Received: 1 November 2016 | Revised: 5 June 2017 | Accepted: 13 July 2017

DOI: 10.1111/jfpe.12608

ORIGINAL ARTICLE

WILEY Journal of Food Process Engineering

Artificial neural network model for the kinetics of canola oil extraction for different seed samples and pretreatments

R. J. Sánchez[1] | M. B. Fernández[1,2] | S. M. Nolasco[1,3]

[1]Núcleo TECSE, Facultad de Ingeniería, Universidad Nacional del Centro de la Provincia de Buenos Aires, Olavarría, Argentina

[2]CIFICEN, Universidad Nacional del Centro de la Provincia de Buenos Aires – CONICET-CIC, Tandil, Argentina

[3]CIC, Comisión de Investigaciones Científicas de la Provincia de Buenos Aires, Argentina

Correspondence M. B. Fernández and S. M. Nolasco, Núcleo TECSE, Facultad de Ingeniería, Universidad Nacional del Centro de la Provincia de Buenos Aires, Avda del Valle 5737, B7400JWI Olavarría, Buenos Aires, Argentina.
Emails: mbfernan@fio.unicen.edu.ar; snolasco@fio.unicen.edu.ar

Funding information
ANPCyT, Grant/Award Number: PICT 2012-1442; Universidad Nacional del Centro de la Provincia de Buenos Aires, Argentina, Grant/Award Number: 03-E170

Abstract

In this work, a multi-layer feedforward artificial neural network (ANN) was used for modeling and predicting the oil extraction yields of three canola samples with three pretreatments (*unpretreatment, hydrothermal, and microwave pretreatment*), considering extraction time and temperature as variables. Based on the results of the training, validation, and testing of the network, a neural network with eleven neurons in one hidden layer was selected as the best architecture for predicting the oil extraction yield response. Results obtained by the ANN model were compared with models from the literature (modified Fick's diffusion models), generally obtaining a more accurate fit with the ANN model.

Practical applications

Existing models of canola oil extraction kinetics have some limitations since they are not able to describe various conditions, such as variability among samples and pretreatments. Artificial neural networks (ANN) are powerful and high-precision computational statistical modeling techniques that can address different problems. The aim of this work was to model the kinetics of canola oil extraction under different conditions (varying temperature, samples of canola, pretreatments applied) with an ANN, which presents several advantages over other reported models, allowing to describe a process that depends on many variables even when the data are incomplete or contain errors, thus facilitating its industrial application.

1 | INTRODUCTION

Canola (rapeseed low in erucic acid) is an oilseed with high oil content cultivated worldwide. In Argentina, winter and spring rapeseed are the most common canola varieties. Canola oil is widely consumed worldwide; it is extracted by pressing, solvent extraction, or a combination of the two. The solvent extraction process has been extensively studied (Fernández, Perez, Crapiste, & Nolasco, 2012; Ramos, Sánchez, DE Figueiredo, Nolasco, & Fernández, 2017; Sánchez, Mateo, Fernández, & Nolasco, 2017; Zárate, Perez, Crapiste, Nolasco, & Fernández, 2015). In recent years there has been a growing search for new methods that increase the extraction yield while maintaining oil quality (hydrothermal pretreatments, microwave irradiation, etc.).

Mathematical models derived from Fick's laws of diffusion have been developed to describe the extraction kinetics for seeds with different pretreatments or untreated. Fernández et al. (2012) used a simplified diffusion model for the extraction of canola oil from seeds of the winter variety. The same model was used by Zárate et al. (2015) to study the extraction of oil from untreated and hydrothermally pretreated canola seeds of the spring variety. In both cases, the independent variable was extraction time, maintaining constant temperatures. Sánchez et al. (2017) studied the extraction kinetics of canola oil for untreated and microwave-pretreated seeds of the winter variety, developing a modified Fick's diffusion model as a function of extraction time and temperature. In all the cases, they reported variation in the model parameters, indicating that the oil extraction process is influenced by temperature, time, seed sample, and the existence of pretreatments. However, no model has yet been developed that describes the canola oil extraction taking into account the process variables and the characteristics of the canola and the pretreatments.

