

Selection index as *a priori* information for using artificial neural networks to classify alfalfa genotypes

I.G. Santos¹, C.D. Cruz¹, M. Nascimento² and R.P. Ferreira³

¹ Departamento de Biologia Geral, Universidade Federal de Viçosa, Viçosa, MG, Brasil

² Departamento de Estatística, Universidade Federal de Viçosa, Viçosa, MG, Brasil

³ Embrapa Pecuária Sudeste, São Carlos, SP, Brasil

Corresponding author: I.G. Santos

E-mail: iaraminas@hotmail.com

Genet. Mol. Res. 18 (2): gmr18221

Received November 26, 2018

Accepted May 09, 2019

Published May 15, 2019

DOI <http://dx.doi.org/10.4238/gmr18221>

ABSTRACT. The efficiency of a selection index generally depends on the quality of the variance matrixes, which demands controlled experiments. Using Artificial Neural Networks (ANNs) trained from a selection index is advantageous for selecting genotypes since an ANN has the capacity to classify genotypes in an automated way. We propose the use of ANNs for the selection of alfalfa genotypes, based on a selection index. Data were collected from 77 alfalfa genotypes evaluated based on nine traits from four cuttings. The traits were divided into forage yield and nutritive value groups. In order for the ANNs to learn the classification pattern, the Tai index was used, which allows secondary traits to be included in the index to improve the gains of the main traits. An index was established for each group of traits, and based on the index scores the genotypes were subdivided into four classes (optimal, good, medium, and bad). After testing different topologies, ANNs were established for each index, according to the apparent error rates. The chosen ANNs were efficient in classifying the genotypes since the highest apparent error rate reached 15%, meaning that the ANNs efficiently captured the

data pattern. Considering the ANN classification for both groups of traits, there was a high degree of agreement with the classification obtained from the Tai index, as expected. Even in the cuttings where the ANNs presented the worst performance, their potential to classify alfalfa genotypes was clear, because the wrong classifications were placed in groups close to the correct ones. This ensured that the best genotypes did not run the risk of being discarded, since they would not be classified in the group of bad genotypes. The ANNs that were developed have good potential for use in alfalfa breeding programs.

Key words: Tai index; *Medicago sativa*; Computational intelligence

INTRODUCTION

Alfalfa breeding programs aim to develop cultivars that, in addition to high yields, have a high nutritional value and tolerance to biotic and abiotic stresses (Shi et al., 2017). For this reason, alfalfa breeding is based on a set of relevant traits, requiring the adoption of selection techniques that maximize the accuracy of selection. Because alfalfa is a perennial crop, selection is performed using data generated from cuttings, estimating the dry matter yield and forage nutritive value throughout the different growing seasons. In this way, the data generated in each season of the year are analyzed and selection is carried out in order to obtain gains in the whole set of traits.

Simultaneous selection has been the most widely used strategy in alfalfa breeding (Basigalup and Odorizzi, 2011). Among the many published selection indexes, the one proposed by Taí (1977) is applicable to selection of alfalfa genotypes, since it allows secondary traits of forage yield and nutritive value to be included in the index to improve the gains of the main traits (Cruz et al., 2014). However, the efficiency of this index relies on the establishment of desired gains, which may be difficult for breeders. There is a need for high quality in variance and covariance matrices, which demands strongly controlled experiments. Thus, the analysis of a large amount of information generated in forage breeding programs requires that the researcher analyze the data repeatedly, a task that could be minimized using computational intelligence techniques such as Artificial Neural Networks (ANNs).

Artificial Neural Networks are learning techniques based on models of biological nature inspired by the functioning of the human brain (Norvig and Russel, 2013). Given their nonlinear structure, ANNs can achieve high efficiency in working with complex traits (Galvão et al., 1999) and are able to capture relationships among variables that are not captured by stochastic models and may generalize information to new cases (Mackay, 1994). Moreover, ANNs do not require very detailed information about the physical processes of the modeled systems. Although the use of ANNs in plant breeding is still restricted, they show high efficiency to solve prediction and classification problems (Ji et al., 2007; Gianola et al., 2011; Nascimento et al., 2013).

