understand the magnitude of the heritability of the characteristics, its associations with other characters, and which of these forms a multivariate profile that, when selecting a certain attribute of interest, the breeder can indirectly select other characteristics in an indirect way (Wright, 1921, Wright, 1923; Carvalho et al., 2018).

There is a lack of information on the genetic fraction determining the phenotypic expression of dual-purpose wheat characters; also, there is a need to implement models that explain heritability profiles in a species of agricultural importance. To this end, we estimated heritability for 36 characters of agronomic importance for the dual-purpose wheat and investigated the inheritable genetic patterns using linear trends, the Euclidean algorithm, factor analysis and Artificial Neural Networks.

MATERIAL AND METHODS

This study was carried out during the crop seasons of 2011, 2012 and 2013 in Frederico Westphalen - RS (27º39'56"S, 53º42'94" W, with 490m altitude). The soil is classified as Anionic Acrudox (Santos et al., 2018), and climate subtropical Cfa type, by Köppen (Kottek et al., 2006).

The experimental design was randomized block, arranged in factorial scheme with three growing seasons (2011, 2012 and 2013) x five dual-purpose wheat genotypes (BRS Tarumã, BRS Umbu, BRS Figueira, BRS Guatambu and BRS 277) x three cutting management (first cutting, second cutting and third cutting), and the treatments were arranged in three replicates. The cuttings were carried out at 0.30 m of height from the soil level until the last expanded leave, and then a manual clip was realized of all experimental unit in the height 0f 0.10 m from the soil level (Carvalho et al., 2015).

The experimental units were composed of 12 rows, 2.00 m long each and spaced by 0.17 m. The population density employed for the dual-purpose wheat genotypes was three million plants per hectare. The management system used for all crop seasons was based on a no-till system with base fertilization of 250 kg.ha^{-1} with NPK in the formulation 10-20-20, by applying a total of 90 kg.ha^{-1} of nitrogen in the amidine form, fractionated in full tillering and after each cutting. The control of pests, diseases and weeds were done in preventive manner. The evaluations were conducted on six central lines of the utile area of the experimental unit, forgetting 0.50 m at each end. In Table 1 are the 36 characters measured in the study, with the abbreviation, description and unit of measurement for each character.

The data obtained were submitted to the presuppositions where the normality and homogeneity of the residual variances were verified. Afterwards it was carried out the Deviance and maximum likelihood ratio (LRT) analysis, with 0.05 of probability by the chi-square test (X^2) in order to identify the model identity. It was estimated the phenotypic, genetic and environmental variances, therefore with these estimates it was obtained the parameter of heritability for each character of interest.

The construction of the inheritable genetic profiles were obtained through linear correlation models with significance based on the t test at 0.05 of probability, Euclidean algorithm with use of the Z statistic attributions. Thus, it was revealed co-optic correlation (CCC) establishing associations between the data matrix and the genetic distances of the heritability, and were estimated the matrix stress index, matrix distortion, Mantel test, kurtosis and standard deviation. After, the distances matrix between the genetic parameters were submitted to the Unweighted Pair Grouping Method with Arithmetic Mean (UPGMA), and for the construction of the heritabilities dendrogram the optimized Tocher method was used, defining the maximum homogeneity within the group and inter-group heterogeneity of parameters. It was processed the analysis of factors for the rotation of the genetic parameters, where the eigenvectors were obtained for each determining factor to form each

profile. Subsequently, the matrix of characters' heritability was submitted to unsupervised computational learning through Artificial Neural Networks using the K-means algorithms and Kohonen Map. (Carvalho, 2018). Statistical analyzes were performed using the Genes software (Cruz, 2013).

Table1. Abbreviations, descriptions and units of characters for the investigation of dual purpose wheat cultivars.

Variable	Description	Unit
NAA	Number of tillers per plant measured in full tillering	units
APA	Plant height measured in the tillering	cm
DIA	Tillers stem diameter measured in full tillering	mm
NAU	Number of tillers per plant before the first cutting	units
APU	Plant height measured before the first cutting	cm
NAD	Number of tillers per plant before the second cutting	units
APD	Plant height before the second cutting	cm
DAD	Tiller stem diameter before the second cutting	mm
NAT	Number of tillers per plant before third cutting	units
APT	Plant height before third cutting	cm
DIT	Stem diameter before the third cutting	mm
MVH	Green matter yield per hectare	kg.ha^{-1}
MSH	Dry matter yield per hectare	kg.ha^{-1}
PMM	Mineral matter in forage	%
PTN	Crude protein in forage	%
AIS	Height of head insertion	cm
ANG	Angle of the flag leaf	Degrees
DIF	Diameter of the tillers' stem	cm
NFF	Number of fertile tillers	units
NEQ	Number of spikes per square meter	units
CSM	Length of the main head	cm
NSM	Number of seeds of the main head	units
MSM	Seeds mass of the main head	g
MMS	Thousand seeds mass	g
RS	Seeds' yield per hectare	kg.ha^{-1}
CSA	Length of tillers head	cm
MEM	Mass of main head	g
NSA	Number of seed of tillers head	units
MEA	Mass of tillers head	g
PCG	First germination counting	units
PG	Germinated seeds	%
RAD	Radicle length	cm
CPA	Seedlings' shoot length	cm
MSP	Seedlings dry matter	g
EC	Field emergence	%
IVE	Index of emergence speed	dimensionless

RESULTSAND DISCUSSION

Deviance and genetic parameters

The deviance analysis or maximum likelihood (LRT) showed significance at the 0.05 of probability of the chi-square test (X^2) for 36 characters in measured in dual-purpose wheat. In this way, the reliability of the estimates of the genetic NAA, APA, DIA, NAU, APU, NAD, APD, DAD, NAT, APT, DIT, MVH, MSH, PMM, PTN, AIS, ANG, DIF, NFF, NEQ, CSM, NSM, MSM, MMS, RS, CSA, MEM, NSA, MEA, PCG, PG, RAD, CPA, MSP, EC and IVE.

