



DOI: <http://dx.doi.org/10.1590/1807-1929/agriambi.v23n1p9-15>

## Application of artificial neural networks in the prediction of sugarcane juice Pol

Anderson P. Coelho<sup>1</sup>, João V. T. Bettiol<sup>1</sup>, Alexandre B. Dalri<sup>1</sup>,  
João A. Fischer Filho<sup>1</sup>, Rogério T. de Faria<sup>1</sup> & Luiz F. Palaretti<sup>1</sup>

<sup>1</sup> Universidade Estadual Paulista/Faculdade de Ciências Agrárias e Veterinárias, Jaboticabal, SP. E-mail: anderson\_100ssp@hotmail.com (Corresponding author) - ORCID: 0000-0003-2472-9704; joaovictor\_bettiol@hotmail.com - ORCID: 0000-0002-7947-7690; alexandre.dalri@unesp.br - ORCID: 0000-0002-3122-1899; joaofischer16@gmail.com - ORCID: 0000-0001-8787-1631; rogerio.faria@unesp.br - ORCID: 0000-0002-1696-7940; luiz.f.palaretti@unesp.br - ORCID: 0000-0001-5107-6038

**ABSTRACT:** Innovative techniques that seek to minimize the costs of production and the laboriousness of certain operations are one of the great challenges in the sugar-energy sector nowadays. Thus, the objective of the present study was to estimate the Pol values of sugarcane juice as a function of °Brix and wet cake weight (WCW) using artificial neural network (ANN) modeling. A database was organized consisting of 204 technological analyses from a field experiment with 15 treatments and 2 years of evaluation. 75% of the data were used for the calibration of the model and 25% for its validation. Multilayer Perceptron ANNs were used for calibration and validation of the data. Before calibration, the variables were normalized. The training algorithm used was backpropagation and the activation function was the sigmoid. The ANNs were established with two hidden layers and the number of neurons ranging from 4 to 20 in each. The 15 ANNs with the lowest root mean square errors were randomly presented by the software, among which 6 were chosen to verify the accuracy. The ANNs had a high accuracy in the estimation of sugarcane juice Pol, both in the calibration phase ( $R^2 = 0.948$ , RMSE = 0.36%) and in the validation ( $R^2 = 0.878$ , RMSE = 0.41%), and can replace the standard method of analysis. Simpler networks can be trained to have the same accuracy as more complex networks.

**Key words:** trs, Brix, sucrose, technological quality

## Aplicação de redes neurais artificiais na predição do Pol do caldo da cana-de-açúcar

**RESUMO:** Técnicas inovadoras que busquem minimizar os custos de produção e a onerosidade de determinadas operações são um dos grandes desafios atualmente no setor sucroenergético. Nesse sentido, objetivou-se estimar os valores do Pol do caldo da cana-de-açúcar, em função do °Brix e do peso do bolo úmido (PBU), utilizando modelagem por redes neurais artificiais (RNAs). Foi organizado um banco de dados constituído de 204 análises tecnológicas provenientes de um experimento de campo com 15 tratamentos e 2 anos de avaliação. Foram utilizados 75% dos dados para a calibração do modelo e 25% para a validação. Foram utilizadas RNAs do tipo Multilayer Perceptron para calibração e validação dos dados. Antes da calibração, as variáveis foram normalizadas. O algoritmo de treinamento utilizado foi o backpropagation e a função de ativação foi a sigmoide. As RNAs foram estabelecidas com duas camadas ocultas e o número de neurônios variando de 4 a 20 em cada. As 15 RNAs com menor raiz do erro quadrático médio foram apresentadas aleatoriamente pelo software, dentre as quais 6 foram escolhidas para verificação da acurácia. As RNAs apresentaram elevada acurácia para a estimativa do Pol do caldo da cana-de-açúcar, tanto na fase de calibração ( $R^2 = 0,948$ ; RMSE = 0,36%) quanto na validação ( $R^2 = 0,878$ ; RMSE = 0,41%), podendo ser alternativa viável ao método padrão de análise. Redes mais simples podem ser treinadas para apresentarem a mesma acurácia que redes mais complexas.

**Palavras-chave:** atr, Brix, qualidade tecnológica, sacarose



**INTRODUCTION**

The world area planted with sugarcane is  $22 \times 10^6$  ha, with cultivation in 120 countries (FAO, 2016). Brazil is the largest global producer of sugarcane with 661 million tons, representing 20.3% of the global production (USDA, 2017). For being a highly competitive activity and since the price paid for the product is a function of its quality, increased profitability is linked to greater sugar yield (Pereira et al., 2014; Bigaton et al., 2015).

Among the technological analyses in sugarcane, only °Brix, juice Pol and wet cake weight (WCW) are determined, while the other technological attributes are calculated (CONSECANA, 2006). Juice Pol is the most laborious for determination because it requires clarifiers and specific equipment for reading. Alternative methodologies have already been tested to determine Pol (Valderrama et al., 2007; Rodrigues Júnior et al., 2013) however, they either have low accuracy or require specific high-cost equipment.