Artificial neural networks (ANNs) are sophisticated modeling techniques inspired by the way biological neurons develop learning and memory functions. ANNs present several advantages over conventional modeling techniques since they do not depend on assumptions on the nature of the phenomenological mechanisms, and they are able to learn linear and nonlinear relationships between variables from a set of examples (Fathi, Mohebbi, & Razavi, 2011). The basic units of ANNs are neurons (analogous to biological neurons) and weights (connections that are comparable to the synapses in a biological system). An ANN is able to learn from examples and generalize from simultaneous parallel calculations of its elements, thus allowing to address different issues, even if the data are incomplete or contain errors (Rafiq, Bugmann, & Easterbrook, 2001). The use of this tool to model mass transfer kinetics is recommended by several authors (Baruch, Genina-Soto, Nenkova, & Barrera-Cortés, 2004; Lertworasirikul & Saetan, 2010; Shokri, Hatami, & Khamforoush, 2011). For example, studies have been conducted using ANNs to predict the extraction kinetics of anise essential oil, with the results showing that the proposed model was more accurate than a mathematical model of diffusion (Shokri et al., 2011). The aim of this work was to develop an ANN structure that allows to model and predict accurately the solvent extraction kinetics of canola oil at different temperatures for three seed samples and three pretreatment conditions.

2 | MATERIAL AND METHODS

2.1 | Experimental data

Data of canola oil extractions for three samples (C_0, C_1, C_2) and various pretreatments (*unpretreatment*: grinding; *hydrothermal*: hydrothermal pretreatment and grinding; *microwave*: microwave pretreatment and grinding) corresponding to the works by Fernández et al. (2012), Zárate et al. (2015), and Sánchez et al. (2017) were used (Table 1). In these studies, the authors used a batch apparatus with a magnetic stirrer (200 rpm) for solid–liquid oil extraction with hexane. The experiments were conducted at different times (from 300 to 64,800 s) and at different temperatures (298–333 K) in order to determine the kinetics of oil extraction for all cases. The hydrothermal pretreatment applied to sample C_2 by Zárate et al. (2015) was carried out in an autoclave at 393 K for 5 min, while Sánchez et al. (2017) pretreated sample C_0 in a domestic microwave oven at 607 W for 5 min.

2.2 | Artificial neural network

A fully connected multilayer perceptron (MLP) feed forward neural network was used. The MLP structure is one of the most common types of ANNs (Fathi et al., 2011; Rafiq et al., 2001; Ramzi, Kashaninejad, Salehi, Mahoonak, & Razavi, 2015; Shokri et al., 2011), and it consists of one or more inputs representing independent variables, one output layer with neurons representing the dependent variables, and one or more hidden layers (Hagan, Demuth, Beale, & DE Jesús, 1996) that

TABLE 1 Characterization of the canola samples (Abas et al., 2013)

	C_0	C_1	C_2
Moisture (%db)	5.7 ± 0.2	8.1 ± 0.1	7.4 ± 0.2
Oil (%db)	46.3 ± 0.3	45.2 ± 0.9	40.6 ± 0.6
Protein (%db)	20.3 ± 0.1	18.7 ± 0.6	25.9 ± 0.9

C_0: Canola used by Sánchez et al. (2017). C_1: Canola used by Fernández et al. (2012). C_2: Canola used by Zárate et al. (2015). %db: Dry basis.

contain neurons to capture the nonlinearity of the system. The complexity of the network depends on the number of layers and the number of neurons in each layer. The selection of appropriate network architecture with optimum number of neurons in the hidden layers is an important factor because it's effects upon the network convergence as well as on the accuracy of estimations. The hidden layers connect inputs x to outputs y through a series of weights w interconnected mathematically by Equation 1 (Shokri et al., 2011):

$$y_i = f\left(\sum_{j=1}^{n} w_{ij} x_i + b_i\right) \tag{1}$$

where w_{ij} is the weight of the ith component of the input vector which is connected to the jth neuron; n is the number of inputs to the neuron; b_i is the bias associated with the jth neuron, adding an extra variable, which can make it more powerful than a network without thresholds (Hagan et al., 1996); and f is the activation function that gives the nonlinear behavior of the neuron. The activation function may be linear or nonlinear, depending on the network topology. In this work, the categories canola sample (x_1) and pretreatment (x_2), and the variables temperature (x_3) and extraction time (x_4) were used as input data, while the oil yield relative to the yield at infinite time (64,800 s) for each experiment was considered as the output data (z), thus constituting the input and output vectors as shown in equation 2.

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, \quad Z = [z] \tag{2}$$

For the categories canola sample (x_1) and pretreatment (x_2), numbers from 0 to 2 were assigned to each condition, as shown in Table 2.