The phenotypic expression of a given character is determined by the intrinsic genetic effects of the genotype characteristics and the effects attributed to the growing environment. In this context, it is essential to use a parameter that can show when the magnitude evidenced by the character is due to the genetic variation. In this way, the heritability (H^2) allows to determine which are the relevant genetic effects in this situation as well, its magnitude can determine the experimental accuracy of trait in evidence (Carvalho et al., 2016). Research by Hallauer and Miranda (1988), classified this parameter as being high ($H^2 > 0.70$), medium or intermediate ($0.30 > H^2 < 0.70$) and low ($H^2 < 0.30$).
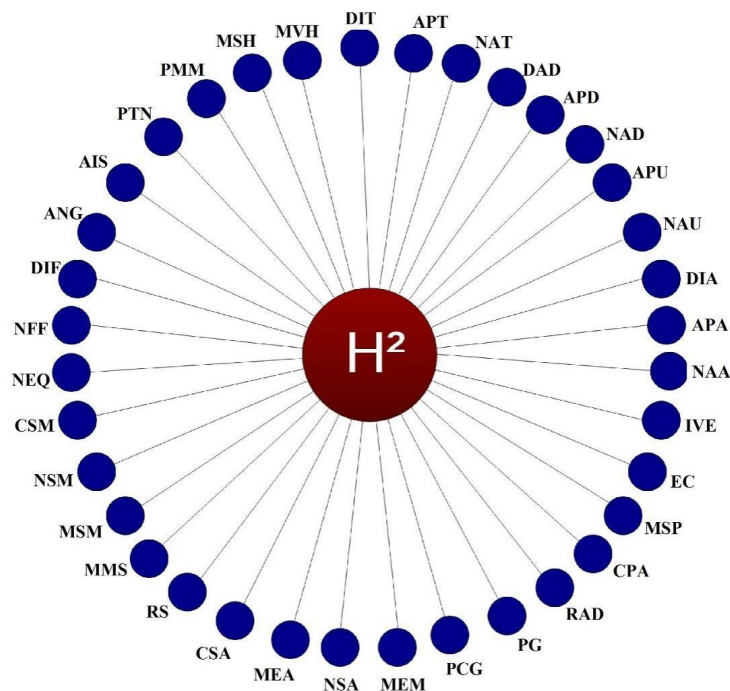

Figure 1. Hypothetical Neural Network for 36 characters measured weighted by the broad sense heritability (H^2) for dual-purpose wheat. Number of tillers per plant measured at full tillering (NAA, units), Plant height measured in the tillering (APA, cm), Tillers' stem diameter measured at full tillering (DIA, mm), Number of tillers per plant before the first cutting (NAU, units), Plant height measured before the first cutting (APU, cm), Number of tillers per plant before the second cutting (NAD, units), Plant height before the second cutting (APD, cm), Tiller stem diameter before the second cutting (DAD, mm), Number of tillers per plant before third cutting (NAT, units), Plant height before third cutting (APT, cm), Stem diameter before the third cutting (DIT, mm), Green matter yield per hectare (MVH, kg.ha⁻¹), Dry matter yield per hectare (MSH, kg.ha⁻¹, Mineral matter in forage (PMM, %), Crude protein in forage (PTN, %), Height of head insertion (AIS, cm), Angle of the flag leaf (AN, Degrees), Diameter of the tillers' stem (DIF, cm), Number of fertile tillers (NFF, units), Number of spikes per square meter (NEQ, units), Length of the main head (CSM, cm), Number of seeds of the main head (NSM, units), Seeds mass of the main head (MSM, g), Thousand seeds mass (MMS, g), Seeds' yield per hectare (RS, kg.ha⁻¹), Length of tillers head (CSA, cm), Mass of main head (MEM, g), Number of seed of tillers head (NSA, units), Mass of tillers head (MEA, g), First germination counting (PCG, units), Germinated seeds (PG, %), Radicle length (RAD, cm), Seedlings' shoot length (CPA, cm), Seedlings dry matter (MSP, g), Field emergence (EC, %), Index of emergence speed (IVE, dimensionless).

In possession of heritability estimates, the important traits of dual-purpose wheat (Figure 2) it was found that CSM, CSA, APD, APU, APT, APA and MSP were considered characters with intermediate heritability, where there was an amplitude of 0.40 to 0.65 indicating that these characters are predominantly determined by genetic effects. However, the remaining

29 characters showed low magnitudes for this parameter, being these mostly affected by the effects of the growing environment, possibly controlled by a large number of genes which tend to differentially respond across the modifications imposed by the environment.

Due to the great magnitude of measured characters and the difficulty of interpreting separately each genetic tendency, a new approach was developed in order to understand the inheritable profile of the characters and to present a logical tendency of association between the estimates of this parameter. In order to prove the efficacy of the definition of genetic profiles, different univariate, multivariate and unsupervised computational learning methods were applied, based on linear correlations (CL), mean Euclidean distance and UPGMA grouping (EM), factor analysis (AF), k-means algorithm (KM), optimized clustering of Tocher (TC) and artificial neural networks applied to Kohonem Map (MK).
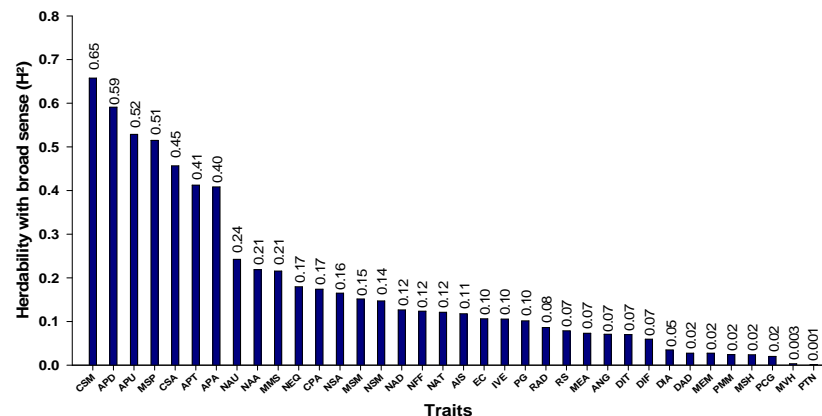

Figure 2. Broad sense heritability for 36 characters in dual-purpose wheat. number of tillers per plant measured in full tillering (NAA, units), Plant height measured in the tillering (APA, cm), Tillers' stem diameter measured in full tillering (DIA, mm), Number of tillers per plant before the first cutting (NAU, units), Plant height measured before the first cutting (APU, cm), Number of tillers per plant before the second cutting (NAD, units), Plant height before the second cutting (APD, cm), Tiller stem diameter before the second cutting (DAD, mm), Number of tillers per plant before third cutting (NAT, units), Plant height before third cutting (APT, cm), Stem diameter before the third cutting (DIT, mm), Green matter yield per hectare (MVH, kg.ha^{-1}), Dry matter yield per hectare (MSH, kg.ha^{-1}, Mineral matter in forage (PMM, %), Crude protein in forage (PTN, %), Height of head insertion (AIS, cm), Angle of the flag leaf (AN, Degrees), Diameter of the tillers' stem (DIF, cm), Number of fertile tillers (NFF, units), Number of spikes per square meter (NEQ, units), Length of the main head (CSM, cm), Number of seeds of the main head (NSM, units), Seeds mass of the main head (MSM, g), Thousand seeds mass (MMS, g), Seeds' yield per hectare (RS, kg.ha^{-1}), Length of tillers head (CSA, cm), Mass of tillers head (MEM, g), Number of seed of tillers head (NSA, units), Mass of tillers head (MEA, g), First germination counting (PCG, units), Germinated seeds (PG, %), Radicle length (RAD, cm), Seedlings' shoot length (CPA, cm), Seedlings dry matter (MSP, g), Field emergence (EC, %), Index of emergence speed (IVE, dimensionless).