The use of ANNs in alfalfa breeding programs can take place in a simple and automated way. In the initial breeding stages, it is possible to base selection on subjective criteria, from the experience of good breeders, or to adopt an appropriate

biometric procedure so that the genetic value of each genotype can be inferred. As breeding programs are dynamic, future stages may need such learning to establish a new selection strategy that can aggregate information from qualified breeders, accurate biometric techniques, and information processed within the context of computational intelligence.

To that end, it is enough to have a data set of alfalfa cuttings with information coming from well-conducted experiments that characterize the genotype response in diverse environmental conditions. From this, it is possible to obtain a selection criterion, for example an index, and to perform the genotype classification as to the performance in forage yield and nutritive value. Using an established classification, an ANN can be trained to identify patterns and to ponder the genotype traits in the same way as is done in the index, and from there, to classify the genotype information that will be collected in future cuttings. The great advantage of using a network is that it is effective even for genotype evaluations with incomplete and noisy information, a fact that makes the extreme rigor normally needed for conducting the experiments optional. In addition, statistical analyses of new cuttings would become simpler and automated. Our objective was to adjust an ANN from an established selection index for use in the simultaneous selection of alfalfa genotypes of the alfalfa breeding program of Embrapa Pecuária Sudeste, a Brazilian government agricultural breeding program.

MATERIAL AND METHODS

Data from four alfalfa cuttings considering 77 genotypes from Embrapa Pecuária Sudeste were used. The cuttings were carried out in the months of November 2015, February, May and August 2016. All cuttings were irrigated, except for the fourth one, on August 2016, which was grown under water deficit.

The traits were divided into two groups, one of which included plant height, measured from the ground up to the inflorescence; dry matter yield ($\text{kg}\cdot\text{ha}^{-1}$), obtained by manually cutting the plants at eight to 10 cm above ground when each cultivar reached 10% of flowering or when the basal shoots reached an average height of three to five centimeters; and susceptibility to diseases, determined according to the percentage of leaf area attacked in each plot (Vasconcelos et al., 2010). In addition, nutritive value traits were measured, including crude protein, *in vitro* dry matter digestibility, neutral detergent fiber, acid detergent fiber, lignin, and stem/leaf ratio.

In order for the ANN to learn the genotypes classification pattern, it was necessary to use previous information, in this case the selection index proposed by Taí (1977), which served as a learning basis for the network. An index was established for selection of productive genotypes and another for the genotypes of high nutritive value. Considering the nutritive value traits, crude protein (CP) and *in vitro* dry matter digestibility (IVDMD) were considered as the main ones, and the others as secondary.

Among the nutritive value traits, crude protein and *in vitro* dry matter digestibility were considered as the main traits and the others as secondary. Thus, the genotypic aggregate (H) and the index (I) were given by:

$$H = a_1g_1 + a_2g_2 \quad (Eq. 1)$$

$$I = b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 \quad (Eq. 2)$$

where: a: economic weights; g: genotypic values; x: phenotypic values and b: trait coefficients in the index. The vector of desired gains was established based on a genetic standard deviation for each trait (Cruz et al., 2014). Vector b was estimated from equations (3) and (4):

$$P\hat{b} = Ga \quad (Eq. 3)$$

$$G\hat{b} = \Delta g_d \quad (Eq. 4)$$

where: G: matrix of genotypic covariance between variables; P: matrix of phenotypic covariance between variables; Δg_d : vector of desired gains; replacing equation (4) in (3):

$$a = GP^{-1}G\Delta g_d \quad (Eq. 5)$$

Considering:

$$a = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \text{ and } \Delta g_d = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$$

where: c_1 , c_2 , d_1 and d_2 are vectors of dimension 2 x 1. And also considering:

$$c_2 = \Phi \omega = \begin{bmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{bmatrix}$$

where: ω_{11} , ω_{12} , ω_{21} and ω_{22} are 2 x 2 matrices. Therefore:

$$a = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = \begin{bmatrix} \omega_{11}d_1 & \omega_{12}d_2 \\ \omega_{21}d_1 & \omega_{22}d_2 \end{bmatrix} \text{ and}$$

$$\Delta g_d = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = \Delta g_d = \begin{bmatrix} d_1 \\ -\omega_{22}^{-1}\omega_{21}d_1 \end{bmatrix}$$