In order to ensure a successful model, two important factors must be considered: the number of layers, and the number of neurons in each hidden layer. Since many practical problems in neural modeling could be solved with a hidden layer (Rafiq et al., 2001), an ANN with two layers (a hidden layer and an output layer) was used with a backpropagation algorithm (BP) as supervised training algorithm because it provides quick adjustments and is easily applicable (Shokri et al., 2011). The weights were adjusted using the Levenberg-Marquardt algorithm, which converges faster than gradient-descent methods (Hagan et al., 1996). In this technique, the data set is divided into three subsets: training, validation, and testing. The training and validation subsets were used to adjust the model parameters, while the testing subset was used to evaluate its predictive power.

Of the total dataset of 132 values, 88, 22, and 22 were used to construct the training, validation, and testing subsets, respectively. The modified Jenkin's method (Rafiq et al., 2001) was applied to each category to select the training data, those selections combined were used to train the network, and the remaining data was randomly divided to form the validation and testing sets. The training was conducted for up to 1000 epochs or until the error of the validation data reached a minimum, thus avoiding over training the network (Hagan et al., 1996). The optimal number of neurons in

TABLE 2 Inputs of the ANN

Canola sample[a]	0	0	1	2	2
Pretreatment[b]	0	2	0	0	1
Temperature range (K)	298–333	313–333	313–333	298–333	298–333
Time range (s)	300–64,800	300–64,800	300–64,800	300–64,800	300–64,800

[a]0:C_0; 1: C_1; 2: C_2.
[b]0: Unpretreated; 1: Hydrothermal; 2: Microwave.

the hidden layer was determined by a process of trial and error (Fathi et al., 2011; Hernandez-Perez, García-Alvarado, Trystram, & Heyd, 2004; Lertworasirikul & Saetan, 2010; Ramzi et al., 2015), minimizing the value of MSE (mean squared error) for the validation data. The activation functions "tansig" and "purelin" of MATLAB were selected for the hidden layer and the output layer, respectively. These functions have also been used by other authors (Shokri et al., 2011). They are described in Equations 3 and 4.

$$\mathrm{tansig}(x) = \frac{2}{1+e^{-2x}} - 1 \quad (3)$$

$$\mathrm{purelin}(x) = x \quad (4)$$

In order to ensure a good prediction capability (goodness of fit and accuracy), it is recommended to correlate the two performance parameters, high quadratic coefficient of determination (R^2) to prove a good fitness and minimum root mean square error (RMSE) to prove the accuracy. The RMSE and the quadratic coefficient of determination (R^2) were used to assess the accuracy and goodness of fitness of the models respectively (Fathi et al., 2011). They were calculated using Equations 5 and 6:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}\left(z_i^{\mathrm{exp}} - z_i^{\mathrm{pred}}\right)^2}{N}} \quad (5)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(z_i^{\mathrm{exp}} - z_i^{\mathrm{pred}}\right)^2}{\sum_{i=1}^{N}\left(z_i^{\mathrm{exp}} - \bar{z}^{\mathrm{mexp}}\right)^2} \quad (6)$$

where z_i^{exp} represents the value of the experimental response for the ith data, z_i^{pred} represents the predicted value for the ith data, $\bar{z}^{\mathrm{mexp}}$ is the value of the average response, and N is the total number of data points. For data processing and the design and training of the ANN, the MATLAB software was used.

2.2.1 | Data pre-processing

It is possible to achieve greater efficiency in the training of the neural network by pre-processing the input and output data before entering them into the network. Since the way the data is presented to the network affects the learning process, normalizing the input and output