Definition of inheritable profiles through linear correlations (CL)

The linear correlation analysis revealed 57 associations with level of significance at 5% of probability by t test (Table 2). The correlations obtained had both directions. The magnitude of the correlations varied from 0.61 to 0.91 in the positive direction and - 0.64 to -0.89 in the negative direction.

The character that received the most influence on the obtained heritability combinations was RS, which was positively correlated with APT (r = 0.84), APU (r = 0.68),

MVH (r = 0.77), NAV (r = 0.66) and PTN (r = 0.74). The influence of the number of tillers on wheat yield is already known; its importance is characterized by their participation as part of yield components of plants and as probable suppliers of products needed by the main stem (Merotto Junior, 1995). It was observed that the height of the plant is positively related to the yield of seeds. Similar results were observed by Mohammad et al. (2001) and Dhonde et al. (2000), which found positive correlations between grain production and plant height, reinforcing the results of the present study. The contribution of biologically active mass accumulation, due to the increase in and plant height, may contribute to seed yield, because of larger photosynthetic active surface. Consequently, it increases the production capacity of photosynthesis' substrates that will be used for the grain filling process.

Seed yield was negatively associated with ANG (r = -0.89), DAD (r = -0.71) and NAD (r = -0.79). Leaf architecture is an important aspect of the canopy structure of the crop because it allows a better use of solar energy, increasing the productive potential (Fioreze, 2011). The use of this solar energy by the plant is dependent on the size, shape, angle and azimuth orientation of the leaves (Castro and Klunge, 1999). In heritability correlations, a high negative correlation of ANG with RS was observed. This fact can be associated in the study by Lozano and Leaden (2002), evaluating a growth reducer in two wheat cultivars, observing significant gains in yield and changes in leaf architecture. According to the authors, the application of the regulator caused a change in the leaf architecture, with the modification of the angulation of the flag leaf, which was more erect, factors that may be linked to the increment in grain yield.

The negative correlation observed between DAD x RS and NAD x RS can be explained by the fact that the studied species are dual-purpose cultivars (forage and grain production), when a greater amount of photoassimilates is used to develop vegetative characteristics, which will be removed in grazing /cutting, reducing the filling capacity of grains, thus compromising seed yield. According to Zilio et al. (2018) some dual-purpose wheat cultivars may present reduction in grain yield after consecutive cuttings, as well the value of the hectoliter weight and the thousand seeds mass.

The most influential character in the combinations was NAD. It was shown to be an important characteristic of heritability, showing positive correlations with DAD (r = 0.66) and with ANG (r = 0.67). Negative correlations with APT (r = -0.76), STP (r = -0.69) and RS (r = -0.79). Still, the plants tillering ability occurs by several factors, including the coordination and synchronization between the development of leaves and tillering, known as occupation sites (Valerio et al., 2008).

The great majority of NAD positive correlations were those that negatively influenced yield components. This fact is due to the stress caused by the first cutting, which caused the apical dominance to stop, making the plant to destine most of its photoassimilates for vegetative growth, compromising important components such as PTN and RS. Similar results were found by Martin et al. (2010) in which THE cutting management reduced all yield components, as well as the grain yield for all cultivars in the analysis.

The MSH presented negative correlation with NFF (r = -0.77) and NSM (r = -0.71). Positive correlations were obtained with NEQ (r = 0.66), CSA (r = 0.78), NSA (0.77) and the highest positive correlation between the pairs of characters (r = 0.91) was CSM. These associations demonstrate that plants with higher vegetative growth provide higher dry matter yield per hectare, satisfactorily contributing to the forage function of the dual-purpose wheat.

Studies reveal that the selection of adequate genotypes to the foraging aspect must express high dry mass per area unit, high protein ratio and adequate animal acceptance (Pariz et al., 2010). However, they present a reduction in yield by their negative correlation with NFF and NSM.

Table 2. Linear correlation coefficients between broad sense heritabilities of dual-purpose wheat characters.

ANG	x	MSP	0.66*	CSM	x	NSA	0.73*	MSH	x	NFF	-0.77*	NAU	x	NAD	-0.70*
ANG	x	EC	0.72*	CSM	x	PCG	-0.82*	MSM	x	NSA	-0.66*	NAU	x	RS	0.66*
ANG	x	RS	-0.89*	CSM	x	RAD	0.64*	MVH	x	ANG	-0.79*	NEQ	x	CSM	0.69*
APA	x	APD	-0.68*	DAD	x	NSM	0.76*	MVH	x	RS	0.77*	NFF	x	NSM	0.66*
APD	x	AIS	0.72*	DAD	x	RS	-0.71*	MVH	x	MEM	-0.75*	NFF	x	CSA	-0.76*
APD	x	PCG	-0.76*	DAD	x	EC	0.76*	MVH	x	PG	-0.71*	NFF	x	NSA	-0.80*
APD	x	PG	-0.67*	DIF	x	PG	0.64*	NAA	x	ANG	0.64*	NSM	x	MMS	-0.79*
APT	x	ANG	-0.80*	DIT	x	AIS	-0.78*	NAA	x	EC	0.67*	PCG	x	RAD	-0.64*
APT	x	RS	0.84*	MMS	x	EC	-0.75*	NAD	x	DAD	0.66*	PMM	x	MEM	0.63*
APT	x	MSP	-0.67*	MMS	x	IVE	0.79*	NAD	x	APT	-0.76*	PTN	x	RS	0.74*
APU	x	NAT	-0.67*	MSH	x	NEQ	0.66*	NAD	x	PTN	-0.69*	RS	x	MEM	-0.75*
APU	x	PTN	0.72*	MSH	x	CSM	0.91*					RS	x	EC	-0.70*
APU	x	RS	0.68*	MSH	x	NSM	-0.71*	NAD	x	ANG	0.67*				
CSA	x	IVE	-0.76*	MSH	x	CSA	0.78*	NAD	x	NSM	-0.68*				
CSM	x	CSA	0.76*	MSH	x	NSA	0.77*	NAD	x	RS	-0.79*				