Thus, vector b was estimated by (3) or (4). Similar reasoning can be made for the forage yield group, in which dry matter yield was considered as the main trait and the others as secondary.

After obtaining the scores, the genotypes were sorted in descending order according to the score obtained by the index. From this, four subgroups for forage yield and four for nutritive value were determined, where the genotypes were classified as optimal, good, medium and bad (Table 1).

Table 1. Classification of the 77 alfalfa genotypes according to the Tai index, regarding performance in nutritive value and forage yield. Numbers 1, 2, 3, and 4 classifies each genotype into optimal, good, medium, and bad classes, respectively, considering each cutting.

Genotype	Nutritive Value				Yield			
	Cut 1	Cut 2	Cut 3	Cut 4	Cut 1	Cut 2	Cut 3	Cut 4
1	1	1	1	1	1	4	4	4
2	1	4	3	3	3	3	2	1
3	2	1	1	2	4	2	3	3
4	2	1	2	4	1	2	1	3
5	2	1	3	3	3	3	2	1
6	3	4	3	1	3	3	3	2
7	2	3	1	1	4	4	4	4
8	3	4	4	2	3	3	3	1
9	4	2	4	4	2	3	1	2
10	4	2	3	3	2	2	1	1
11	4	1	3	3	2	3	2	2
12	4	3	4	3	1	1	1	3
13	4	3	2	2	4	2	2	1
14	3	4	4	2	1	1	1	2
15	1	1	1	1	4	4	4	4
16	3	2	3	3	4	1	1	2
17	2	1	1	1	1	2	3	2
18	1	1	1	1	4	4	4	4
19	3	2	2	3	2	3	4	3
20	2	2	3	2	4	1	2	2
21	2	3	3	4	2	2	1	1
22	2	2	3	3	3	1	1	2
23	1	1	1	1	2	4	4	4
24	3	2	2	2	2	4	3	3
25	1	3	2	2	2	1	2	1
26	3	4	4	4	4	1	2	1
27	3	2	3	3	3	1	3	2
28	3	3	3	3	4	2	2	1
29	1	3	2	2	3	4	3	2
30	3	2	4	2	4	3	2	2
31	3	4	3	2	4	3	2	2
32	2	2	2	2	1	3	4	4
33	2	2	2	2	3	4	3	3
34	4	3	3	3	2	2	2	1
35	1	1	1	1	4	4	4	4
36	1	3	2	4	3	2	4	3
37	1	4	1	3	2	1	3	4
38	2	3	3	4	3	2	1	3
39	2	4	4	4	1	2	2	3
40	4	4	4	3	1	1	1	1
41	2	3	2	1	2	1	2	2
42	3	3	3	2	1	2	2	2
43	1	4	2	1	3	2	2	3
44	3	2	1	2	1	3	3	4
45	4	3	4	4	3	3	1	3
46	1	4	2	3	4	1	3	3
47	2	1	1	2	2	4	3	4
48	4	3	2	1	1	4	2	3
49	1	1	1	1	4	2	4	4
50	1	1	1	1	3	4	4	4
51	1	1	2	2	3	4	4	3
52	4	3	2	2	4	3	3	2
53	4	2	2	1	2	4	4	4
54	3	2	3	3	1	3	2	2
55	1	1	1	1	4	4	4	4
56	3	1	2	4	1	1	3	2