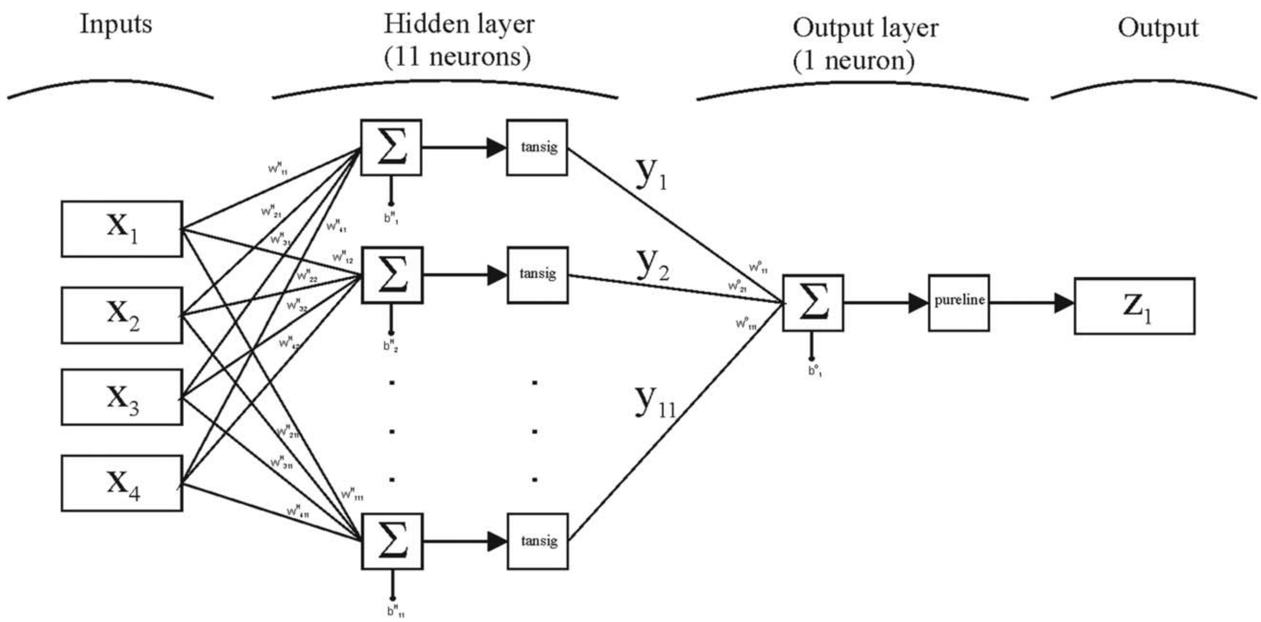

FIGURE 1 Schematic diagram of the optimal multilayer feedforward ANN model. x_i: ith input; w_{hij}: weight corresponding to the ith input of the jth neuron in the hidden layer; b_{hj}: bias of the jth neuron in the hidden layer; w_{ojk}: weight corresponding to jth input of the kth neuron in the output layer; b_{ok}: bias of the kth neuron in the output layer; tansig: tangent sigmoid function; purelin: linear activation function; y_j: output of j neuron in the output layer; z: output of ANN

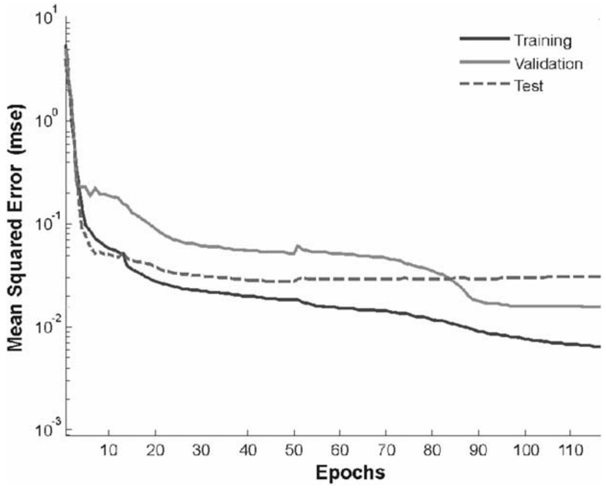

FIGURE 2 Early stopping criteria and the trends of the mean squared error for the training, validation and testing subsets as a function of epochs

data before submitting them to the network is recommended (Rafiq et al., 2001). In the present work, the data were normalized for each input and output to zero mean and unit standard deviation using the mapstd function of the MATLAB software.

3 | RESULTS AND DISCUSSION

3.1 | Optimal ANN

The ANN with 11 neurons in the hidden layer gave the best results. Figure 1 shows a schematic diagram of the architecture of the adopted network.

Figure 2 shows the evolution of the MSE of the network during training, starting at a high value, and decreasing along the process. Three curves corresponding to the training, validation, and test sets are presented in the graph. When the validation error reached a minimum, the training process was stopped, and the weights and biases were obtained.

The weights (w) and biases (b) for each neuron are reported in Table 3. The hidden layer contains 44 weights and 11 biases, while the output layer contains 11 weights and 1 bias.

A comparison between the experimental data and the values predicted by the network is presented in Figure 3. The results show a correlation straight line with a unit linear coefficient and independent terms equal to zero for all cases, except for sample C_2, which presents a linear coefficient value of 0.99 for the *unpretreated* sample, and a linear coefficient value of 0.96 and an independent term equal to 0.05 for the *hydrothermally* pretreated sample. In turn, the correlation between the experimental and the predicted values for the testing subset are shown in Figure 4. A correlation coefficient of 0.971 can be observed, with a RMSE of 0.0356.