*significant at 5% of probability by t test. number of tillers per plant measured in full tillering (NAA, units), Plant height measured in the tillering (APA, cm), Tillers' stem diameter measured in full tillering (DIA, mm), Number of tillers per plant before the first cutting (NAU, units), Plant height measured before the first cutting (APU, cm), Number of tillers per plant before the second cutting (NAD, units), Plant height before the second cutting (APD, cm), Tiller stem diameter before the second cutting (DAD, mm), Number of tillers per plant before third cutting (NAT, units), Plant height before third cutting (APT, cm), Stem diameter before the third cutting (DIT, mm), Green matter yield per hectare (MVH, kg.ha^{-1}), Dry matter yield per hectare (MSH, kg.ha^{-1}, Mineral matter in forage (PMM, %), Crude protein in forage (PTN, %), Height of head insertion (AIS, cm), Angle of the flag leaf (AN, Degrees), Diameter of the tillers' stem (DIF, cm), Number of fertile tillers (NFF, units), Number of spikes per square meter (NEQ, units), Length of the main head (CSM, cm), Number of seeds of the main head (NSM, units), Seeds mass of the main head (MSM, g), Thousand seeds mass (MMS, g), Seeds' yield per hectare (RS, kg.ha^{-1}), Length of tillers head (CSA, cm), Mass of tillers head (MEM, g), Number of seed of tillers head (NSA, units), Mass of tillers head (MEA, g), First germination counting (PCG, units), Germinated seeds (PG, %), Radicle length (RAD, cm), Seedlings' shoot length (CPA, cm), Seedlings dry matter (MSP, g), Field emergence (EC, %), Index of emergence speed (IVE, dimensionless).

Definition of inheritable profiles across the mean Euclidean distance (EM)

The analysis of the mean Euclidean distance obtained a co-phonetic correlation coefficient of the UPGMA method of 0.89 (Figure 3), which was significant by the Mantel test (P >0.01) indicating good distances representation in the hierarchical clustering (Figure 4). This coefficient is in accordance with Sokal and Rohlf (1962), indicating good reliability of the established clusters.

In the dendrogram, the Euclidian distance analysis using broad sense heritability as a reference, by the UPGMA method, to define profiles of the multivariate heritability showed eight parameters in relation to its inheritable trend (Figure 2). The mean (μ: 0.0778) of the distances matrix was evidenced as the separation criterion of the heritability profiles with restricted sense.

The profile I grouped DIT, ANG, MEA, DIF, EC, RS, RAD and PG. The profile II comprised NAD, AIS, NEQ, MSM, NSA, CPA, NAT and NSM. The profile III grouped NAA, NAU and MMS. The profile IV assembled only the characters MSH and PMM, where DIA, DAD, MEA and PCG were gathered together at the profile V. For instance, the profile VI,

gathered APD, CSM, APU, APT, CSA and MSP. The last, profile VIII, grouped two characters, being MVH and PTN.

As for the characters, those that were less distant were DIT and ANG, and those that distanced themselves from the others were MVH, PTN and CMS. The characterization of groups of genetic behaviors allows the grouping of agronomic characteristics that show uniqueness as to the portion of the character variation, which depends on the genetic factor. Assumptions about the nature of the expression, amount of genes involved and quantification of the genes expression may be associated with this singular behavior of the variables.
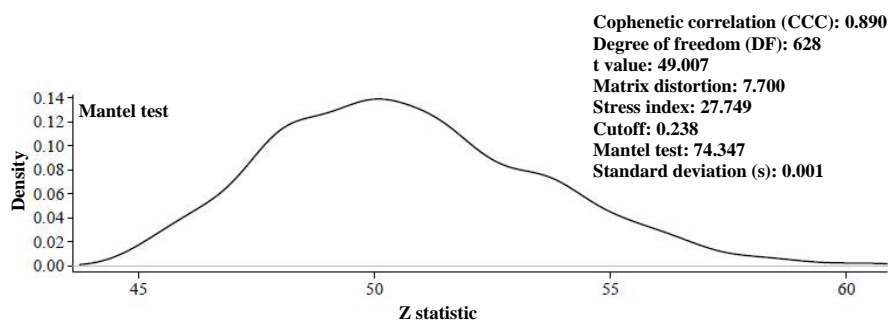

Cophenetic correlation (CCC): 0.890
Degree of freedom (DF): 628
t value: 49.007
Matrix distortion: 7.700
Stress index: 27.749
Cutoff: 0.238
Mantel test: 74.347
Standard deviation (s): 0.001

Figure 3. Mantel multivariate test, cophenetic correlation and stress index of the matrix from 36 broad sense heritability in dual-purpose wheat.
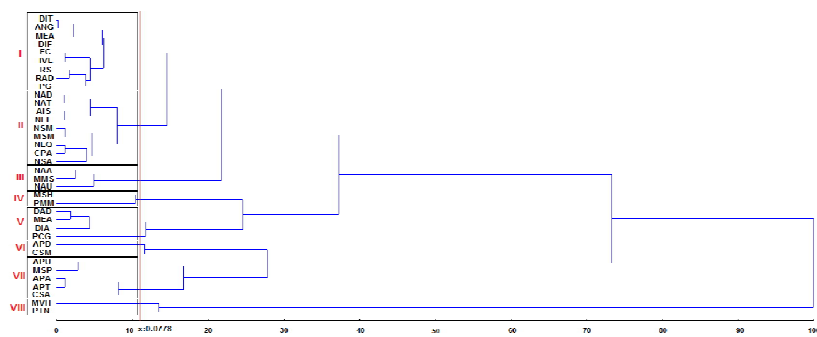

Figure 4. Hierarchical clustering obtained through Mean Euclidian distance and UPGMA clustering for the broad sense heritability of 36 characters in dual-purpose wheat. number of tillers per plant measured in full tillering (NAA, units), Plant height measured in the tillering (APA, cm), Tillers' stem diameter measured in full tillering (DIA, mm), Number of tillers per plant before the first cutting (NAU, units), Plant height measured before the first cutting (APU, cm), Number of tillers per plant before the second cutting (NAD, units), Plant height before the second cutting (APD, cm), Tiller stem diameter before the second cutting (DAD, mm), Number of tillers per plant before third cutting (NAT, units), Plant height before third cutting (APT, cm), Stem diameter before the third cutting (DIT, mm), Green matter yield per hectare (MVH, kg.ha^{-1}), Dry matter yield per hectare (MSH, kg.ha^{-1}, Mineral matter in forage (PMM, %), Crude protein in forage (PTN, %), Height of head insertion (AIS, cm), Angle of the flag leaf (AN, Degrees), Diameter of the tillers' stem (DIF, cm), Number of fertile tillers (NFF, units), Number of spikes per square meter (NEQ, units), Length of the main head (CSM, cm), Number of seeds of the main head (NSM, units), Seeds mass of the main head (MSM, g), Thousand seeds mass (MMS, g), Seeds' yield per hectare (RS, kg.ha^{-1}), Length of tillers head (CSA, cm), Mass of tillers head (MEM, g), Number of seed of tillers head (NSA, units), Mass of tillers head (MEA, g), First germination counting (PCG, units), Germinated seeds (PG, %), Radicle length (RAD, cm), Seedlings' shoot length (CPA, cm), Seedlings dry matter (MSP, g), Field emergence (EC, %), Index of emergence speed (IVE, dimensionless).