Genotype	Nutritive Value				Yield			
	Cut 1	Cut 2	Cut 3	Cut 4	Cut 1	Cut 2	Cut 3	Cut 4
57	2	3	3	4	1	3	3	1
58	2	1	1	1	3	4	4	3
59	1	1	1	1	2	3	4	3
60	4	2	4	4	1	2	1	1
61	3	3	4	4	1	1	1	1
62	4	4	4	4	4	3	1	1
63	4	4	3	4	3	1	3	3
64	3	4	4	4	1	1	1	2
65	3	1	1	2	3	4	3	3
66	4	4	4	4	2	1	1	1
67	1	2	2	4	1	2	2	4
68	4	2	1	1	1	3	4	4
69	1	2	1	1	4	4	4	4
70	2	1	1	1	1	4	4	4
71	1	3	2	2	3	2	2	4
72	2	4	4	3	2	2	1	1
73	4	3	4	3	3	3	3	3
74	2	2	3	4	4	1	3	2
75	4	4	4	3	2	2	1	1
76	3	4	4	3	2	1	1	1
77	4	4	4	4	2	1	1	1

In order to obtain an ANN with good generalization capacity, the forage yield and nutritive value data, considering the classification according to the Tai index (1977) in each cutting, were amplified. Using this method, the original data of each index class generated 50 new observations, resulting in 800 new data points with yield information and 800 with nutritive value information (Table 2).

Table 2. Amplification scheme of alfalfa data for the selection of the best Artificial Neural Networks for forage yield and nutritive value. Considering each cutting, each class of genotypes obtained by the Tai index (optimal, good, medium and bad) generated 50 new observations.

Group	Cut	Classification	Amplification	Group	Cut	Classification	Amplification
Forage Yield	C1	1	50	Nutritive Value	C1	1	50
		2	50			2	50
		3	50			3	50
		4	50			4	50
	C2	1	50		C2	1	50
		2	50			2	50
		3	50			3	50
		4	50			4	50
	C3	1	50		C3	1	50
		2	50			2	50
		3	50			3	50
		4	50			4	50
	C4	1	50		C4	1	50
		2	50			2	50
		3	50			3	50
		4	50			4	50

The amplification process consisted of generating a random variable with normal distribution according to Silva et al. (2014). The values were simulated, conserving the means, variances and covariances of the original data of each cutting separately. Thus, the

data set was subdivided into 80% for training and 20% for the validation process. The amplification was carried out using the GENES program (Cruz, 2013).

Artificial Neural Network modeling

Two ANNs were established, one for the forage yield group and another for the nutritive value group. Network conformations were defined from the test of different topology possibilities. In all cases, the architecture used Multilayer Perceptron. The hyperbolic tangent function (tansig) and the logistic function (logsig) were tested. For the output layer, the threshold function (purelin) was used. In addition, the trainbr - Backpropagation training algorithm with a number of epochs equal to 7000 was used. The input layer of each network corresponded to the number of traits of each group.

The number of hidden layers tested ranged from one to four and the number of neurons in each layer ranged from three to six neurons. The output layer was composed of a single neuron, represented by a vector of known elements since the learning process in this case was supervised.

After testing the various topology possibilities, the networks that showed the lowest apparent error rates were chosen. The whole process of training and validation was performed using information generated in the amplification procedure. After the training and validation steps, ANNs were used to classify the real data of the four cuttings evaluated, simulating their practical use in breeding programs. The ANNs were processed using Matlab software (Mathworks, 2012) from GENES application scripts (Cruz, 2013).

RESULTS AND DISCUSSION

ANN topology for groups of traits

The best ANN topology for each group of traits was chosen considering 5832 possibilities, adopting as a criterion the apparent error rate. Of the 800 genotypes generated through data amplification, 640 were used for network training and 160 for validation. Thus, for the forage yield trait group, the topology of lower apparent error rate included six neurons in the first hidden layer, three in the second, and six in the third, and the activation functions were tansig, tansig and logsig, respectively (Figure 1A). For the nutritive value trait group, the topology of lower apparent error rate included six neurons in the first hidden layer, six in the second, and three in the third, and the activation function for all of them was logsig (Figure 1B).