Modeling in agriculture is used for various purposes, such as estimation of harvest season, yield, soil attributes and quality of products (Demattê et al., 2011; Jones et al., 2016). In this context, the use of artificial neural networks (ANNs) is more accurate than linear and nonlinear models (Wang et al., 2010; Yilmaz & Kaynar, 2011; Castro et al., 2017). This is because ANNs acquire knowledge through experience and are able to recognize patterns and make inferences, imitating the behavior of a human brain (Haykin, 2001).

Thus, this study aimed to estimate the Pol of sugarcane juice through modeling by artificial neural networks, using values of °Brix and WCW as input variables, besides indicating whether the network architecture complexity interferes with its accuracy.

**MATERIAL AND METHODS**

Technological data of a sugarcane experiment conducted in two seasons (2015/2016 and 2016/2017) were used to estimate the juice Pol. The database comprised 204 analyses, and 75% (153) were used for model calibration by ANNs and 25% (51) for validation. The data were composed of 15 treatments, with 5 sugarcane cultivars (CTC 4, IACSP93-3046, RB86-7515, IACSP95-5000 and IAC91-1099) and 3 water managements, showing wide variability of conditions.

The experiment was conducted at the Faculty of Agrarian and Veterinary Sciences -FCAV, UNESP, Campus of

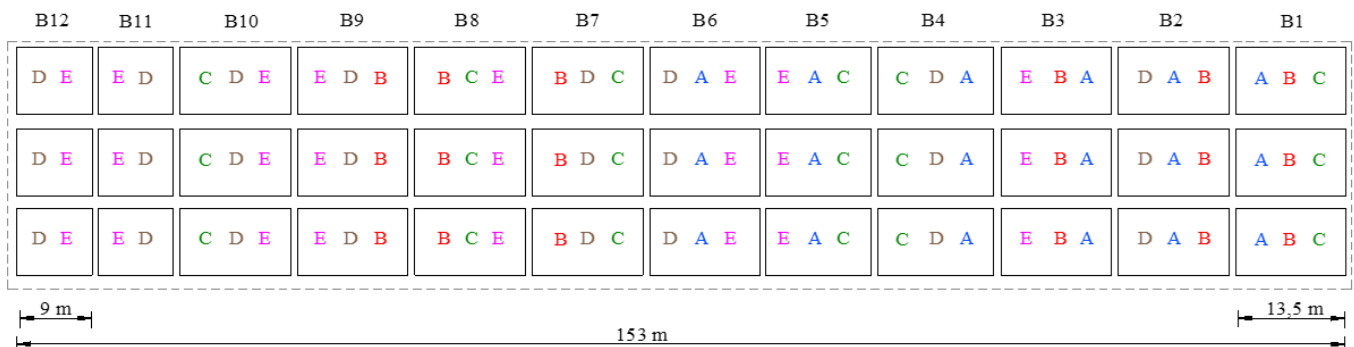
Jaboticabal, SP, Brazil (21° 14' 50" S, 48° 17' 5" W and altitude of 570 m), from May 2015 to May 2016 and from May 2016 to July 2017 (2<sup>nd</sup> and 3<sup>rd</sup> cycles, respectively). The climate of the region is Cwa (according to Köppen's classification), characterized by mean annual rainfall of 1415.7 mm (1975-2015, Agroclimatological Station of FCAV), with mean total of 255 mm for the rainiest month (December) and 25 mm for the driest month (July) (Alvares et al., 2013). The soil in the experimental area is classified as eutroferic Red Latosol (EMBRAPA, 2013).

The experimental design used was partially balanced incomplete blocks (PBIB) in split plots, with three plots and three cultivars per block (Figure 1). The experiment consisted of three factors: supplementary irrigation (100% ETc), deficit irrigation (50% ETc) and non-irrigated (NI), distributed into 12 blocks. These factors were allocated in the plots and the cultivars in the subplots. The PBIB is characterized for being a good option to reduce the magnitude of the experiments. PBIB designs were introduced by Bose & Nair (1939) as a way to evaluate a large number of treatments in small areas.