Definition of inheritable profiles through the factor analysis (AF)

The analysis of factors related to heritability of dual-purpose wheat characters is shown in Table 3. Estimates of the eigenvalues were determined using the criterion of more than 80% of the total existing of variation explanation (Cruz et al., 2014). Thus, 83.9% of the variation was obtained with V factors. The factor's load was used as a criterion for character grouping and magnitude of loads. If above 0.70, it reveals that the characters are correlated and allowed to be grouped; loads between 0.50 and 0.70 do not allow to be defined as groups; and loads of less than 0.50 do not make it possible to group characters (Johnson; Wichern, 1992).

Table 3. Factor analysis applied to broad sense heritability of 36 characters in dual-purpose wheat.

Eigenvectors estimates				Characters	Final Factors after rotation					Communality
	Root	%Root	%Accumulate		Factor I	Factor II	Factor III	Factor IV	Factor V	
I	9.934	0.276	27.6	NAA	-0.41	0.16	-0.38	0.44	0.58	0.86
II	7.633	0.212	48.8	APA	0.01	-0.15	0.56	-0.13	-0.26	0.43
III	5.771	0.160	64.8	DIA	-0.03	-0.01	0.83	0.09	0.16	0.72
IV	3.718	0.103	75.2	NAU	0.74	0.14	-0.31	-0.47	0.24	0.94
V	3.135	0.087	83.9	APU	0.72	0.22	-0.11	0.41	-0.37	0.89
VI	2.375	0.066	90.5	NAD	-0.84	0.44	0.19	-0.01	-0.19	0.97
VII	1.505	0.042	94.6	APD	-0.01	-0.04	-0.89	-0.02	0.12	0.80
VIII	1.032	0.029	97.5	DAD	-0.78	0.01	-0.48	0.06	-0.33	0.94
IX	0.897	0.025	100.0	NAT	-0.47	0.28	0.47	-0.21	0.48	0.79
X	.	.	100.0	APT	0.78	-0.39	-0.17	-0.16	-0.22	0.86
XI	.	.	100.0	DIT	0.07	-0.26	0.28	0.09	0.89	0.96
XII	.	.	100.0	MVH	0.76	0.21	-0.21	-0.39	-0.29	0.90
XIII	.	.	100.0	MSH	-0.14	0.78	0.04	0.57	0.15	0.98
XIV	.	.	100.0	PMM	-0.29	-0.39	-0.48	-0.20	0.28	0.58
XV	.	.	100.0	PTN	0.84	-0.05	0.00	0.37	0.20	0.90
XVI	.	.	100.0	AIS	0.43	0.04	-0.08	0.11	-0.85	0.93
XVII	.	.	100.0	ANG	-0.83	0.10	-0.12	0.24	0.29	0.85
XVIII	.	.	100.0	DIF	-0.01	0.14	0.78	0.14	0.21	0.69
XIX	.	.	100.0	NFF	0.17	-0.91	-0.02	-0.07	0.20	0.90
XX	.	.	100.0	NEQ	-0.21	0.20	0.08	0.81	0.24	0.81
XXI	.	.	100.0	CSM	0.09	0.67	-0.14	0.69	0.13	0.97
XXII	.	.	100.0	NSM	-0.43	-0.14	-0.64	0.11	-0.47	0.85
XXIII	.	.	100.0	MSM	0.30	-0.72	-0.14	-0.16	0.08	0.67
XXIV	.	.	100.0	MMS	0.41	-0.15	0.58	-0.55	0.17	0.86
XXV	.	.	100.0	RS	0.96	-0.05	0.04	-0.07	-0.22	0.98
XXVI	.	.	100.0	CSA	-0.10	0.65	-0.07	0.54	-0.38	0.88
XXVII	.	.	100.0	MEM	-0.76	-0.51	-0.06	0.17	0.03	0.87
XXVIII	.	.	100.0	NSA	0.14	0.95	-0.17	0.05	0.15	0.98
XXIX	.	.	100.0	MEA	0.05	-0.02	0.30	0.90	-0.17	0.93
XXX	.	.	100.0	PCG	-0.40	-0.32	0.31	-0.69	0.07	0.85
XXXI	.	.	100.0	PG	-0.38	-0.20	0.75	0.09	0.31	0.84
XXXII	.	.	100.0	RAD	0.48	0.49	0.16	0.44	-0.07	0.70
XXXIII	.	.	100.0	CPA	-0.05	-0.54	0.31	0.30	-0.53	0.76
XXXIV	.	.	100.0	MSP	-0.66	0.10	-0.16	-0.27	0.05	0.55
XXXV	.	.	100.0	EC	-0.67	0.21	-0.57	0.24	0.18	0.92
XXXVI	.	.	100.0	IVE	0.44	-0.36	0.25	-0.58	0.41	0.89

+ number of tillers per plant measured in full tillering (NAA, units), Plant height measured in the tillering (APA, cm), Tillers' stem diameter measured in full tillering (DIA, mm), Number of tillers per plant before the first cutting (NAU, units), Plant height measured before the first cutting (APU, cm), Number of tillers per plant before the second cutting (NAD, units), Plant height before the second cutting (APD, cm), Tiller stem diameter before the second cutting (DAD, mm), Number of tillers per plant before third cutting (NAT, units), Plant height before third cutting (APT, cm), Stem diameter before the third cutting (DIT, mm), Green matter yield per hectare (MVH, kg.ha[-1]), Dry matter yield per hectare (MSH, kg.ha[-1], Mineral matter in forage (PMM, %), Crude protein in forage (PTN, %), Height of head insertion (AIS, cm), Angle of the flag leaf (AN, Degrees), Diameter of the tillers' stem (DIF, cm), Number of fertile tillers (NFF, units), Number of spikes per square meter (NEQ, units), Length of the main head (CSM, cm), Number of seeds of the main head (NSM, units), Seeds mass of the main head (MSM, g), Thousand seeds mass (MMS, g), Seeds' yield per hectare (RS, kg.ha[-1]), Length of tillers head (CSA, cm), Mass of tillers head (MEM, g), Number of seed of tillers head (NSA, units), Mass of tillers head (MEA, g), First germination counting (PCG, units), Germinated seeds (PG, %), Radicle length (RAD, cm), Seedlings' shoot length (CPA, cm), Seedlings dry matter (MSP, g), Field emergence (EC, %), Index of emergence speed (IVE, dimensionless).

Factor I was the one that presented the greatest number of characters, being: NAU, APU, NAD, DAD, APT, MVH, PTN, ANG, RS and MEM. Factor II exhibited higher loads for the MSH, NFF, MSM and NSA characters. Factor III composed the characters DIA, DIF, PG and APD. The factor IV included the characters NEQ and MEA. The factor V, the characters DIT 0.89 and negative charge for AIS -0.85. The highest commonality values were found for the characters NAD 0.97, DIT 0.96, RS 0.98 and NSA 0.98.