The apparent error rates for the Forage Yield ANN reached 15% in the validation and 12% in the training. For the Nutritive Value ANN the apparent error rate reached 12% in the validation and 7% in the training. Higher error rates for the validation step were observed than in the training and the maximum apparent error rate was 15%. These results are compatible with results reported by other researchers such as Sant'Anna (2015) in classificatory studies, reaffirming the great potential for generalization of neural networks as mentioned by Braga et al. (2011). In addition, the network was more efficient to classify genotypes for nutritive value than for forage yield, even presenting topologies of similar complexity, a fact justified by the characteristics of the data set, more distinguishable in terms of nutritive value than forage yield.

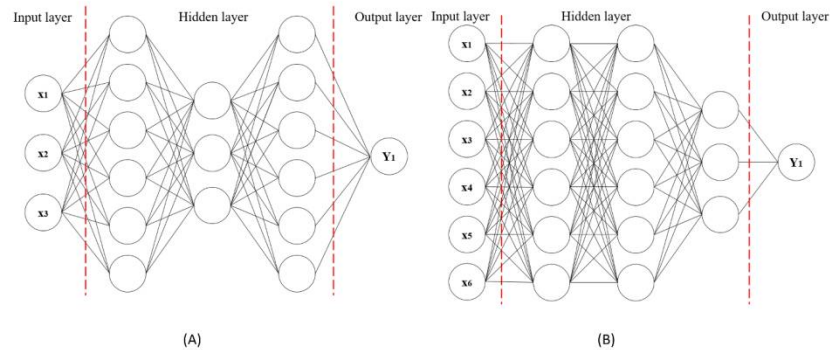


Figure 1. Architectures of Artificial Neural Network (ANN) that best fitted the alfalfa data considering the forage yield (A) and nutritive value traits (B). Inputs (x1) to (x3) and (x1) to (x6) represent each trait for forage yield and nutritive value, respectively. The forage yield network presented three hidden layers consisting of six, three and six neurons, with the activation functions logsig in all layers. The nutritive value network presented six, six and three neurons at each hidden layers, with the activation functions logsig in all layers. In the output layer (Y1), the ANN returns the classification of each genotype.

Efficiency of using artificial neural networks to classify alfalfa genotypes

The ANNs obtained for the forage yield and nutritive value groups were used to classify the real data of the 77 genotypes evaluated in each cutting. Considering the genotype classification for the forage yield traits (Table 3), there was strong agreement with the index, as can be seen, for example, in the first cutting, where the ANN correctly classified all genotypes belonging to group one (optimal genotypes).

Table 3. Classification of the 77 alfalfa genotypes obtained by the Artificial Neural Network for the forage yield traits, considering four cuttings, based on the Tai index. The correct classification occurred when the class of a genotype obtained by the ANN was the same class established by the Tai index.

Cut	Tai index classification	Artificial Neural Network classification			
		1	2	3	4
1	1	20	0	0	0
	2	3	14	2	0
	3	0	4	15	0
	4	0	0	2	16
2	1	17	0	2	1
	2	3	13	1	2
	3	0	6	11	2
	4	1	3	2	13
3	1	10	9	1	0
	2	1	10	7	1
	3	0	3	13	3
	4	0	0	1	18
4	1	11	8	1	0
	2	3	11	4	1
	3	0	2	15	2
	4	0	0	1	18

In the fourth cutting, in which genotypes were kept under water deficit, the ANN correctly classified 18 of the 19 bad genotypes (group four) and 15 of the 19 medium genotypes (group three). Even though the percentage of correct classifications in the optimal and good classes was lower (the ANN correctly classified 11 of the 20 optimal and 11 of the 19 good genotypes), the high accuracy in the bad and medium classes guarantees the selection efficiency by eliminating inferior genotypes. In the other cuttings, accuracy was lower, but the wrong classifications were placed into groups close to the correct ones, as occurred in the third cutting, in which the ANN classified 10 genotypes correctly considering group one, and nine optimal genotypes (which did not enter group one) were classified as good (group two).