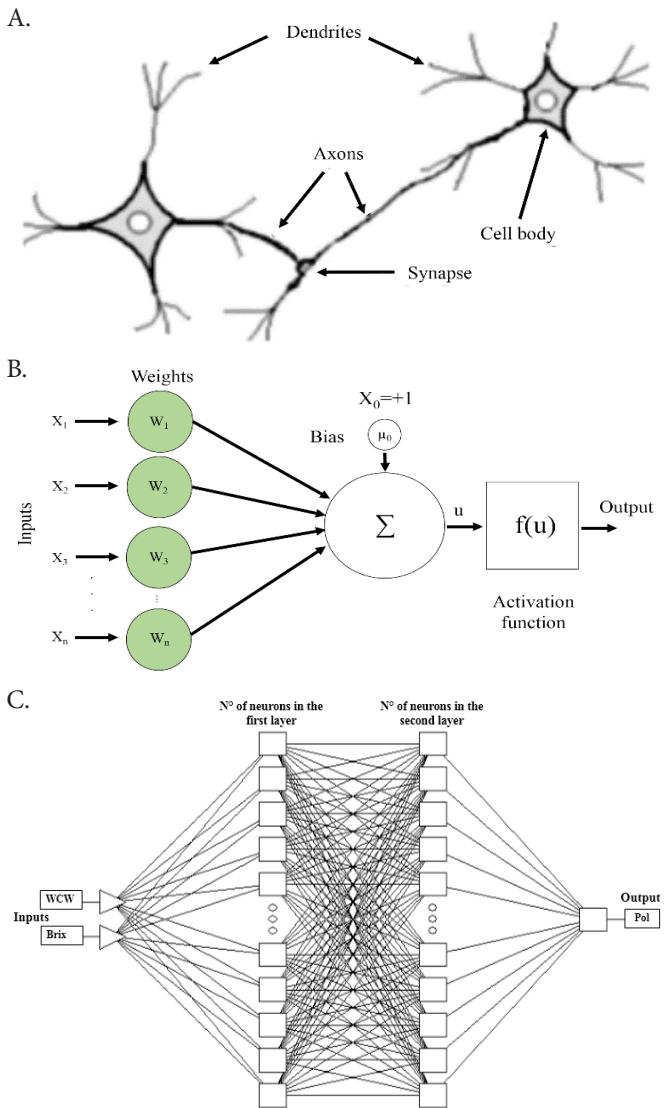
The experimental subplots were composed of four sugarcane rows with 4.5 m length each. Both lateral rows, plus 1 m on each end of the central rows, were considered as borders, and 2.5 m of each central row were used for evaluations. Sugarcane was planted using meristem-grown plantlets, at spacing of 50 cm between plantlets and 1.5 m between rows, which is equivalent to 13,333 plantlets ha<sup>-1</sup>. The second harvest was carried out in May 2016 and the third in July 2017.

The technological quality analysis was performed by harvesting 10 stalks per subplot. The bundles of stalks were properly identified and sent to the laboratory for technological analysis, according to the methodology proposed by CONSECANA (2006). Juice Pol represents the percentage of apparent sucrose contained in a sugar solution of normal weight determined based on the deflection caused by the solution on the vibration plane of the polarized light (Lavanholi, 2010).

Figure 2 shows a representation of a biological neuron (A), artificial neuron (B) and the architecture of artificial neural networks (C). In Figure 2B,  $\Sigma$  represents the sum of the products of inputs by the respective weights, i.e.,  $u = \Sigma x_i W_i$  +bias (the sum of bias provides the increase in degrees of freedom, allowing a better adaptation, by the neural network, to the knowledge presented to it). Lastly,  $f(u)$  is the activation function to obtain the network output.



**Figure 1.** Experimental design in partially balanced incomplete blocks (PBIB). Uppercase letters indicate the cultivars: A - CTC4; B - IACSP93-3046; C - RB86-7515; D - IACSP95-5000 and E - IAC91-1099; B1 - Block 1



$$f(v) = \frac{1}{1 + e^{-av}} \quad (1)$$

where:

- $f(v)$  - sigmoid activation function;
- $a$  - estimate of the parameter which determines the slope of the sigmoid curve; and,
- $v$  - activation potential of the function.

The statistical indices used to evaluate the accuracy of the juice Pol estimation model were the Pearson's linear correlation (Eq. 2), coefficient of determination ( $R^2$ ), root mean square error (RMSE) (Eq. 3), mean error (ME) (Eq. 4), Index of agreement (d) (Willmott, 1981) (Eq. 5) and confidence coefficient (c) (Eq. 6) (Camargo & Sentelhas, 1997). The outliers were removed according to methodology proposed by Belsley et al. (1980).

$$r = \frac{\left( N \times \sum_{i=1}^N (X_{obs_i} \times Y_{obs_i}) \right) - \sum_{i=1}^N X_{obs_i} \times \sum_{i=1}^N Y_{obs_i}}{\sqrt{\left( \left( N \times \sum_{i=1}^N (X_{obs_i})^2 \right) - \left( \sum_{i=1}^N X_{obs_i} \right)^2 \right) \times \left( \left( N \times \sum_{i=1}^N (Y_{obs_i})^2 \right) - \left( \sum_{i=1}^N Y_{obs_i} \right)^2 \right)}} \quad (2)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (Y_{obs_i} - Y_{est_i})^2}{N}} \quad (3)$$

$$ME = \frac{\sum_{i=1}^N (Y_{obs_i} - Y_{est_i})}{N} \quad (4)$$

$$d = 1 - \frac{\sum_{i=1}^N (Y_{obs_i} - Y_{est_i})^2}{\sum_{i=1}^N \left( |Y_{est_i} - \bar{Y}| + |Y_{obs_i} - \bar{Y}| \right)^2} \quad (5)$$

$$c = r \times d \quad (6)$$

where:

- $N$  - number of data;
- $X_{obs_i}$  - observed values of  $X$ ;
- $Y_{obs_i}$  - observed values of  $Y$ ; and,
- $Y_{est_i}$  - estimated values of  $Y$ .

## RESULTS AND DISCUSSION

The data obtained in the calibration and validation phases of the artificial neural networks (ANNs), as well as the statistical indices are presented in Table 1. All models were statistically significant ( $p < 0.01$ ). Descriptive statistics showed that all samples were within the minimum standard of quality required (CONSECANA, 2006).