Definition of inheritable profiles through the k-means clustering algorithm (KM)

The analysis of main components, using the K-means clustering, resulted in the formation of eight groups with two main components that make up a total variance of 96.55% (Figure 5). Group I was composed of the largest number of characters, totaling 14 variables (NSA, NEQ, AIS, EC, NAT, RAD , NAD, DIT, MEM, ANG, NSM, PG, NFF, MSM), suggesting that they are part of a single heterotic group, but compose the lowest values for the two main components presented.
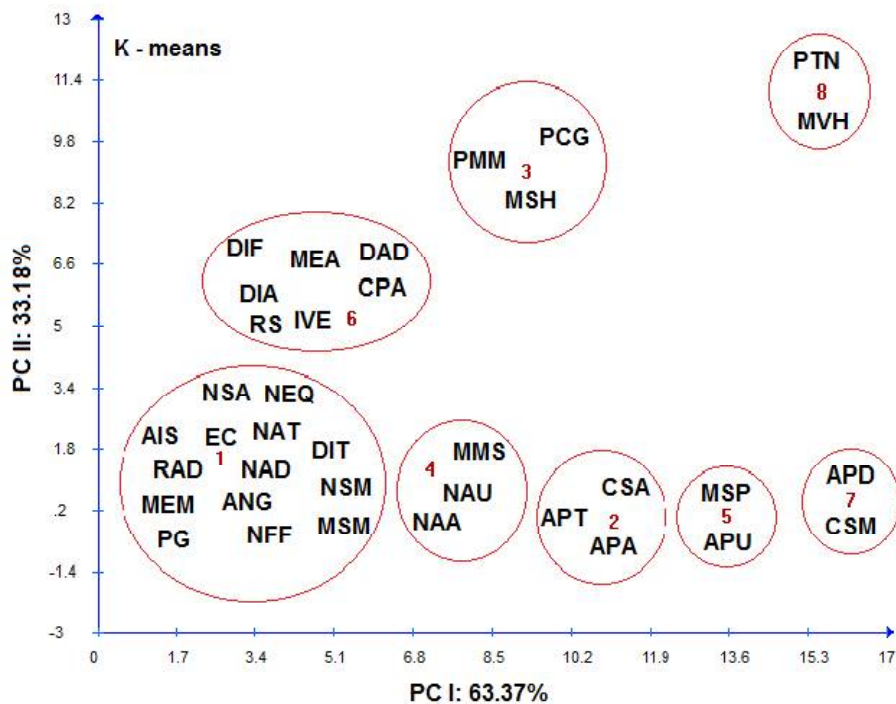

Figure 5. Main compounds analysis using K-means clustering referent to the broad sense heritability of 36 characters in dual-purpose wheat.

Group VI was the second group with the largest number of characters, totaling seven variables (DIF, MEA, DAD, DIA, CPA, RS, IVE), followed by groups II (CSA, APT APA), III (PCG, PMM, MSH) and IV (MMS, NAU, NAA) respectively. The groups V, VII

(MSP, APU) and VIII (PTN, MVH) had only two characters next in the group, including Group VIII, which presents the two characters with the highest values of the variance presented between the two components.

Definition of Inheritable Profiles Through Optimized Tocher Grouping (TC)

Tocher's optimized cluster afforded the fragmentation of the 36 characters of the dual-purpose wheat genotypes into eight profiles of distinct groups. This method establishes groups so that there is homogeneity within the group and heterogeneity between groups. In addition to this factor, it performs grouping keeping the criterion that intra-group distances are always smaller than intergroups (Cruz and Regazzi, 2001). As a result, the formation of these groups is relevant information on the choice of s characteristics of agronomic interest in breeding programs (Table 4).

Table 4. Tocher Optimized Clustering and Artificial Neural Network estimated by Kohonem Map for the broad sense heritability of dual-purpose wheat.

Tocher Optimized Clustering

Perfil I	DIT	ANG	MEA	RAD	RS	DIA	IVE	EC	PG	AIS	NFF	NAT	NAD	NSM	MSM
Perfil II	APA	APT	CSA												
Perfil III	NEQ	CPA	NSA	MMS	NAA	NAU									
Perfil IV	DAD	MEM	DIF	PCG											
Perfil V	APU	MSP	APD												
Perfil VI	NSH	PMM													
Perfil VII	MVH	PTN													
Perfil VIII	CSM														

Artificial Neural Network - Kohonen Map

Perfil I	APA	APU	APD	APT	CSM	CSA	MSP				
Perfil II	MVH	PTN									
Perfil III	DIA	DAD	MSH	PMM	MEA	PCG					
Perfil IV	NAA	NAU	NAD	NAT	NEQ	NSM	MSM	NSA	CPA	MMS	
Perfil V	DIT	AIS	ANG	DIF	NFF	RS	PG	RAD	EC	IVE	MEM

number of tillers per plant measured in full tillering (NAA, units), Plant height measured in the tillering (APA, cm), Tillers' stem diameter measured in full tillering (DIA, mm), Number of tillers per plant before the first cutting (NAU, units), Plant height measured before the first cutting (APU, cm), Number of tillers per plant before the second cutting (NAD, units), Plant height before the second cutting (APD, cm), Tiller stem diameter before the second cutting (DAD, mm), Number of tillers per plant before third cutting (NAT, units), Plant height before third cutting (APT, cm), Stem diameter before the third cutting (DIT, mm), Green matter yield per hectare (MVH, kg.ha-1), Dry matter yield per hectare (MSH, kg.ha-1, Mineral matter in forage (PMM, %), Crude protein in forage (PTN, %), Height of head insertion (AIS, cm), Angle of the flag leaf (AN, Degrees), Diameter of the tillers' stem (DIF, cm), Number of fertile tillers (NFF, units), Number of spikes per square meter (NEQ, units), Length of the main head (CSM, cm), Number of seeds of the main head (NSM, units), Seeds mass of the main head (MSM, g), Thousand seeds mass (MMS, g), Seeds' yield per hectare (RS, kg.ha -1), Length of tillers head (CSA, cm), Mass of tillers head (MEM, g), Number of seed of tillers head (NSA, units), Mass of tillers head (MEA, g), First germination counting (PCG, units), Germinated seeds (PG, %), Radicle length (RAD, cm), Seedlings' shoot length (CPA, cm), Seedlings dry matter (MSP, g), Field emergence (EC, %), Index of emergence speed (IVE, dimensionless).