For all cases R^2 was higher than 0.950, indicating a good fitness of the model, and for all cases RMSE was lower than 0.0500 (Figures 3 and 4), which is the maximum coefficient of variation of the experimental data (5%, Fernández et al., 2012; Sánchez et al., 2017; Zárate et al., 2015) indicating high accuracy of the models.

In general, a good fit of the experimental data was obtained with this model for all the experiments, indicating that the ANN model developed in this work could adequately predict the extraction kinetics of canola oil with hexane for the range of variables studied.

Table 4 presents a comparison between the ANN model and the modified Fick's diffusion models (DMs) for each experiment. In general, a more accurate fit provided by the ANN model with respect to the corresponding DMs is observed. These results are similar to those

TABLE 3 Parameters of the optimal ANN

Hidden layer parameters						Output layer parameters	
						b	−2.1780
w						*w*	
Neuron	Canola sample	Pretreatment	Temperature	Time	*b*	Neuron	*w*
1	−1.6334	0.1902	−0.7136	0.4086	4.4449	1	1.0714
2	0.7717	−3.2816	0.2659	2.5444	−2.2438	2	−2.5873
3	−2.6158	−3.2858	−0.2524	−1.0787	−0.8224	3	−2.1501
4	0.6275	1.0305	−0.2429	1.9419	2.6097	4	−2.6155
5	2.9243	−2.3986	−0.2912	7.4568	5.5626	5	2.0292
6	1.0417	3.2865	1.9430	−1.9068	0.1845	6	0.2422
7	2.3303	6.8163	0.0243	−3.3845	0.8097	7	−5.5820
8	3.9305	2.6794	−0.2965	0.1054	0.1791	8	2.4594
9	7.1583	−4.3185	0.1781	−8.2811	2.2701	9	−4.4016
10	−1.6833	4.5054	−0.1672	6.0815	8.9030	10	2.5599
11	−1.2818	−1.5794	−0.0946	2.3795	−1.0210	11	−2.6021