All models had a confidence coefficient (c) above 0.90, being classified as "very good" (Camargo & Sentelhas, 1997). The average  $R^2$  for the ANNs was 0.948 for the calibration phase (Figure 1) and 0.878 for the validation (Figure 2), thus proving to be accurate to estimate the Pol of sugarcane juice.

**Figure 2.** Neural network: Biological neuron (A), Artificial neuron (B). Illustration of artificial neural network with two input variables (inputs), 2 layers, with the number of neurons ranging from 4 to 20 in each, with the corresponding result (Output) (C)

Juice Pol was estimated through artificial neural networks (ANNs) using the software Statistica 7<sup>o</sup>. The networks were calibrated (trained) in the Intelligent Problem Solver mode. This mode allows one to randomly train several ANNs with different architectures, returning the previously established quantity based on those which had the lowest error. For this, ANNs were established with two hidden layers and 4 to 20 neurons in each layer (Figure 2C), and 15 ANNs were returned by the program. The 6 ANNs with the smallest root mean square errors (RMSE) were selected for analysis, aiming to demonstrate that simpler ANNs can have high accuracy in the prediction.

Multilayer Perceptron ANNs were used for data calibration and validation. The training algorithm used was the backpropagation, which is quite robust for simulating linear and nonlinear models (Haykin, 2001). The criterion used for stopping the calibration algorithm was the total number of cycles equal to 500 or mean squared error lower than 1%. The activation function used was sigmoid (or logistic) (Eq. 1).

**Table 1.** Summary of the statistical analyses carried out for the various models of artificial neural networks and descriptive statistics of the technological data used

Model	Phase	R <sup>2</sup>	RMSE		ME	d	c	F	p-value
			(%)		(%)				
Network 1	Calibration	0.945	0.365		-0.017	0.986	0.958	2489.953**	<0.0001
2-18-05 <sup>1</sup>	Validation	0.878	0.409		-0.030	0.966	0.905	335.984**	<0.0001
Network 2	Calibration	0.945	0.365		-0.018	0.986	0.958	2490.564**	<0.0001
2-20-5	Validation	0.878	0.407		-0.023	0.967	0.906	339.823**	<0.0001
Network 3	Calibration	0.947	0.360		-0.017	0.986	0.959	2562.031**	<0.0001
2-20-13	Validation	0.876	0.412		-0.036	0.966	0.905	333.292**	<0.0001
Network 4	Calibration	0.947	0.361		-0.018	0.986	0.959	2564.646**	<0.0001
2-20-10	Validation	0.874	0.415		-0.022	0.966	0.903	325.715**	<0.0001
Network 5	Calibration	0.949	0.353		-0.021	0.987	0.961	2677.339**	<0.0001
2-20-11	Validation	0.882	0.403		-0.023	0.968	0.910	351.400**	<0.0001
Network 6	Calibration	0.950	0.351		-0.017	0.987	0.962	2710.050**	<0.0001
2-20-15	Validation	0.879	0.410		-0.028	0.968	0.907	340.659**	<0.0001
Attribute	Mean	Minimum	Maximum	Median	Standard error	Standard deviation			
<sup>o</sup> Brix (%)	20.31	16.26	23.10	20.34	0.09	1.23			
WCW (g)	124.73	98.56	147.14	123.94	0.66	9.27			
Pol (%)	17.89	13.32	20.76	17.87	0.11	1.48			

\*\*Significant at 0.01 probability level; <sup>1</sup>Sequence of numbers indicates number of inputs (<sup>o</sup>Brix and WCW), n° of neurons in the first layer and n° of neurons in the second layer; R<sup>2</sup> – Coefficient of determination; RMSE – Root mean square error; ME – Mean error of the estimate; d – Willmott’s index of agreement; c – Confidence coefficient; WCW – Wet cake weight

In the overall mean, RMSE was equal to 0.36% for calibration and to 0.41% for validation. RMSE indicates the actual value of the error produced by the model, being a good indication of model performance, and should be associated with other metrics for better evaluation (Chai & Draxler, 2014). In general, the lower the values obtained for RMSE, the better the performance of the models. The bias of the model (ME) was virtually non-existent, that is, the ANNs do not either underestimate or overestimate the observed values of Pol.

All ANNs had high accuracy, low error and low bias, indicating high accuracy. It can be noted that the selected networks have number of neurons ranging from 18 to 20 in the first layer. However, this number varies from 5 to 15 in the second layer. This indicates that simpler networks, that is, with fewer neurons in the second layer, can be trained and have the same accuracy to predict Pol as more complex networks, since the correlations between the

values estimated by each ANN was almost 1 in all situations (Table 2). In addition, highly complex ANNs may generalize the results and priority should be given to the selection of simpler networks that have high accuracy (Kiani et al., 2010).

The higher the Pol content, the better and greater the amount of sugar obtained by the industry. Juice Pol can be determined both by readings in saccharimeters and using near infrared spectroscopy (CONSECANA, 2006).

However, these methodologies require equipment and clarifications which represent extra cost to the laboratories. In addition, the clarified mixture has aluminum in its composition and may cause environmental contamination as the wastes are disposed off (Silva et al., 2012).