In the first group formed 41.66% of the analyzed characters are concentrate (DIT, ANG, MEA, RAD, RS, DAY, IVE, CE, PG, ACN, NFF, NAT, NAD, NSM and MSM) suggesting that these are part of the same heterocyclic group. The profile II (APA, APT and CSA) and V (APU, MSP and APD) grouped 8.33% of the evaluated characters each, and profile III (NEQ, CPA, NSA, MMS, NAA and NAU) 16.7%. The profile IV (DAD, MEM, DIF and PCG) by 11.11%,

VI (NSH and PMM) and VII (MVH, PTN) by 5.55% each. Profile VIII (CSM) was the only one that exhibited only one character of interest with 2.77%.

Definition of inheritable profiles through artificial neural networks - Map of Kohonem (MK)

Artificial neural networks (ANNs) are indicated to deal with complex systems (Jana et al., 2012; Soares et al., 2015). For this study, the Kohonem Map was employed using 36 entries (neurons) represented by the dual-purpose wheat characters. The phenotypic matrix was submitted to the iterative procedures that defined a neural network with the topology of five centroids, establishing associative patterns among the characters (Figure 6). The patrons established by means of the characters MSP, CSA, CSM, APT, APD and APA (Centroid I), PTN, MVH (Centroid II), PCG, MEA, PMM, MSH, DAD and DIA (Centroid III) MSM, MMS, NSA CPA, NAU, NAD, NSM and NAT (Centroid IV), DID, DIF, NFF, RS, MEM, PG, RAD, EC, IVE, and ANG (Centroid V).
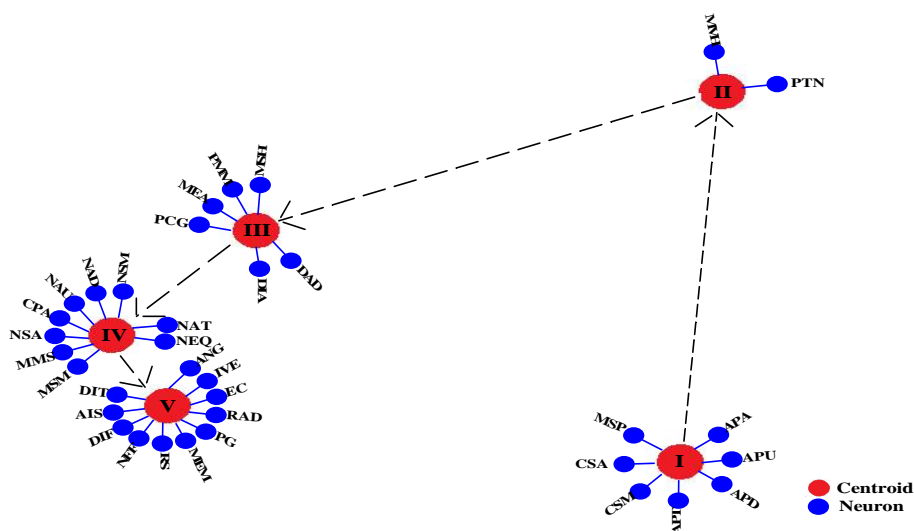

Figure 6. Artificial Neural Networks (ANNs) obtained by Kohonen Map defining the centroids (red) and neurons (blue) and synaptic links (blue lines).

For Nascimento et al. (2013), neural networks may draw the most complex features of existing data sets, therefore, are successfully used in a multitude of studies as Oliveira et al. (2013), which made use of artificial neural networks in the classification of banana autotetraploids. In wheat the approach with artificial network were performed for yield prediction and classification (Kadir, et al. 2014, Yasas, et al. 2016), these approaches show superiority for classification genetic (Sant'Anna et al., 2015). Arruda et al. (2013) used RNAs to obtain a digital map of soils using correlations between mapping units and environmental covariates. From this computational approach, five genetic patterns were defined and, therefore, selections can be directed to certain agronomic aspects of the dual-purpose wheat, being possible to identify the centroid of interest and consequently the group of the corresponding characters or neurons.

CL x MS x AF x KM x TC x MK

The grouping of genotypes by the UPGMA method was similar to the Tocher method, in relation to groups formed between more divergent characters. The agreement between these two techniques of analysis can be verified by the characters MVH and PTN belonging to Profile VII and CMS belonging to Tocher Profile VII, which were the same as those of the higher distance groupings by UPGMA (Table 3 and Figure 2). However, in relation to the distances between the characters, the Tocher grouping provided the formation of profiles made up of several characters, while in the UPGMA method the presentation of distances in the dendrogram made it possible to visualize more similar characters groupings. Similar results were found by Bertan et al. (2006), in a study comparing the morphological distance between wheat genotypes. The authors attributed this difference to the fact that the Tocher grouping always advocated the largest distances between groups in relation to the distance within the groups.

Regarding the other methods, similarities were observed between all the methods used. We can emphasize the presence of the MVH and PTN characters being always allocated in the same group for the Tocher, UPGMA and K-means clustering methods. Neural networks and factor analysis gathered the 36 characters in a smaller number of groups (V) in relation to the other methods that stipulated VIII groups. The V group (Centroids) obtained by the neural networks, grouped more similar characters in relation to the other methods, this fact can be justified by the centroid I, in which characters related to the heights of the plant after the cuttings were observed.

CONCLUSIONS

The length of the head of the main plant, plant height before the first second cutting and dry mass of the seedlings showed high variability. The 36 measured characters express linear genetic dependence, with genetic standards, obtained by Euclidean Algorithm; similar to what was found with Tocher Optimized Clustering and Artificial Neural Networks K-means methods. Similar genetic trends for the heritability profiles were obtained for factor analysis and Artificial Neural Networks by the Kohonem method. The use of Artificial Neural Networks through the Kohonem method had greater efficacy in the definition of the genetic profiles needed for the recommended agronomic ideotype for the improvement of dual-purpose wheat.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Arruda GP, Demattê JAM and Chagas CS (2013). Digital soil mapping by artificial neural networks based on soil-landscape relationships. *Rev. Bras. Ciênc. Solo*. 37(2): 327-338. http://dx.doi.org/10.1590/S0100-06832013000200004
Bertan I, Carvalho FIF, Oliveira AC, Vieira E, et al. (2006). Comparação de métodos de agrupamento na representação da distância morfológica entre genótipos de trigo. *Rev. Bras. Agrociência.* 12(3): 279-286. HTTP://DX.DOI.ORG/10.18539/CAST.V12I3.4554

Carvalho IR, de Souza VQ, Nardino M, Follmann DN, et al. (2015). Correlações canônicas entre caracteres morfológicos e componentes de produção em trigo de duplo propósito. *Pesq. Agropec. Bras.* 50(8): 690-697. http://www.scielo.br/scielo.php?script=sci_abstract&pid=S0100204X2015000800690&lng=en&nrm=iso&tlng=pt

Carvalho IR, Nardino M, Ferrari M, de Pelegrin AJ, et al. (2016). Genetic variability among common black bean (*Phaseolus vulgaris* L.,) accessions in southern Brazil. *Aust J Crop Sci.* 10: 1474-1483. DOI: 10.21475/ajcs.2016.10.10.p7786

Carvalho IR, Szareski VJ, Demari GH, Barbosa MH, et al. (2018). Artificial Neural Network and Multivariate Models Applied to Morphological Traits and Seeds of Common Beans Genotypes. *J. Agric. Sci.* 10: 572-580. 10.5539/jas.v10n11p572

Carvalho IR (2018). Biometria e modelos mistos aplicados à seleção de genótipos de feijão. 2018. 104 f. Monograph (Specialization) - Curso de Pós-graduação em Ciência e Tecnologia de Sementes, UFPEL, Pelotas.