There was high concordance between the ANN and the index for nutritive value traits as observed for forage yield (Table 4). In the first cut, the ANN failed to classify two genotypes, and in the second one there were five missing pieces of information. However, this fact did not reduce the generalization power of the ANN in the other scenarios. The accuracy for the group of optimal genotypes for cuttings one and two was high, a fact that was repeated in the other cuttings. In the fourth cutting, the ANN classified correctly all of the optimal genotypes, and most of the good, medium and bad ones. The efficiency of the ANN to discriminate bad genotypes in all cuttings was also high. The worst ANN performance occurred in the third cutting. However, similar to the ANN yield classification, the allocation of the genotypes that had wrong classification occurred in the near classes.

Table 4. Classification of the 77 alfalfa genotypes obtained by the Artificial Neural Network for the nutritive value traits, considering four cuttings, based on the Tai index. The correct classification occurred when the class of a genotype obtained by the ANN was the same class established by the Tai index.

Cut	Tai index classification	Artificial Neural Network classification			
		1	2	3	4
1	1	15	1	2	0
	2	2	12	2	3
	3	0	1	13	5
	4	0	0	1	18
2	1	18	0	0	0
	2	4	10	4	0
	3	0	5	11	2
	4	0	0	2	16
3	1	10	9	1	0
	2	1	10	7	1
	3	0	3	13	3
	4	0	0	1	18
4	1	20	0	0	0
	2	3	15	1	0
	3	0	0	18	1
	4	0	0	2	17

Even in the cuttings where the ANNs showed worse performance (when compared to the others), it was possible to observe their potential to classify alfalfa genotypes, because the wrong classifications were placed in groups close to the correct ones. This ensures that the best genotypes (optimal group) are not at risk of being discarded, since they will not be classified in the group of bad genotypes. Satisfactory performance of ANNs in classificatory analyses has been reported by several researchers, such as Barbosa et al. (2011), who studied eight traits of

papaya (*Carica papaya*); they evaluated the ANN performance compared to an Anderson discriminant analysis, and found ANNs with 94% accuracy, superior to that obtained by a discriminant analysis. A similar conclusion had already been reported by Braga et al. (2011), who recognized the superiority of ANNs in relation to stochastic models. For alfalfa crops, classificatory studies using ANNs were made by Nascimento et al. (2013) and Barroso et al. (2013), who compared the efficiency of adaptability and genotype stability methodologies with ANNs. Yet, as a disadvantage of ANNs in some situations, Barroso et al. (2013) highlight the time spent on data processing and the lack of knowledge about the influence of each trait on the response variable.

Considering that alfalfa breeding programs' aim to develop cultivars with high yields and high nutritive value, the complex genetic control of most of the traits evaluated in this study, and considering that the accuracy of selective process is highly influenced by environmental factors (Li and Brummer, 2012), the use of selection indexes proves to be a very useful tool to carry out selection of alfalfa genotypes. Once the researcher establishes a superior index that considers satisfactory weights of the traits, it is possible to use ANNs to automate the selection process in alfalfa breeding programs. The information obtained *a priori* by the index can be used to train an ANN so that it is able to classify new observations with the efficiency required in the selection process. For this to work, it is crucial that there be efficient training of the network, providing examples that characterize several situations.

For the ANNs developed in this study, the training drew on the supply of examples of cuttings obtained at different times of the year, and this was sufficient for the ANNs to obtain superior performance in all of them. Even if such ANNs are tested in cuttings that were not included in the amplification process, it is expected that the performance in totally unknown cuttings also be high, since the data used for training were amplified from cuttings performed in representative conditions of all seasons of the year, under irrigated conditions and under water deficit.

The ANNs developed in this study showed great potential for use in new alfalfa cuttings. The automated classification will help researchers in their decision making. Besides, due to the dynamism and plasticity of ANNs, they are able to improve their performance when new training examples are presented (Haykin, 2008).