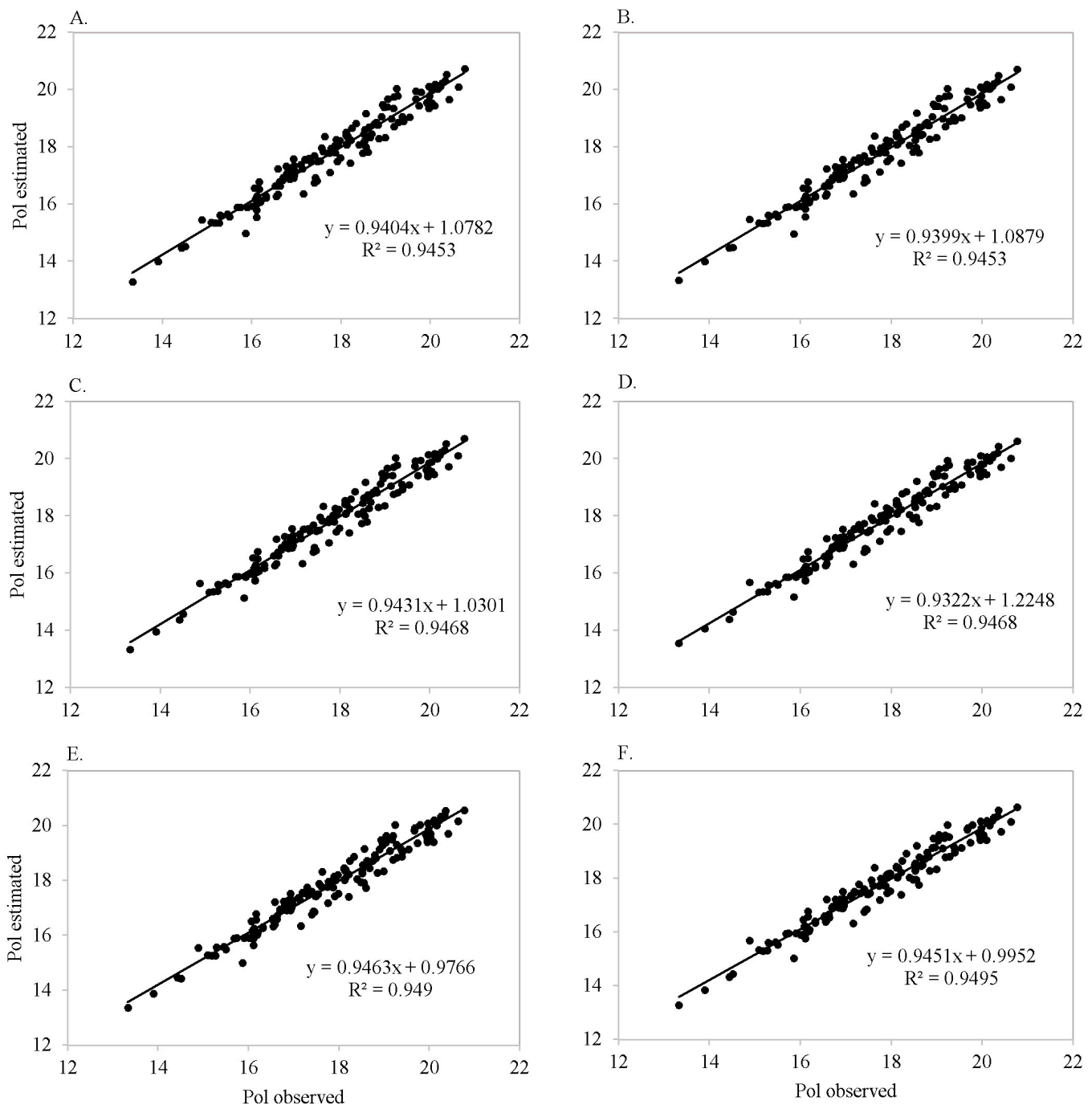
The high accuracy levels obtained in Pol estimation using ANNs (Figures 3 and 4) associated with parameters that are already commonly determined in routine analyses represent

**Table 2.** Correlation matrix between Pol observed and Pol estimated in artificial neural networks (ANNs), correlation between the values estimated by the ANNs, and descriptive statistics of Pol for the calibration and validation phases of the models