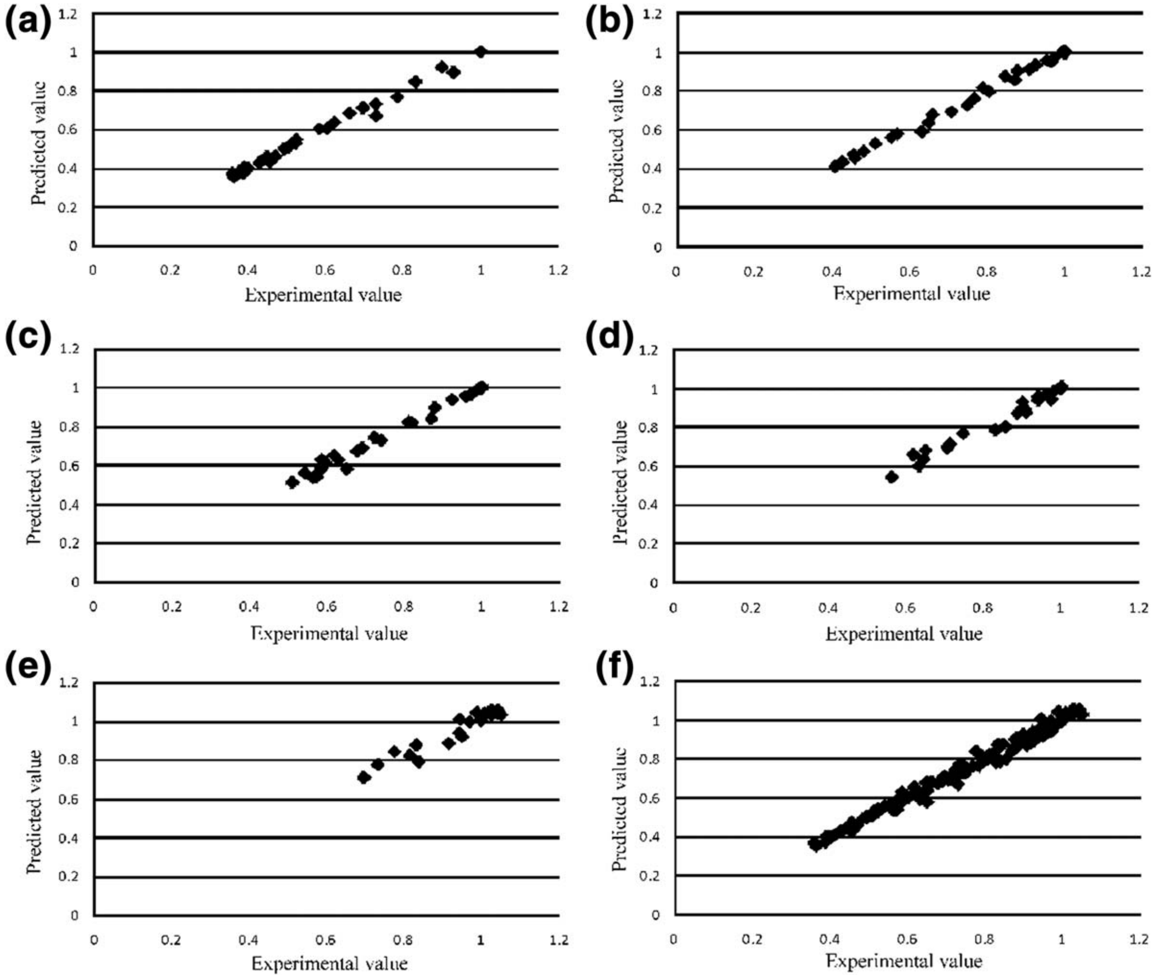

FIGURE 3 Experimental data and values predicted by ANN. *m*: linear coefficient. *b*: independent term. (a) Unpretreated sample C_0. $m = 1$, $b = 0$, $R^2 = 0.993$; RMSE = 0.0181. (b) microwave-pretreated sample C_0. $m = 1$, $b = 0$, $R^2 = 0.993$; RMSE = 0.0157. (c) Unpretreated sample C_1. $m = 1$, $b = 0$, $R^2 = 0.984$; RMSE = 0.0218. (d) Unpretreated sample C_2. $m = 0.99$, $b = 0$, $R^2 = 0.973$; RMSE = 0.0248. (e) Hydrothermally pretreated sample C_2. $m = 0.96$, $b = 0.05$, $R^2 = 0.923$; RMSE = 0.0301. (f) Total data. $m = 1$, $b = 0$, $R^2 = 0.988$; RMSE = 0.0222

reported by other authors who modeled the mass transfer during the dehydration of kaffir lime peel (Lertworasirikul & Saetan, 2010) and the near critical carbon dioxide extraction of anise (Shokri et al., 2011) by means of ANN and mathematical models, obtaining a better accuracy with ANN than with the mathematical models. It is important to note that the ANN allows to adjust the data for all the studied conditions, capturing the relationship between the variables and the behavior of the phenomena with a single model, even processing incomplete data, which would facilitate its potential industrial application.

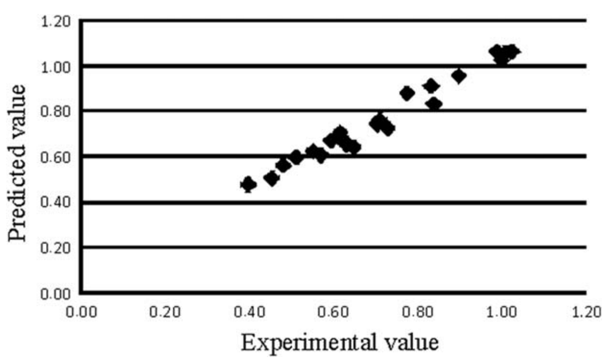

FIGURE 4 Experimental data and values predicted by ANN for test data. *m*: linear coefficient. *b*: Independent term. $m = 1.051$, $b = 0.045$, $R^2 = 0.971$; RMSE = 0.0356

3.2 | Data simulated using the ANN and DM models

In order to determine the correspondence between the ANN and DM models, responses (z_1) for both models were simulated for randomly selected variables within their range of validity (Table 2). The results are shown in Table 5.

It can be observed that the maximum difference between the values predicted by both models is 0.07, and that the coefficients of variation between them do not exceed 5.6%. Although the ANN does not require any assumptions on the nature of the phenomenological mechanisms, it showed good correspondence with the modified Fick's diffusion model (DM).

TABLE 4 Comparison of RMSE and R^2 values between the ANN and DM models for different operating variables