CONCLUSIONS

The Artificial Neural Networks approach is an efficient alternative for classificatory studies. ANNs are highly efficient for use in alfalfa breeding programs.

ACKNOWLEDGMENTS

The authors would like to thank Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Embrapa Pecuária Sudeste. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

REFERENCES

Barbosa CD, Viana AP, Quintall SSR and Pereira MG (2011). Artificial neural network analysis of genetic diversity in *Caricacica papaya* L. *Crop. Breed. Appl. Biotechnol.* 11: 224-231.

- Barroso LMA, Nascimento M, Nascimento ACC, Silva FF, et al. (2013). Uso do método de Eberhart e Russell como informação a priori para aplicação de redes neurais artificiais e análise discriminante visando a classificação de genótipos de alfafa quanto à adaptabilidade e estabilidade fenotípica. *Rev. Bras. Biom.* 31: 176-188.
- Basigalup DH and Odorizzi AS (2011). Melhoramento genético da alfafa. In: Melhoramento Genético da Alfafa (Ferreira RP, Basigalup DH and Gioco JO, eds.). Embrapa Pecuária Sudeste, São Carlos.
- Braga AP, Ferreira ACPL and Ludermir TB (2011). Redes Neurais Artificiais: teoria e aplicações. LTC, Rio de Janeiro.
- Cruz CD (2013). GENES - a software package for analysis in experimental statistics and quantitative genetics. *Acta Sci.* 35: 271-276.
- Cruz CD, Carneiro PCS and Regazzi AJ (2014). Modelos Biométricos Aplicados ao Melhoramento Genético - Vol 2. 3rd ed. Editora UFV, Viçosa.
- Galvão CO, Valença MJS, Vieira VPPB, Lacerda EGM, et al. (1999). Sistemas inteligentes: Aplicações a recursos hídricos e ciências ambientais. UFRGS/ABRH, Porto Alegre.
- Gianola D, Okut H, Weigel KA and Rosa GJM (2011). Predicting complex quantitative traits with Bayesian neural networks: a case study with Jersey cows and wheat. *BMC. Genet.* 12: 87-100.
- Haykin S (2008) Neural Networks and Learning Machines. 3rd edn. McMaster University, Ontario.
- Ji B, Sun Y, Yang S and Wan J (2007). Artificial neural networks for rice yield prediction in mountainous regions. *J. Agric. Sci.* 145: 249-261.
- Li X and Brummer EC (2012). Applied Genetics and Genomics in Alfalfa Breeding. *Agronomy* 2: 40-61.
- Mackay DJC (1994). Bayesian non-linear modelling for the prediction competition. In: Maximum Entropy and Bayesian Methods. Fundamental Theories of Physics (Heidbreder GR, ed.). Springer, Dordrecht.
- Mathworks (2012). MATLAB – The Language of Technical Computing. Available at [<http://www.mathworks.com/products/matlab>].
- 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.
- Norvig P and Russell S (2013). Inteligência Artificial. 3rd edn. Elsevier, Rio de Janeiro.
- Sant'Anna IC, Tomaz RS, GN Silva, Nascimento M, et al. (2015). Superiority of artificial neural networks for a genetic classification procedure. *Genet. Mol. Res.* 14: 9898-906.
- Shi S, Nan L and Smith KF (2017). The Current Status, Problems, and Prospects of Alfalfa (*Medicago sativa* L.) Breeding in China. *Agronomy*. 7: 1-11.
- Silva GN, Tomaz RS, Sant'Anna IC, Nascimento M, et al. (2014). Neural networks for predicting breeding values and genetic gains. *Sci. Agric.* 71: 494-498.
- Tai GCC (1977). Index selection with desired gains. *Crop. Sci.* 17: 182-183.
- Vasconcelos ES, Ferreira RP, Cruz CD, Moreira A, et al. (2010). Estimativas de ganho genético por diferentes critérios de seleção em genótipos de alfafa. *Rev. Ceres.* 57: 205-210.