	Pol Obs	Pol Est R1	Pol Est R2	Pol Est R3	Pol Est R4	Pol Est R5	Pol Est R6
Pol Obs <sup>1</sup>	1	0.972	0.972	0.973	0.973	0.974	0.974
Pol Est <sup>2</sup> R1 <sup>3</sup>	0.937	1	0.999	0.999	0.999	0.998	0.998
Pol Est R2	0.937	0.999	1	0.999	0.999	0.998	0.998
Pol Est R3	0.936	0.999	0.999	1	0.999	0.998	0.998
Pol Est R4	0.935	0.999	0.999	0.999	1	0.998	0.998
Pol Est R5	0.939	0.999	0.999	0.999	0.998	1	0.999
Pol Est R6	0.937	0.998	0.999	0.999	0.999	0.999	1
Descriptive Statistics - Calibration							
Mean (%)	17.816	17.815	17.817	17.815	17.816	17.819	17.815
Minimum (%)	13.320	13.287	13.336	13.337	13.561	13.386	13.288
Maximum (%)	20.760	20.743	20.710	20.716	20.614	20.575	20.648
Median (%)	17.860	17.818	17.817	17.800	17.858	17.793	17.758
Standard error (%)	0.129	0.125	0.124	0.125	0.123	0.125	0.125
Standard deviation (%)	1.567	1.510	1.509	1.513	1.496	1.517	1.514
Descriptive Statistics - Validation							
Mean (%)	18.121	18.151	18.144	18.157	18.143	18.144	18.148
Minimum (%)	15.230	15.250	15.236	15.330	15.335	15.205	15.232
Maximum (%)	20.630	20.667	20.634	20.647	20.502	20.679	20.634
Median (%)	17.910	18.182	18.170	18.192	18.189	18.156	18.143
Standard error (%)	0.168	0.159	0.159	0.161	0.160	0.164	0.166
Standard deviation (%)	1.176	1.112	1.111	1.128	1.122	1.147	1.161

For the correlation matrix, values below the diagonal line refer to the calibration phase of the ANNs and values above the diagonal refer to the validation phase of the ANNs; <sup>1</sup>Observed values; <sup>2</sup>Estimated values; <sup>3</sup>Number of the artificial neural network





**Figure 3.** Performance graph for Pol observed versus Pol estimated in the calibration phase. ANN 1 (A), ANN 2 (B), ANN 3 (C), ANN 4 (D), ANN 5 (E) and ANN 6 (F)

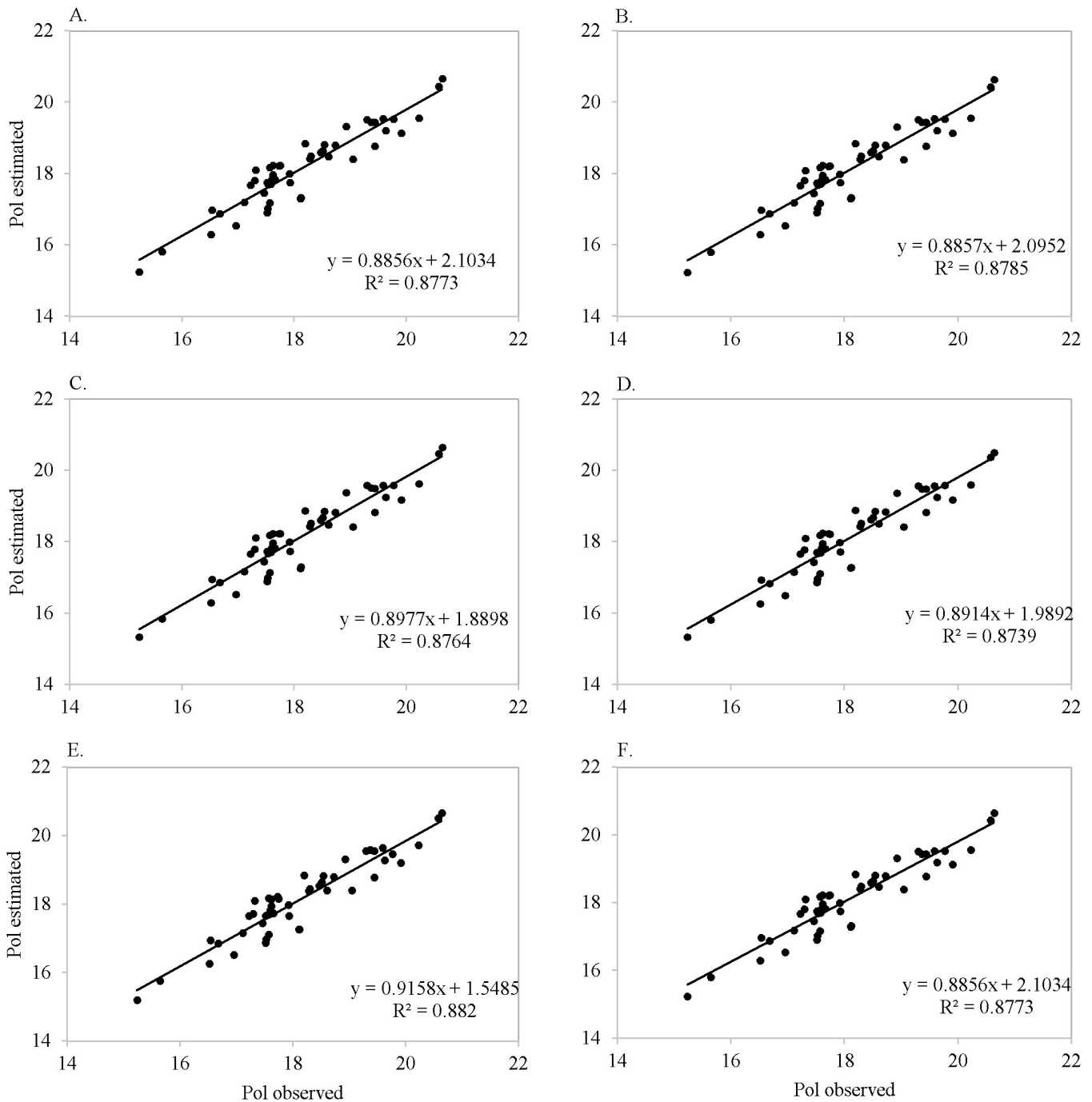
faster procedure and reduced costs for the laboratory. Evaluating the use of infrared spectroscopy in the prediction of sugarcane juice Pol, Nawi et al. (2012) observed  $R^2$  values of 0.78 and 0.71 and RMSE values of 1.89 and 2.11% for calibration and validation phases, respectively. These values indicate less accurate models than the ones obtained by the ANNs in this study. With the same methodology as Nawi et al. (2012), Mat et al. (2014) observed  $R^2$  values of 0.90 for the technological characteristics of sugarcane, an accuracy close to that obtained in this experiment.