			ANN		Modified Fick's diffusion model (DM)		
Canola sample	Pretreatment	T(K)	R^2	RMSE	Model	R^2	RMSE
C_0 (Sánchez et al., 2017)	Unpretreated	298–333	0.993	0.0181	$z(x_3,x_4)=1-0.462\sum_{n=1}^{8}\frac{1}{n^2}e^{-\left(0.233e^{-2571.6\frac{1}{x_3}}\right)n^2x_4}$	0.990	0.0207
	Microwave	298–333	0.994	0.0157	$z(x_3,x_4)=1-0.450\sum_{n=1}^{8}\frac{1}{n^2}e^{-\left(0.677e^{-2611.3\frac{1}{x_3}}\right)n^2x_4}$	0.992	0.0174
C_1 (Fernández et al., 2012)	Unpretreated	298	0.997	0.0132	$z(x_4)=1-0.464e^{-0.000100x_4}$	0.977	0.0487
	Unpretreated	313	0.964	0.0337	$z(x_4)=1-0.464e^{-0.000131x_4}$	0.952	0.0330
	Unpretreated	323	0.990	0.0173	$z(x_4)=1-0.464e^{-0.000160x_4}$	0.965	0.0176
	Unpretreated	333	0.991	0.0151	$z(x_4)=1-0.464e^{-0.000246x_4}$	0.950	0.0237
C_2 (Zárate et al., 2015)	Unpretreated	313	0.978	0.0280	$z(x_4)=1-0.455e^{-0.000237x_4}$	0.940	0.0275
	Unpretreated	323	0.979	0.0186	$z(x_4)=1-0.455e^{-0.000416x_4}$	0.930	0.0344
	Unpretreated	333	0.969	0.0253	$z(x_4)=1-0.455e^{-0.000559x_4}$	0.910	0.0222
	Hydrothermal	313–333	0.923	0.0301	$z(x_4)=1-0.304e^{-0.000734x_4}$	0.850	0.0324

z: relative oil extraction yield; x_3: temperature (K); x_4: extraction time (s).

4 | CONCLUSIONS

In this work, the kinetics of canola oil extraction with hexane at different temperatures (298–333 K) for three samples of canola seeds and different pretreatments (*unpretreatment*, *microwave*, *hydrothermal*) was modeled using an artificial neural network. The data was divided into three subsets: 2/3 for training and to adjust the model parameters, 1/6 for validation, to avoid overtraining the network, and 1/6 for testing the model. The best result was obtained with a neural network with a hidden layer and 11 neurons in the hidden layer, using the Levenberg-Marquardt algorithm to adjust the parameters, with the activation functions "tansig" for the hidden layer and "purelin" for the output layer. The high correlation coefficients (R^2) between the values predicted by the network and the corresponding experimental data, as well as the low values of the root mean square error (RMSE) demonstrate the feasibility of using the ANN for modeling the canola oil extraction process, presenting in general better accuracy than the DM models. The ANN model presented the advantage of including qualitative variables such as simple origin and pretreatment directly as independent variables, while in DM models these attributes should be studied separately. This unique ANN model represented the oil extraction kinetics within the range of the variables studied, allowing to describe a process that depends on multiple variables from a set of incomplete data. In this context, this tool has great potential at industrial level considering that

TABLE 5 Comparison between the responses simulated using the ANN model and modified diffusive models (DM) for different operating variables

Canola sample	Pretreatment	Temperature (K)	Time (s)	Response z ANN	Response z DM	Difference |ANN-DM|
C_0	Unpretreated	307	4,331	0.59	0.59	0.01
		326	8,363	0.74	0.78	0.04
		341	12,394	0.89	0.87	0.02
	Microwave	311	300	0.43	0.43	0.00
		330	4,313	0.83	0.82	0.01
		341	16,425	0.97	1.00	0.02
C_1	Unpretreated	298	3,000	0.61	0.66	0.05
		313	500	0.55	0.57	0.02
		323	2,000	0.68	0.66	0.02
		333	1,400	0.69	0.67	0.02
		313	6,000	1.00	0.96	0.04
		323	800	0.93	0.90	0.03
		333	13,000	1.06	1.00	0.06
C_2	Unpretreated	313	800	0.59	0.62	0.04
		323	3,000	0.84	0.87	0.03
		333	11,000	0.93	1.00	0.07
	Hydrothermal	323	1,000	0.90	0.85	0.05
		328	64,800	1.01	1.00	0.01
		313	400	0.83	0.80	0.03
		323	15,000	1.03	0.97	0.07
		333	5,000	0.98	0.91	0.07

process data can be collected continuously during the operation of an industrial process.

ACKNOWLEDGMENTS

The authors acknowledge the financial support from ANPCyT (Agencia Nacional de Promoción Científica y Tecnológica) (Grant Number: PICT 2012-1442) and Universidad Nacional del Centro de la Provincia de Buenos Aires, Argentina (Grant Number: 03-E170).