Castro PRC and Klunge RA (1999). Ecofisiologia de cultivos anuais: trigo, milho soja, arroz e mandioca. Nobel, São Paulo. 128 p.

Companhia Nacional de Abastecimento. (2019). Acompanhamento da safra brasileira de grãos, v. 11 Safra 2017/18 - Boletim de grãos Janeiro de 2019, Brasília, 126p. Available at: < http://www.conab.gov.br >. Accessed 2 January, 2019.

Cruz CD, Carneiro PCS and Regazzi AJ (2014). Modelos biométricos Aplicados ao Melhoramento genético. UFV, Viçosa. 668p.

Cruz CD, Regazzi AJ and Carneiro PCS (2012). Modelos Biométricos Aplicados ao Melhoramento Genético. 4.ed. Editora UFV, Viçosa. 514p.

Cruz CD and Regazzi AJ (2001). Modelos biométricos aplicados ao melhoramento genético. UFV. 390p.

Cruz CD (2013). Genes: a software package for analysis in experimental statistics and quantitative genetics. *Acta Sci. Agron.* 35(3): 271-276. http://www.scielo.br/scielo.php?pid=S180786212013000300001&script=sci_abstract&tlng=pt

Dhonde SR, Kute NS, Kanawade DG and Sarode ND (2000). Variability and characters association in wheat (*Triticum aestivum*). *Agric Sci. Dig.* 20(2): 99-101.

Fioreze LS (2011). Comportamento produtivo do trigo em função da densidade de semeadura e da aplicação de reguladores vegetais.Doctoral thesis. UNESP, Botucatu. 86p.

Hallauer ARJB and Miranda F (1988). Quantitative Genetics in Maize Breeding, 2.ed. Iowa State University Press, Ames, 650p.

Hartwig I, Carvalho FIF, Oliveira AC, Vieira E, et al. (2007). Estimativa de coeficientes de correlação e trilha em gerações segregantes de trigo hexaplóide. *Bragantia* 66(2): 203-218. http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0006-87052007000200004

Jana RB, Mohanty BP and Sheng Z (2012). Upscaling soil hydraulic parameters in the Picacho Mountain region using Bayesian Neural Networks. *Trans ASABE*. 55(2): 463- 473.

Johnson RA and Wichern DW (1992). Multivariada aplicada análise estatística. Prentice Hall, New Jersey. 642p.

Kenga R, et al. (2006). Genetic and phenotypic association between yield components in hybrid sorghum (*Sorghum bicolor* (L.) Moench) populations. *Euphytica* 150(3): 319-326.

Kottek M, et al. (2006). World map of the Köppen-Geiger climate classification updated. *Meteorol Z.* 15(3): 259-263.

Lozano CM and Leaden MI (2002). Efecto de reguladores de crecimiento sobre el rendimiento y altura en dos cultivares de trigo. In: Congresso Nacional De Trigo, 5, Simpósio Nacional De Cereales De Siembra Otoño Invernal, 3. INTA, Argentina.

Martin TN, Simionatto CC, Bertoncelli P, Ortiz S, et al. (2010). Fitomorfologia e produção de cultivares de trigo duplo propósito em diferentes manejos de corte e densidades de semeadura. *Cienc. Rural*. 40(8): 1695-1701. http://www.scielo.br/scielo.php?pid=S010384782010000800004&script=sci_abstract&tlng=pt

Merotto Júnior A (1995). Processo de afilhamento e crescimento de raízes de trigo afetados pela resistência do solo. 132p. Doctoral thesis. Programa de Pós-graduação em Agronomia, Universidade Federal do Rio Grande do Sul.

Nascimento M, Peternelli LA, Cruz CD, Nascimento ACC, et al. (2013). Artificial neural networks for adaptability and stability evaluation in alfalfa genotypes. Crop Breed. *Appl. Biotechnol.* 13: 152-156. http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1984-70332013000200008

Oliveira ACLD, Pasqual M, Pio LAS, Lacerda W, et al. (2013). Utilização da modelagem matemática (redes neurais artificiais) na classificação de autotetraploides de bananeira (Musa acuminata COLLA). *Biosci. J.* 29(3): 617-622.

Pariz CM, Ferreira RL, Sá ME, Andreotti M, et al. (2010). Qualidade fisiológica de sementes de Brachiaria e avaliação da produtividade de massa seca, em diferentes sistemas de integração lavoura pecuária sob irrigação. *Pesq. Agropec. Tropical*. 40(3): 330-340.

Santos HG, Jacomine PKT, Anjos LHC, Oliveira VA, et al. (2018). Sistema Brasileiro de classificação de solos. Embrapa, Brasília, DF.

Smanhotto A, Nóbrega LH, Opazo MA and Prior M (2006). Características físicas e fisiológicas na qualidade industrial de cultivares e linhagens de trigo e triticale. *Rev. Bras. Eng. Agríc. Ambiental* 10(4) p.867-872.

Soares FC, Robaina AD, Peiter MX and Russi JL (2015). Predição da produtividade da cultura do milho utilizando rede neural artificial. *Ciênc Rural*. 45(11): 1987-1993. http://www.scielo.br/scielo.php?pid=S010384782015001101987&script=sci_abstract&tlng=pt

Sokal RR and Rohlf F (1962). The comparison of dendrograms by objective methods. *Taxon*. 11(2): 33-40.

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

Valério IP, et al. (2008). Desenvolvimento de afilhos e componentes do rendimento em genótipos de trigo sob diferentes densidades de semeadura. *Pesq. agropec. bras.* 43(3): 319-326.

Wright S (1921). Correlation and causation. *J. Agric. Res.* 20: 557-585.

Wright S (1923). Theory of path coefficients. *Genetics.* 8: 239-285.

Zilio M, Peloso JÁ and Mantovani A (2017). Produção de forragem e de grãos de trigo de duplo propósito submetido a diferentes densidades de semeadura, adubação nitrogenada e manejos de corte. *Rev. Ciênc. Agrovet.*16(4): 367-375.