According to Nawi et al. (2014), traditional technologies to assess sugarcane quality have limitations, because they require techniques that are costly and of low practicality. In addition, specific devices to determine technological attributes have

a high cost and may cause problems for the plants such as maintenance procedures and failures.

In alternative methodologies to estimate Pol, the accuracy of the model in the calibration phase is often high. However, it becomes little accurate and with high bias in the validation (Melquiades et al., 2012). This is because the samples do not have contrasting conditions and the model ends up generalizing the results. For this study, due to the variability of cultivation conditions, caused by either the number of cultivars, water conditions or data of two years of evaluation, this did not occur because there were high accuracy and low bias of the models by ANNs in the validation phase (Figure 4).

It is worth pointing out that, when the models are used in data sets that were obtained under conditions other than the



**Figure 4.** Performance graph for Pol observed versus Pol estimated in the validation phase. ANN 1 (A), ANN 2 (B), ANN 3 (C), ANN 4 (D), ANN 5 (E) and ANN 6 (F)

original ones, low accuracy levels may be obtained. Thus, it is fundamental to train and validate models for the conditions in which they will actually be used. Since sugar-alcohol plants have a wide range of sugarcane technological data, calibration and validation of models by ANNs can have higher accuracy than that obtained in this study. Therefore, this study aims to indicate an alternative high-accuracy methodology to obtain sugarcane juice Pol by the sugar-alcohol plants, reducing the time and costs with the analyses.

## CONCLUSIONS

1. Artificial neural networks showed high accuracy to estimate sugarcane juice Pol from the °Brix and wet cake

weight (WCW), and can be a viable alternative to the standard method of analysis.

2. The average accuracy of the artificial neural networks was high for both the calibration phase ( $R^2 = 0.948$ ; RMSE = 0.36) and the validation phase ( $R^2 = 0.878$ ; RMSE = 0.41%).

3. Simpler artificial neural networks can be trained and show the same accuracy as more complex networks.

## LITERATURE CITED

Alvares, C. A.; Stape, J. L.; Sentelhas, P. C.; Gonçalves, J. L. de M.; Sparovek, G. Köppen's climate classification map for Brazil. *Meteorologische Zeitschrift*, v.22, p.711-728, 2013. <https://doi.org/10.1127/0941-2948/2013/0507>