NOMENCLATURE

C_0	Canola sample used by Sánchez et al. (2017).
C_1	Canola sample used by Fernández et al. (2012).
C_2	Canola sample used by Zárate et al. (2015).
y	Output of a neuron in the hidden layer.
n	Number of inputs to a neuron.
f	Activation function.
w	Weight associated with an input.
b	Bias associated with a neuron.
Z	Output vector of the neural network.
z	Oil extraction yield response relative to the yield at infinite time.
X	Input vector of the neural network.
x_1	Canola sample.
x_2	Pretreatment.
x_3	Temperature variable (K).
x_4	Extraction time variable (s).
tansig(x)	Sigmoid tangent activation function.
purelin(x)	Linear activation function.
RMSE	Root mean square error.
MSE	Mean square error.
m	Linear coefficient of the correlation straight line.
b	Independent term of the correlation straight line.

Subscripts

0,1,2,. . ..,n	series terms.
t	at time t.

Superscripts

exp	Experimental data.
pred	Predicted data.
mexp	Mean experimental data

ORCID

M. B. Fernández http://orcid.org/0000-0001-6910-5474

REFERENCES

Abas, F., Khatib, A., Shitan, M., Shaari, K., & Lajis, N. H. (2013). Comparison of partial least squares and artificial neural network for the prediction of antioxidant activity in extract of Pegaga (Centella) varieties from 1 H nuclear magnetic resonance spectroscopy. *Food Research International*, *54*, 852–860.

Baruch, I., Genina-Soto, P., Nenkova, B., & Barrera-Cortés, J. (2004). Neural model of osmotic dehydration kinetics of fruits cubes. In C. Bussler & D. Fensel (Eds.), *Artificial Intelligence: Methodology, Systems, and Applications: 11th International Conference, AIMSA 2004, Varna, Bulgaria, September 2–4 2004, Proceedings* (pp. 312–320). Berlin: Springer-Verlag.

Fathi, M., Mohebbi, M., & Razavi, S. M. A. (2011). Application of image analysis and artificial neural network to predict mass transfer kinetics and color changes of osmotically dehydrated kiwifruit. *Food and Bioprocess Technology*, *4*, 1357–1366.

Fernández, M. B., Perez, E. E., Crapiste, G. H., & Nolasco, S. M. (2012). Kinetic study of canola oil and tocopherol extraction: Parameter comparison of nonlinear models. *Journal of Food Engineering*, *111*, 682–689.

Hagan, M. T., Demuth, H. B., Beale, M. H., & DE Jesús, O. (1996). *Neural network design*. Boston: PWS Publishing Company.

Hernandez-Perez, J., García-Alvarado, M., Trystram, G., & Heyd, B. (2004). Neural networks for the heat and mass transfer prediction during drying of cassava and mango. *Innovative Food Science & Emerging Technologies*, *5*, 57–64.

Lertworasirikul, S., & Saetan, S. (2010). Artificial neural network modeling of mass transfer during osmotic dehydration of kaffir lime peel. *Journal of Food Engineering*, *98*, 214–223.

Rafiq, M., Bugmann, G., & Easterbrook, D. (2001). Neural network design for engineering applications. *Computers & Structures*, *79*, 1541–1552.

Ramos, L. B., Sánchez, R. J., DE Figueiredo, A. K., Nolasco, S. M., & Fernández, M. B. (2017). Optimization of microwave pretreatment variables for canola oil extraction. *Journal of Food Process Engineering*, *40*, e12431.

Ramzi, M., Kashaninejad, M., Salehi, F., Mahoonak, A. R. S., & Razavi, S. M. A. (2015). Modeling of rheological behavior of honey using genetic algorithm–artificial neural network and adaptive neuro-fuzzy inference system. *Food Bioscience*, *9*, 60–67.

Sánchez, R., Mateo, C., Fernández, M., & Nolasco, S. (2017). Bidimensional modeling applied to oil extraction kinetics of microwave-pretreated canola seeds. *Journal of Food Engineering*, *192*, 28–35.

Shokri, A., Hatami, T., & Khamforoush, M. (2011). Near critical carbon dioxide extraction of Anise (Pimpinella Anisum L.) seed: Mathematical and artificial neural network modeling. *The Journal of Supercritical Fluids*, *58*, 49–57.

Zárate, V., Perez, E. E., Crapiste, G. H., Nolasco, S. M., & Fernández, M. B. (2015). Oil extraction kinetics of hydrothermally pretreated canola seeds. *The Canadian Journal of Chemical Engineering*, *93*, 842–848.

How to cite this article: Sánchez RJ, Fernández MB, Nolasco SM. Artificial neural network model for the kinetics of canola oil extraction for different seed samples and pretreatments. *J Food Process Eng*. 2017;e12608. https://doi.org/10.1111/jfpe.12608