- Belsley, D. A.; Kuh, E.; Welsch, R. E. Regression diagnostics: Identifying influential data and sources of collinearity. 1.ed. New York: John Wiley & Sons, 1980. 100p. <https://doi.org/10.1002/0471725153>
- Bigaton, A.; Danelon, A. F.; Torres, H. J. da S.; Xavier, C. E. O.; Marques, P. V. Indicadores técnicos e custos de produção de cana-de-açúcar, açúcar e etanol no Brasil: Fechamento de safra 2013/14. *Revista iPecege*, v.1, p.103-107, 2015.
- Bose, R. C.; Nair, K. R. Partially balanced incomplete block designs. *Sankhyā*, v.4, p.337-372, 1939.
- Camargo, A. P.; Sentelhas, P. C. Avaliação do desempenho de diferentes métodos de estimativas da evapotranspiração potencial no estado de São Paulo, Brasil. *Revista Brasileira de Agrometeorologia*, v.5, p.89-97, 1997.
- Castro, C. A. D. O.; Resende, R. T.; Kuki, K. N.; Carneiro, V. Q.; Marcatti, G. E.; Cruz, C. D.; Motoike, S. Y. High-performance prediction of macauba fruit biomass for agricultural and industrial purposes using Artificial Neural Networks. *Industrial Crops & Products*, v.108, p.806-813, 2017. <https://doi.org/10.1016/j.indcrop.2017.07.031>
- Chai, T.; Draxler, R. R. Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, v.7, p.1247-1250, 2014. <https://doi.org/10.5194/gmd-7-1247-2014>
- CONSECANA - Conselho dos Produtores de Cana-de-Açúcar, Açúcar e Álcool do Estado de São Paulo. Manual de instruções. 5.ed. Piracicaba: CONSECANA, 2006. 112p.
- Demattê, J. A. M.; Bortoletto, M. A. M.; Vasques, G. M.; Rizzo, R. Quantificação de matéria orgânica do solo através de modelos matemáticos utilizando colorimetria no sistema Munsell de cores. *Bragantia*, v.70, p.590-597, 2011. <https://doi.org/10.1590/S0006-87052011005000006>
- EMBRAPA – Empresa Brasileira de Pesquisa Agropecuária. Sistema brasileiro de classificação de solos. 3.ed. Brasília: Embrapa Informação Tecnológica, 2013. 353p.
- FAO - Food and Agriculture Organization of the United Nations. Crops. 2016. Available on: < <http://www.fao.org/faostat/en/#data/QC>>. Access on: Dec. 2017.
- Haykin, S. *Networks neurais: Princípios e prática*. 2.ed. Porto Alegre: Bookma, 2001. 900p.
- Jones, J. W.; Antle, J. M.; Basso, B.; Boote, K. J.; Conant, R. T.; Foster, I.; Keating, B. A. Brief history of agricultural systems modeling. *Agricultural Systems*, v.155, p.240-254, 2016. <https://doi.org/10.1016/j.agsy.2016.05.014>
- Kiani, M. K. D.; Ghobadian, B.; Tavakoli, T.; Nikbakht, A. M.; Najafi, G. Application of artificial neural networks for the prediction of performance and exhaust emissions in SI engine using ethanol-gasoline blends. *Energy*, v.35, p.65-69, 2010. <https://doi.org/10.1016/j.energy.2009.08.034>
- Lavanholi, M. D. G. D. P. Qualidade da cana-de-açúcar como matéria-prima para produção de açúcar e álcool. In: Miranda, L. L. D.; Vasconcelos, A. C. M.; Landell, M. G. D. A. *Cana-de-açúcar*. 1.ed. Campinas: Instituto Agrônomo de Campinas, 2010. Cap.8, p.882-890.
- Mat, N. N.; Rowshon, K. M.; Guangnan, C.; Troy, J. Prediction of sugarcane quality parameters using visible-shortwave near infrared spectroradiometer. *Agriculture and Agricultural Science Procedia*, v.2, p.136-143, 2014. <https://doi.org/10.1016/j.aaspro.2014.11.020>
- Melquiades, F. L.; Bortoletto, G. G.; Marchiori, L. F. S.; Bueno, M. I. M. S. Direct determination of sugar cane quality parameters by X-ray spectrometry and multivariate analysis. *Journal of Agricultural and Food Chemistry*, v.60, p.10755-10761, 2012. <https://doi.org/10.1021/jf302471b>
- Nawi, N. M.; Chen, G.; Jensen, T. In-field measurement and sampling technologies for monitoring quality in the sugarcane industry: A review. *Precision Agriculture*, v.15, p.684-703, 2014. <https://doi.org/10.1007/s11119-014-9362-9>
- Nawi, N. M.; Jensen, T.; Chen, G. The application of spectroscopic methods to predict sugarcane quality based on stalk cross-sectional scanning. *American Society of Sugar Cane Technologists Journal*, v.32, p.16-27, 2012.
- Pereira, G. G. S.; Albrecht, A. J. P.; Fausto, D. A.; Migliavacca, R. A. Custo de produção de cana-de-açúcar no estado do Mato Grosso do Sul. *Revista iPecege*, v.1, p.81-102, 2014.
- Rodrigues Júnior, F. A.; Magalhães, P. S. G.; Franco, H. C. J. Soil attributes and leaf nitrogen estimating sugar cane quality parameters: Brix, pol and fibre. *Precision Agriculture*, v.14, p.270-289, 2013. <https://doi.org/10.1007/s11119-012-9294-1>
- Silva, F. de N.; Teixeira, D. da S.; Paiva, O.; Zioto, P.; Marchioro, S.; Saick, K. W.; Loos, R. Riscos relacionados à intoxicação por alumínio. *Infarma-Ciências Farmacêuticas*, v.24, p.120-124, 2012.
- USDA - United States Department of Agriculture. World agricultural production. 2017. Available on: <<https://apps.fas.usda.gov/psdonline/circulars/production.pdf>>. Access on: Dec. 2017.
- Valderrama, P.; Braga, J. W.; Poppi, R. J. Validation of multivariate calibration models in the determination of sugar cane quality parameters by near infrared spectroscopy. *Journal of the Brazilian Chemical Society*, v.18, p.259-266, 2007. <https://doi.org/10.1590/S0103-50532007000200003>
- Wang, X.; Xia, A.; Wang, J. Determination of brix and POL in sugar cane juice by using near infrared spectroscopy coupled with BP-ANN. *Spectroscopy and Spectral Analysis*, v.30, p.1759-1762, 2010.
- Willmott, C. J. On the validation of models. *Physical Geography*, v.2, p.184-194, 1981. <https://doi.org/10.1080/02723646.1981.10642213>
- Yilmaz, I.; Kaynar, O. Multiple regression, ANN (RBF, MLP) and ANFIS models for prediction of swell potential of clayed soils. *Expert Systems with Applications*, v.38, p.5958-5966, 2011. <https://doi.org/10.1016/j.eswa.2010.11.027>